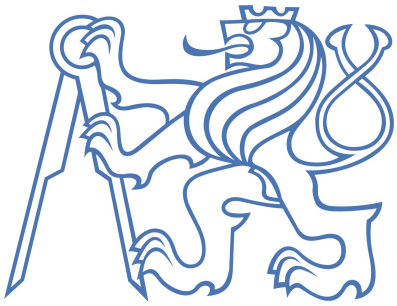
CZECH TECHNICAL UNIVERSITY
IN PRAGUE
FACULTY OF CIVIL ENGINEERING

DEPARTMENT OF MECHANICS



EVALUATION OF ACCURACY AND EFFICIENCY OF NUMERICAL METHODS FOR CONTACT PROBLEMS

MASTER'S THESIS

Author: Bc. Evžen Korec

Supervisor: prof. Ing. Milan Jirásek, DrSc.

2019/2020

ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE
Fakulta stavební
Thákurova 7, 166 29 Praha 6

ZADÁNÍ DIPLOMOVÉ PRÁCE

I. OSOBNÍ A STUDIJNÍ ÚDAJE

Příjmení: Korec Jméno: Evžen Osobní číslo: 439012

Zadávající katedra: Katedra mechaniky

Studijní program: Stavební inženýrství

Studijní obor: Konstrukce a dopravní stavby

II. ÚDAJE K DIPLOMOVÉ PRÁCI

Název diplomové práce: Vyhodnocení přesnosti a účinnosti numerických metod pro kontaktní úlohy

Název diplomové práce anglicky: Evaluation of Accuracy and Efficiency of Numerical Methods for Contact Problems

Pokyny pro vypracování:

Na základě rešerše literatury vyberte vhodné metody pro vynucení kontaktních podmínek při popisu mechanického kontaktu mezi pružnými tělesy a popište jejich teoretické základy. Vybrané metody implementujte do systému FEniCS, proveďte ověření správnosti implementace pomocí simulace vhodně zvolené úlohy s analytickým řešením a porovnejte přesnost a výpočetní náročnost jednotlivých metod.

Seznam doporučené literatury:

Wriggers, P. Computational Contact Mechanics. Chichester: Wiley, 2002
Yastrebov, Vladislav A. Numerical Methods in Contact Mechanics. Numerical Methods in Engineering Series. London : Hoboken, N.J.: ISTE : Wiley, 2013

Jméno vedoucího diplomové práce: Prof. Ing. Milan Jirásek, DrSc.

Datum zadání diplomové práce: 4.10.2019 Termín odevzdání diplomové práce: 17. 5. 2020
Údaj uveďte v souladu s datem v časovém plánu příslušného ak. roku

Podpis vedoucího práce Podpis vedoucího katedry

III. PŘEVZETÍ ZADÁNÍ

Beru na vědomí, že jsem povinen vypracovat diplomovou práci samostatně, bez cizí pomoci, s výjimkou poskytnutých konzultací. Seznam použité literatury, jiných pramenů a jmen konzultantů je nutné uvést v diplomové práci a při citování postupovat v souladu s metodickou příručkou ČVUT „Jak psát vysokoškolské závěrečné práce" a metodickým pokynem ČVUT „O dodržování etických principů při přípravě vysokoškolských závěrečných prací".

24.2.2020

Datum převzetí zadání Podpis studenta(ky)

Declaration of Authorship

I declare that this master's thesis was carried out by me under the supervision of Prof Milan Jirásek and only with the use of materials that are stated in the literature sources.

In ..... on ..... signature: .....

Acknowledgements

I would like to wholeheartedly thank Prof Milan Jirásek (CTU in Prague) for being an excellent, ceaselessly motivating supervisor and mentor, who kindly supported me throughout my whole engineering studies and whose vast knowledge and deep passion sparked me to study mechanics.

I would also like to offer my special thanks to Prof Garth N. Wells (University of Cambridge) for giving me an opportunity to study contact mechanics and FEniCS in his group at the University of Cambridge and to participate in thrilling ASiMoV project for a few months. Also, I would like to express my very great appreciation to the members of his group and his co-workers - Yuxuan Chen (University of Cambridge) for his passionate explanations of multigrid methods, Dr Nathan Sime (Carnegie Institution for Science), Dr Matthew Scroggs (University of Cambridge), Joseph Dean (University of Cambridge), Dr C. N. Richardson (University of Cambridge) and especially Jørgen S. Dokken (University of Cambridge) for their advice and support with FEniCS. Many thanks also goes to Dr Samuel Groth (University of Cambridge) for stimulating mathematical discussions and his friendliness and humanity.

I wish to thank Jaroslav Schmidt (CTU in Prague) for many thrilling discussions of mechanics and FEniCS and for his friendly encouragement. The moral support and help of Prof Aleš Nekvinda (CTU in Prague), Prof Jan Zeman (CTU in Prague), Dr Alena Zemanova (CTU in Prague), Dr Martin Horák (CTU in Prague) and Dr Emilio Martinez-Paneda (Imperial College London) were deeply appreciated.

My thanks are also extended to Dr Jeremy Bleyer (École des Ponts ParisTech) for his helpful advice on the implementation of integration schemes in FEniCS and his excellent website `https://comet-fenics.readthedocs.io`, which provides invaluable help to those new to FEniCS.

Advice on the Nitsche method provided by Dr Konstantinos Poulios (Technical University of Denmark) was greatly appreciated.

Finally, I wish to thank my parents for their love, unbelievable patience and their trust in me.

Anotace

Tato diplomová práce se zabývá variačními metodami, které umožňují formulovat problém kontaktu lineárně pružného tělesa bez tření jako nepodmíněnou variační rovnost, která může být posléze diskretizována a řešena metodou konečných prvků. Hlavní důraz je kladen na Nitscheho metody podle Wriggerse a Zavariseho [56] a podle Fabrého, Pousina a Renarda [15]. V současné době nejrozšířenější konečněprvkové softwarové balíky, jako jsou ANSYS, ABAQUS a COMSOL, využívají pro modelování kontaktu především standardní metody penalty a smíšené metody [57, Kapitola 1.1.1, s.7]. Ukazuje se, že právě Nitscheho metody mají potenciál překonat klasické obtíže spojené se standardními metodami penalty a smíšenými metodami. Na rozdíl od metod penalty jsou Nitscheho metody konzistentní a kontaktní okrajové podmínky jsou vynuceny přesně (na teoretické úrovni). Je také možné využít mnohem menší hodnotu parametru penalty, čímž se lze vyhnout problémům spojeným se špatným podmíněním úlohy, charakteristickým pro metody penalty. Nitscheho metoda ale současně nevyžaduje přidání žádných dalších neznámých (Lagrangeových multiplikátorů) a výsledný diskrétní systém tak není nadbytečně rozšířen, jako je tomu v případě smíšených metod. Oproti smíšeným metodám také není třeba věnovat pozornost splnění Babuškovy-Brezziho podmínky. V této diplomové práci se ukazuje, že analyzované Nitscheho metody úzce souvisejí s metodami penalty a metodou augmentovaného lagrangiánu. V práci jsou prezentovány slabé formulace těchto metod a zkoumají se rozdíly mezi formulací Nitscheho metody podle Wriggerse a podle Fabrého, Pousina a Renarda. Všechny metody jsou implementovány do prostředí FEniCS (výpočetní platforma pro řešení parciálních diferenciálních rovnic metodou konečných prvků) a jejich přesnost a výkonnost se testuje na různých dvourozměrných a trojrozměrných problémech kontaktu lineárně pružného tělesa s dokonale tuhou rovinou. Na jednoduchém dvourozměrném příkladu je ukázáno, že funkce, kterou získáme jako levou stranu diskretizované slabé formy Wriggersovy varianty Nitscheho metody, není spojitá vzhledem k neznámým stupňům volnosti. Tento poznatek vysvětluje problémy s konvergencí Newtonovy metody při řešení Wriggersovou variantou Nitscheho metodou, které jsme zaznamenali při numerických experimentech.

Klíčová slova

výpočetní kontaktní mechanika, kontakt bez tření, lineární pružnost, Nitscheho metoda, variační nerovnosti, nepodmíněné optimalizační metody, metoda penalty, metoda augmentovaného lagrangiánu, FEniCS

Annotation

This thesis is concerned with various methods that allow us to formulate the frictionless linear elastic contact problems as an unconstrained variational equality, which is then discretised and solved with the finite element method. The main focus is on Nitsche methods in the forms used respectively by Wriggers and Zavarise [56] and Fabré, Pousin and Renard [15]. Currently, standard penalty and mixed methods are dominant in the modern leading finite element software packages such as ANSYS, ABAQUS and COMSOL [57, Chapter 1.1.1, p.7]. Nitsche methods display a potential to overcome classic drawbacks of the penalty and mixed methods. Unlike penalty methods, Nitsche methods are consistent, and contact boundary conditions are enforced precisely (on the theoretical level). Also, a significantly smaller value of the penalty parameter is necessary and the possible ill-conditioning, so characteristic for penalty methods, is thus avoided. At the same time, no additional unknowns (Lagrange multipliers) are introduced; thus, the corresponding discrete system is not enlarged, and one does not have to worry about the Babuška-Brezzi condition. In this thesis was shown that the analysed Nitsche methods are closely related to penalty methods and the augmented Lagrangian method. The weak forms of all these methods are presented, and differences between Wriggers' version and Fabré, Pousin and Renard's version of Nitsche method are investigated. All methods are implemented in FEniCS (the computational platform for solving partial differential equations with the finite element method), and their accuracy and efficiency is tested on various two- and three-dimensional numerical examples of contact of an elastic body with a rigid plane. By means of the simple two-dimensional example it is shown that the function obtained as the left-hand side of the discretised weak form of the Nitsche-Wriggers method is not continuous with respect to the unknown displacement DOFs. This finding explains the convergence problems (of Newton's method) that the Nitsche-Wriggers method suffers from, unlike other investigated methods.

Keywords

computational contact mechanics, frictionless contact, linear elasticity, Nitsche method, variational equalities, unconstrained optimization methods, penalty method, augmented Lagrangian method, FEniCS

Contents

1 Introduction

This thesis treats the mathematical modelling of the mechanical phenomenon called contact. As its very name suggests, contact describes the interaction of bodies that can touch each other on their boundaries and exchange energy such as heat or electric charge [57, Chapter 1, p.1]. Even though we will consider contact as a single process happening on a macroscopic scale, it is rather a system of many simultaneous processes on various scales ranging from microscopical to microscopical ones. Therefore, the nature of contact is multiscale and multiphysical [57, Chapter 1, p.1], which makes understanding contact a rather complicated task. Moreover, many contact problems involve friction, which makes the whole problem even more complex. There are several reasons for this. Firstly, the friction is a non-conservative phenomenon and secondly, the optimal description of friction is a matter of scientific discussion in physical and engineering community up to these days.

However, despite its inconveniently complex nature, scientists struggle for centuries to uncover the secrets of this complex phenomenon. This is for the reason that contact is virtually omnipresent. Our every movement, be it simple walking or running, would never be possible without frictional contact. Contact is also crucial for our technology. Let us take the example of the wheel, one of humanity's greatest inventions, which is completely based on the existence of frictional contact. For mechanical engineers, the understanding of contact is necessary when designing gears in mechanical devices, bearings, engines, turbines or car brakes. The applications also involve drilling, metal forming or cutting processes (like sheet metal or bulk forming) and the crash analysis of cars. The civil engineer would probably mention many problems of geotechnical engineering such as foundations including piles, lift-off of the foundation from the soil due to eccentric forces, driving of piles into the soil, bearing problems of steel constructions, the connecting of structural members by bolts or screws or the impact of cars against building structures [55, Chapter 1, p.1-5], [57, Chapter 1, p.1]. Because it is relatively complicated to investigate contact experimentally and analytical models exist only for very simple cases, it is very desirable to develop methods to model contact numerically.

The adequately precise numerical modelling of contact can save a tremendous amount of time and resources which would have to be otherwise wasted on experimental testing (which sometimes does not have to be even possible). The excellent example of successful numerical modelling is the computational crash analysis of cars that reduces the development time and costs of modern cars and increases their safety.

However, modelling of contact is a demanding matter as well. Firstly, because the objects are deformable, we do not know the area in which bodies in contact (let us imagine tyre and road) interact with each other. Even if all the other relations describing the problem are linear, this aspect makes the resulting boundary value problem non-linear. In addition to this theoretical aspect, there are many other complications. Therefore, many authors (for instance [57, Chapter 1.1, p.5], [55, Chapter 1, p.1]) and engineering professionals state that even nowadays, certain complex frictional problems pose a formidable challenge and fully robust algorithms for them are not known.

Due to its industrial importance and vast possibilities of applications, contact has been attracting scientists for centuries. The names of early researchers investigating frictional contact involve famous physicists and mathematicians like Leonardo Da Vinci, Leonhard Euler, Charles-Augustin de Coulomb and Heinrich Hertz [55, Chapter 1, p.6-8]). And it is Heinrich Hertz who is considered to be the father of the modern mechanical theory of contact, which started in 1882 with his famous paper 'On the Contact of Elastic Solids' [27]. His interest was not primarily mechanical but rather optical, as he researched optical interference between glass lenses. As Johnson mentions [33], Hertz worked out his theory during the Christmas vacation 1880 at the age

of twenty-three, and it immediately aroused considerable interest. With the rapid development of engineering in the late 19th and early 20th century, the interest in contact mechanics continued, and new works emerged, as industrial applications as railways, reduction gears and rolling contact bearings required it.

Yastrebov writes [57, Chapter 1.1, p.5] that the history of modern computational contact started in 1933, when Italian mathematician and civil engineer Antonio Signorini formulated the general problem of the equilibrium of a linearly elastic body in frictionless contact with a rigid foundation [49]. Ingeniously, he formulated the problem mathematically consistently based on variational principles of mechanics as a variational inequality, which was a very new structure at the time. The story of this problem is reported by Fichera [17], [51]. During his life, Signorini very much desired to know whether his problem is well-posed and so there exists a unique solution. This continued until 1962 when Signorini started to suffer from health problems and with the vision of approaching death, he expressed a wish to his friends to know the answer to his question before he deceases. As his student and close friend Fichera was working on this problem, another close friend to Signorini, Picone, started to chase Fichera to find a solution. With tremendous effort, Fichera finally found it on the first days of January 1963. To the great joy and relief of dying Signorini, a preliminary research announcement was written and submitted exactly a week before his death. The last conversation of Signorini with his family doctor Damiano Aprile was reported to be [17], [51]:

'My disciple Fichera gave me great contentment.', said Signorini.

'But you had many, Professor, during your life.', replied the doctor.

'But this is the greatest one.', which were Signorini's last words.

As Yastrebov mentions [57, Chapter 1.1, p.5], since the time of NASTRAN code in 1965, the finite element method proved itself to be the most efficient numerical method to treat contact.

For its vast industrial relevance, contact mechanics still attracts the attention of researchers and their industrial partners, who desperately seek for more reliable and robust codes for the modelling of computational contact. Such example is the recent Rolls-Royce investment of £14.7M (in partnership with other companies) to the high-performance computing research project ASiMoV (Advanced Simulation and Modelling of Virtual Systems)[11], [43]. The aim of this project is [11]: '...to simulate an entire aircraft jet engine in operation at very high fidelity, with a goal that one day the civil aviation authorities would be confident enough to certify the virtual design.' An important part of the project is, of course, modelling of contact, which should be implemented in FEniCS-X, which is a Python-based finite element simulation package.

An essential aspect of every project including computational modelling of contact, such as ASiMoV, is the choice of the method for enforcing contact boundary conditions or in other words, the mathematical formulation of the contact problem which is then discretised and solved with finite element method. In all the leading modern finite element software packages such as ANSYS, ABAQUS and COMSOL, the method for modelling contact is based on some variational equality, in which the contact constraints are enforced with some additional term of the weak form [57, Chapter 1.1.1, p.7]. Traditionally, there are two standard classes of such methods, known from the optimisation theory — penalty methods and mixed methods. The problem is that both of them suffer from certain drawbacks, especially in the case of large-scale problems such as the modelling of an entire jet engine in operation. As their name suggests, penalty methods penalise the penetration, but even on the theoretical level, they cannot prevent it entirely. The amount of penetration depends on the value of the chosen penalty parameter. The higher it is, the smaller penetration is allowed. The problem is that with too

large penalty values, the discrete problem starts to be seriously ill-conditioned. Mixed methods, on the other hand, are based on introducing an additional unknown called Lagrange multiplier, which represents the normal contact traction. Although this formulation allows no penetration on the theoretical level, the number of DOFs is increased with new DOFs representing Lagrange multiplier values. This means that the discrete problem is significantly enlarged with new equations connected with Lagrange multiplier DOFs. Also, in the case of mixed methods, the choice of interpolation functions (in terms of the finite element method) is not entirely arbitrary but has to satisfy the Babuška-Brezzi condition.

The mentioned drawbacks of classic approaches motivate the search for a new class of methods which would fulfil contact conditions exactly (on the theoretical level), require lower values of the penalty parameter to avoid ill-conditioning but at the same time, would not require introducing any additional unknown except displacements. And really, it turns out that so-called Nitsche methods show a potential to fulfil all these requirements. Discovered in 1970 by Joachim A. Nitsche [44], these methods were first used for treating the Dirichlet boundary or interface conditions in a weak sense. Despite the fact that they gradually gained considerable popularity in the finite element community, it was not until the beginning of a new millennium, when the first attempts to apply Nitsche methods to contact emerged. These were by Hansbo, A., Hansbo, P. and co-authors [25], [26] and by Wriggers and Zavarise [56]. Despite the effort of these authors, Nitsche methods seemed to be somehow overlooked for treating contact problems up to 2013, when Chouly, Hild and Renard presented a new formulation of Nitsche method [29]. The authors provided a very consistent and thorough mathematical analysis of this method for Signorini's problem, which attracted the attention of other researchers. Nitsche method from [29] was generalised to the problem of contact of two elastic bodies by Fabré, Pousin and Renard in [15], and this method has been thoroughly investigated until these days. The complex overview of recent research on this version of Nitsche method could be found in [31].

For their great potential, in this thesis, we will focus on the two versions of Nitsche methods — one by Wriggers and Zavarise [56] and one by Fabré, Renard and Pousin in [15]. Both methods will be theoretically analysed, some differences between them identified and finally, their performance tested on numerical examples implemented in FEniCS. Also, the augmented Lagrangian method and one version of the penalty method will be presented, and their connection with the analysed Nitsche methods will be described. We will limit ourselves to the case of linear elastic bodies, small deformations and frictionless contact. Despite the methods themselves will be formulated for the case of two bodies in contact, the numerical examples will be limited to the contact of one body with a rigid plane. This is to eliminate the potential influence of the contact detection algorithm on the numerical performance of the methods. The choice of FEniCS is motivated by the recent efforts of the mentioned ASiMoV project to implement sophisticated contact algorithms to this very software. Also, FEniCS allows us to directly provide the variational formulation coded in UFL language as an input. This is very convenient when investigating the physically same problem described with several variational formulations, which we intend to compare.

2 The frictionless contact problem

In this chapter, we will present the continuum formulation of the frictionless contact problem of two elastic bodies. Firstly, we will introduce the necessary notation. Equipped with this, we will present the equations of elastic continua, and we will formulate contact conditions. This will establish the contact problem in its strong-form. After that, we will formulate the problem weakly as a variational inequality. For the sake of simplicity, only two elastic bodies in contact will be considered.

Let us present the relevant notation first. The whole domain Ω consists of two deformed elastic bodies in $\mathbb{R}^d$, where d is the dimension of the problem, which is either 2 or 3. Because we will follow a master-slave approach in the following chapters, one (arbitrarily chosen) elastic body will be considered a master body Ω^M and the other one a slave body Ω^S. In the master-slave approach, the so-called non-penetration conditions (describing the possibility of contact as will be discussed further) are prescribed on the boundary of the slave body and integration is also performed over the slave surface. Although it is very customary to consider the master-slave approach when studying contact (and we will do that too in this thesis), it is questionable whether it is necessary. On the one hand, Konyukhov and Izi mention in their book [35, Chapter 4, p.45-46] that: 'Historically, through the development of the currently popular finite element codes (ABAQUS, ANSYS, LS-DYNA, etc.) the master-slave approach has been proved to be the most efficient approach in computational contact mechanics.' On the other hand, Mlika, Renard and Chouly comment in [41], [32] that considering the master-slave approach causes difficulties in the cases of self-contact and multibody contact. In [41] we can find that: '...in case of self-contact and multibody contact where it is impossible or impractical to a priori nominate a master surface and a slave one. Automating the detection and the separation between slave and master surfaces in these cases may generate a lack of robustness since it may create detection problems.' The authors of [41], [32] also present an alternative approach to the master-slave strategy, which they call the unbiased strategy. This strategy utilizes the Nitsche method [30] (described in the following chapters) and both surfaces in contact are treated in the same way. Integration is performed over both of them too. The interested reader is recommended to see [41], [32] for more details on this strategy.

However, in this thesis, we will stick to the more standard master-slave approach and so we can write $\Omega = \Omega^M \cup \Omega^S$. We wish to point out that the symbol Ω always refers to the actual (deformed) configuration. However, in the numerical examples described in further chapters, the deformed configuration will be replaced with the undeformed one to simplify the computations.

The boundary of the slave and the master body will be denoted as Γ^S and Γ^M respectively, their union is marked as Γ, where $\Gamma = \Gamma^S + \Gamma^M$. The boundaries of both bodies consist of three possible types of subsets. Firstly $\Gamma^{D,S}$, $\Gamma^{D,M}$ where Dirichlet boundary conditions are prescribed, secondly $\Gamma^{\sigma,C}$, $\Gamma^{\sigma,M}$ where traction is prescribed and lastly $\Gamma^{C,S}$, $\Gamma^{C,M}$ where the further described contact boundary conditions (the non-penetration conditions) are defined. Generally it holds $\Gamma = \Gamma^D + \Gamma^\sigma + \Gamma^C = \Gamma^{D,S} + \Gamma^{D,M} + \Gamma^{\sigma,M} + \Gamma^{\sigma,S} + \Gamma^{C,M} + \Gamma^{C,S}$. Even though Γ^D and Γ^σ are supposed to be known a priori, in contact problems, Γ^C denotes the region where contact is allowed to happen, but not where it actually happens. For this reason, Γ^C will be called the potential contact zone while the region $\overline{\Gamma}^C$ which is in contact will be denoted as the active contact zone. It is supposed that $\overline{\Gamma}^C \subset \Gamma^C$. Again, the active contact zone can be separated to the parts on master and slave bodies as $\overline{\Gamma}^C = \overline{\Gamma}^{C,S} + \overline{\Gamma}^{C,M}$. We also presume that Γ^D, Γ^σ and Γ^C are non-overlapping regions on each body separately (but when in contact, $\overline{\Gamma}^{C,S}$ and $\overline{\Gamma}^{C,M}$ coincide). Each of the regions is supposed to be of non-zero measure. As the objective of this thesis in not to deal with problems with complicated geometry but rather investigate the methods for enforcing contact boundary conditions, let us suppose that the boundaries Γ^S and Γ^M are

piecewise C^1 differentiable.

The objective of the formulation of the frictionless contact problem in the continuum is to prescribe such boundary conditions on Γ^C that no non-physical penetration of one body into another is allowed. This goal suggests that we need to define some metric on the boundary Γ^C to measure this penetration or in other words, 'gap' between both bodies. To measure the 'gap' in each point of $\Gamma^{C,S}$ and $\Gamma^{C,M}$ we need to find a projection of one boundary to another, in other words to find to every point from $\Gamma^{C,S}$ a point from $\Gamma^{C,M}$. The 'gap' is then a distance of those points. In this text, we will adopt a very standard definition of the normal gap using the mentioned master-slave approach. It should be pointed out that many other definitions of the gap function are possible, see for instance [57, Chapters 2.1, 2.2, p.15-35]. The further mentioned notation will be inspired by the one presented in [15], and we will adopt similar simplifying assumptions as well.

In order to properly define the gap function, we will first define an orthogonal projection Π from the slave boundary potential contact zone $\Gamma^{C,S}$ to the master boundary potential contact zone $\Gamma^{C,M}$. The definition of an orthogonal projection Π reads

$$\Pi : \begin{array}{ll} \Gamma^{C,S} & \to \Gamma^{C,M} \\ \mathbf{x}^S & \mapsto \Pi(\mathbf{x}^S) = \mathbf{x}^M \end{array} \tag{1}$$

Here we would like to point out that $\mathbf{x} \in \Omega$ represents the Cartesian coordinates of some material point of the deformed body. We presume that to all the material points $\mathbf{x} \in \Omega$ initially existed a Cartesian coordinates $\mathbf{X} \in \Omega^0$, such that $\mathbf{x} = \mathbf{X} + \mathbf{u}$, where $\Omega_0 = \Omega_0^S + \Omega_0^M$ is the initial (undeformed) configuration of both bodies. Because we suppose small deformation, it does not matter whether we differentiate with respect to $\mathbf{X}$ or $\mathbf{x}$. Also, because of the small deformation assumption, we can integrate over the undeformed (original) configuration Ω_0 instead of the deformed (actual) configuration Ω. To simplify our analysis, let us suppose that operator Π is a C^1 differentiable one to one mapping on $\Pi\left(\Gamma^{C,S}\right)$. As [15] mentions, this holds, for instance, when $\Gamma^{C,S}$, $\Gamma^{C,M}$ are convex and C^1 differentiable.

For the definition of the gap function, it is also necessary to define the outward unit normal vector $\mathbf{n}$. Because we consider only the case of small deformation, $\mathbf{n}$ is supposed to independent of $\mathbf{u}$. In terms of the master-slave approach, the normal vector is chosen from the master surface $\Gamma^{C,M}$ such that

$$\mathbf{n}^M : \begin{array}{ll} \Gamma^{C,S} & \to \mathbb{R}^d \\ \mathbf{x}^S & \mapsto \mathbf{n}(\Pi(\mathbf{x}^S)) \end{array} \tag{2}$$

Equipped with the previously described orthogonal projection Π and the outward unit normal $\mathbf{n}$ we can now finally define the gap function as

$$g_n : \begin{array}{ll} \Gamma^{C,S} & \to \mathbb{R} \\ \mathbf{x}^S & \mapsto (\mathbf{x}^S - \Pi(\mathbf{x}^S)) \cdot \mathbf{n}^M = (\mathbf{x}^S - \mathbf{x}^M) \cdot \mathbf{n}^M \end{array} \tag{3}$$

Because $\mathbf{u}$ is the only unknown variable in the definition of the gap function, we will use it as an argument and write

$$g_n(\mathbf{u}) = \mathbf{X}^S \cdot \mathbf{n}^M + \mathbf{u}^S \cdot \mathbf{n}^M - \mathbf{X}^M \cdot \mathbf{n}^M - \mathbf{u}^M \cdot \mathbf{n}^M \tag{4}$$

The gap function, in the form we presented it, allow us to detect whether the elastic bodies in the continuum are in contact or not. The problem is that in this thesis, we intend to numerically solve the contact problems with finite element method, where both bodies in contact are discretised. Here, contact detection can become somewhat tricky and requires implementing some special techniques. The most simple case is when both meshes are so-called matching or conforming so that the nodes in contact assume the same coordinates. However, generating matching meshes in the case of complicated geometries could be very complicated and

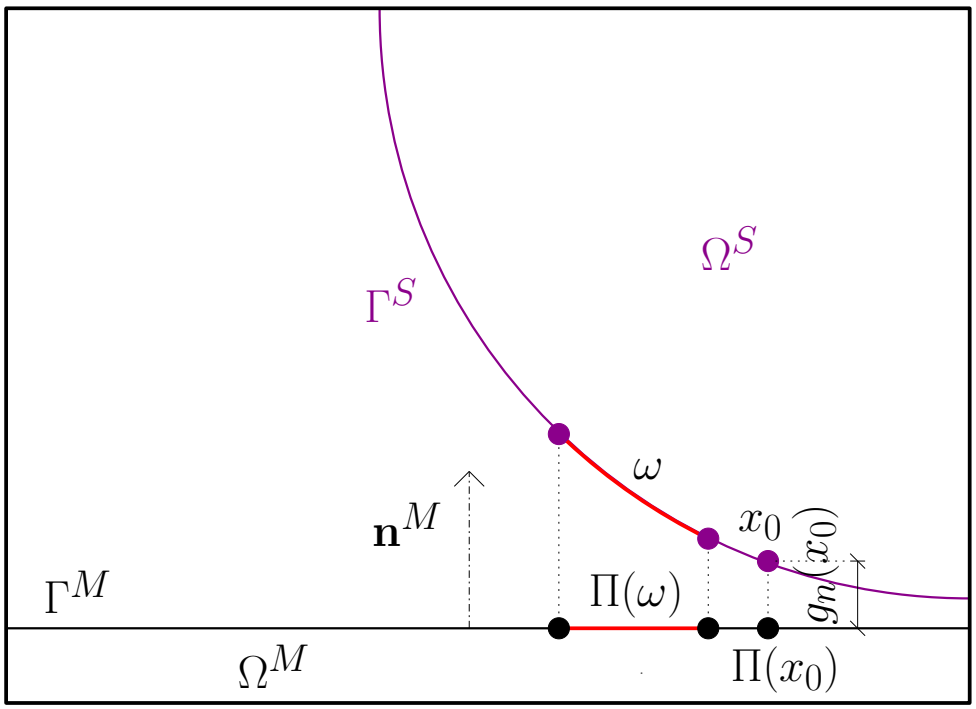


Figure 1: Illustration of geometrical quantities in the master-slave approach(inspired by [15])

ineffective. The same problem also comes with parallel computations. For this reason, many strategies for contact detection in non-matching meshes have been developed. However, it is not an objective of this thesis to concentrate on various detection strategies but rather on the methods for enforcing the contact boundary conditions. Therefore, even though we will formulate the methods for the case of contact of two elastic bodies, in the numerical examples we will focus only on the contact of one elastic body with a rigid obstacle (a rigid plane more precisely). For the reader interested in more sophisticated contact detection techniques, we recommend to see, for example, Yastrebov's book [57, Chapter 3, p.71-102].

In order to set up properly the contact problem we also need to specify the way stress and strain are calculated and supply the appropriate definitions of contact tractions. We presume both bodies are undergoing small elastic deformation. For this reason classic definitions of 'small strain' strain tensor $\boldsymbol{\varepsilon}(\mathbf{u})$ and Cauchy stress tensor $\boldsymbol{\sigma}(\mathbf{u})$ are adopted. So $\boldsymbol{\varepsilon}(\mathbf{u}) = \frac{1}{2}\left(\boldsymbol{\nabla}\mathbf{u} + (\boldsymbol{\nabla}\mathbf{u})^T\right)$ and $\boldsymbol{\sigma}(\mathbf{u}) = \mathbf{D}\boldsymbol{\varepsilon}(\mathbf{u})$ where $\mathbf{D}$ is a fourth-order symmetric elasticity tensor. The normal tractions are defined as $\mathbf{t} = \mathbf{n}\cdot\boldsymbol{\sigma} = \sigma_n\mathbf{n} + \boldsymbol{\sigma}_t$. However, as we consider only the case of normal frictionless contact, $\boldsymbol{\sigma}_t = \mathbf{0}$ on Γ^C and so $\mathbf{t} = \sigma_n\mathbf{n}$ on Γ^C. Also for the purpose of clarity, considering the master-slave approach, let us define the normal tractions on the potential contact zone Γ^C on the master and slave body as

$$\mathbf{t}^S = \mathbf{n}^S \cdot \boldsymbol{\sigma}^S = -\sigma_n^S \mathbf{n}^M \tag{5}$$

where

$$\sigma_n^S = -\mathbf{n}^S \cdot \boldsymbol{\sigma}^S \cdot \mathbf{n}^M \tag{6}$$

and

$$\mathbf{t}^M = \mathbf{n}^M \cdot \boldsymbol{\sigma}^M = \sigma_n^M \mathbf{n}^M \tag{7}$$

where

$$\sigma_n^M = \mathbf{n}^M \cdot \boldsymbol{\sigma}^M \cdot \mathbf{n}^M \tag{8}$$

We would like to remark that relations for tractions on the master surface (7) and on the slave surface (5), are defined using only $\mathbf{n}^M$ - the unit normal vector to the master surface. This is caused by the master-slave

approach we adopted. Obviously, this is reasonable on the active contact zone $\overline{\Gamma}^C$ where the master and slave surfaces coincide and $\mathbf{n}^M = -\mathbf{n}^S$. However, on $\Gamma^C \setminus \overline{\Gamma}^C$, $\mathbf{n}^M \neq -\mathbf{n}^S$. Despite this fact, relations (5)–(8) still hold on $\Gamma^C \setminus \overline{\Gamma}^C$ because $\sigma_n^M = \sigma_n^S = 0$ on this part of the boundary. For this reason, (5)–(8) hold on the whole Γ^C.

The values of contact tractions σ_n^S and σ_n^S are bound by the following relation, which results from Newton's second law (the action-reaction principle) and the assumption of small deformation

$$\sigma_n^S = \sigma_n^M \tag{9}$$

Equation (14) can be derived in the following way. We can express Newton's second law on an arbitrary elementary surface ω from $\Gamma^{C,S}$ as

$$\begin{aligned} \forall \omega \subset \Gamma^{C,S}, \quad \int_\omega \mathbf{n}^S \cdot \boldsymbol{\sigma}^S \mathrm{d}\Gamma &= -\int_{\Pi(\omega)} \mathbf{n}^M \cdot \boldsymbol{\sigma}^M \mathrm{d}\Gamma \\ &= -\int_\omega \mathbf{n}^M \cdot \boldsymbol{\sigma}^M \mathrm{d}\Gamma \end{aligned} \tag{10}$$

In 10, we presume that because of the small deformation assumption, we can replace the integration over ω with the integration over $\Pi(\omega)$. Let us remark, that in Figure 1, the difference in the size of the regions ω and $\Pi(\omega)$ seems to be significant, but this is only for the purposes of graphic visibility.

$$\mathbf{n}^S \cdot \boldsymbol{\sigma}^S + \mathbf{n}^M \cdot \boldsymbol{\sigma}^M = 0 \tag{11}$$

or in other words

$$\mathbf{t}^S + \mathbf{t}^M = 0 \tag{12}$$

With relations (5)–(8) we can rewrite (12) as

$$-\sigma_n^S \mathbf{n}^M + \sigma_n^M \mathbf{n}^M = 0 \tag{13}$$

and so

$$\sigma_n^S = \sigma_n^M \tag{14}$$

Now, we can proceed to the strong formulation of the frictionless elastic contact problem. Firstly, let us mention a few words on the notation and boundary conditions. To avoid to intriguing and lengthy notation, we will omit the superscript $\bullet^M$ or $\bullet^S$ of certain quantities. For example, we will write $\mathbf{u}$ by which we mean $\mathbf{u} = \mathbf{u}^S$ on Ω^S and $\mathbf{u} = \mathbf{u}^M$ on Ω^M.[1] Using the same notation we presume that bodies are subjected to volume forces $\overline{\mathbf{f}}$ where $\overline{\mathbf{f}} = \rho^S \overline{\mathbf{b}}^S$ on Ω^S and $\overline{\mathbf{f}} = \rho^M \overline{\mathbf{b}}^M$ on Ω^M, to surface loads $\overline{\mathbf{t}}$ where $\overline{\mathbf{t}} = \overline{\mathbf{t}}^S$ on $\Gamma^{\sigma,S}$ and $\overline{\mathbf{t}} = \overline{\mathbf{t}}^M$ on $\Gamma^{\sigma,M}$. Let us remark that ρ^S, ρ^M are the densities of the slave and the master body respectively. Moreover, Dirichlet boundary conditions $\overline{\mathbf{u}} = \overline{\mathbf{u}}^S$ and $\overline{\mathbf{u}} = \overline{\mathbf{u}}^M$ are prescribed on $\Gamma^{D,S}$ and $\Gamma^{D,M}$ respectively. With these, all boundary conditions have been prescribed except the ones on the potential contact zones $\Gamma^{C,S}$. We remind, that because we adopted the master-slave approach, contact conditions are prescribed only on $\Gamma^{C,S}$. The contact boundary conditions for frictionless contact can be formulated in the following way [55, Chapter 5, p.97]

$$\begin{aligned} g_n &\geq 0 \quad \text{(i)} \\ \sigma_n &\leq 0 \quad \text{(ii)} \\ \sigma_n g_n &= 0 \quad \text{(iii)} \end{aligned} \tag{15}$$

[1] We would like to note that another notations are possible. For example the authors of [15] understand the displacement as a vector $\mathbf{u} = (\mathbf{u}^S, \mathbf{u}^M)$ and define this object on the Cartesian product $\Omega^S \times \Omega^M$ However, this notation is relatively unusual in the engineering literature. In order to avoid confusion, we adopted more intuitive notation of an 'engineering' character.

These conditions are also known as the Hertz–Signorini–Moreau conditions for frictionless contact. They also have the structure of the Kuhn-Tucker-Karush conditions, well known from the theory of optimization. For the sake of simplicity, we will sometimes refer to them as to the contact conditions.

Equipped with the knowledge of boundary conditions, we can define the elastic frictionless contact problem now. Remembering the well-known equations describing the linear elastic problem in a continuum, we require such a displacement $\mathbf{u} \in \mathbf{U}^2$, $\mathbf{U}^2 = \{\mathbf{u} \in H^2(\Omega)^d; \mathbf{u} = \overline{\mathbf{u}} \text{ on } \Gamma^D\}$ [2] that the equations and inequalities below are satisfied

$$\begin{aligned} -\boldsymbol{\nabla} \cdot \boldsymbol{\sigma}(\mathbf{u}) &= \rho \overline{\mathbf{b}} && \text{in} \quad \Omega \\ \boldsymbol{\sigma}(\mathbf{u}) &= \mathbf{D}\boldsymbol{\varepsilon}(\mathbf{u}) && \text{in} \quad \Omega \\ \boldsymbol{\varepsilon}(\mathbf{u}) &= \frac{1}{2}\left(\boldsymbol{\nabla}\mathbf{u} + (\boldsymbol{\nabla}\mathbf{u})^T\right) && \text{in} \quad \Omega \\ \mathbf{u} &= \overline{\mathbf{u}} && \text{on} \quad \Gamma^D \\ \mathbf{n} \cdot \boldsymbol{\sigma} &= \overline{\mathbf{t}} && \text{on} \quad \Gamma^\sigma \end{aligned} \tag{16}$$

$$\begin{aligned} g_n &\geq 0 \quad \text{on} \quad \Gamma^{C,S} \\ \sigma_n &\leq 0 \quad \text{on} \quad \Gamma^{C,S} \\ \sigma_n g_n &= 0 \quad \text{on} \quad \Gamma^{C,S} \end{aligned} \tag{17}$$

With 16 and 17, the strong form of the elastic frictionless contact problem is formulated. When referring to this problem, we will sometimes omit some adjectives and refer to it just as to strong form of the contact problem.

However, as we intend to solve our contact problem numerically with the finite element method, we need to find the so-called weak form of our problem. For this purpose, let us define V as the space of test functions $\mathbf{v}$ (where $\mathbf{v} = \mathbf{v}^S$ on Ω^S and $\mathbf{v} = \mathbf{v}^M$ on on Ω^M) such that

$$\mathbf{V} = \left\{\mathbf{v} \in H^1(\Omega)^d; \mathbf{v} = \mathbf{0} \text{ on } \Gamma^D\right\} \tag{18}$$

We will follow a standard procedure and multiply both sides of the first equation in (16) and integrate the whole equation over $\Omega = \Omega^M \cup \Omega^S$. The result can be rewritten as

$$\int_\Omega (\boldsymbol{\nabla} \cdot \boldsymbol{\sigma}(\mathbf{u})) \cdot \mathbf{v} \, \mathrm{d}V + \int_\Omega \rho \overline{\mathbf{b}} \cdot \mathbf{v} \, \mathrm{d}V = 0 \quad \forall \mathbf{v} \in \mathbf{V} \tag{19}$$

Applying the famous Gauss-Ostrogradsky theorem and multiplying the whole equation by -1 gives us

$$\int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\nabla}\mathbf{v} \, \mathrm{d}V - \int_\Gamma \mathbf{n} \cdot \boldsymbol{\sigma}(\mathbf{u}) \cdot \mathbf{v} \, \mathrm{d}A - \int_\Omega \rho \overline{\mathbf{b}} \cdot \mathbf{v} \, \mathrm{d}V = 0 \quad \forall \mathbf{v} \in \mathbf{V} \tag{20}$$

Now let us separate integrals over Γ to the integrals over the corresponding subdomains.

$$\begin{aligned} &\int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\nabla}\mathbf{v} \, \mathrm{d}V - \int_\Omega \rho \overline{\mathbf{b}} \cdot \mathbf{v} \, \mathrm{d}V \\ &- \int_{\Gamma^\sigma} \mathbf{n} \cdot \boldsymbol{\sigma}(\mathbf{u}) \cdot \mathbf{v} \, \mathrm{d}A - \int_{\Gamma^C} \mathbf{n} \cdot \boldsymbol{\sigma}(\mathbf{u}) \cdot \mathbf{v} \, \mathrm{d}A = 0 \quad \forall \mathbf{v} \in \mathbf{V} \end{aligned} \tag{21}$$

Utilizing the prescribed Neumann boundary condition $\overline{\mathbf{t}} = \mathbf{n} \cdot \boldsymbol{\sigma}$ on Γ^σ and separating the integral over Γ^C into the integrals over the supposed contact zones of the master and slave body, we can write

$$\begin{aligned} &\int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\nabla}\mathbf{v} \, \mathrm{d}V - \int_\Omega \rho \overline{\mathbf{b}} \cdot \mathbf{v} \, \mathrm{d}V - \int_{\Gamma^\sigma} \overline{\mathbf{t}} \cdot \mathbf{v} \, \mathrm{d}A \\ &- \int_{\Gamma^{C,M}} \mathbf{n}^M \cdot \boldsymbol{\sigma}^M(\mathbf{u}) \cdot \mathbf{v}^M \, \mathrm{d}A - \int_{\Gamma^{C,S}} \mathbf{n}^S \cdot \boldsymbol{\sigma}^S(\mathbf{u}) \cdot \mathbf{v}^S \, \mathrm{d}A = 0 \quad \forall \mathbf{v} \in \mathbf{V} \end{aligned} \tag{22}$$

[2] Here, $H^s(\Omega)^d$ is the Sobolev space, i.e. the space of functions which derivatives up to the order s belong to the Lebesgue space $L^2(\Omega)^d$. The symbol d denotes the geometrical dimension of the problem, which is either 2 or 3.

Utilizing (5)–(8), (22) can be rewritten as

$$\begin{aligned}&\int_\Omega \boldsymbol{\sigma}\left(\mathbf{u}\right) : \boldsymbol{\nabla}\mathbf{v}\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\mathbf{v}\,\mathrm{d}V - \int_{\Gamma^\sigma}\overline{\mathbf{t}}\cdot\mathbf{v}\,\mathrm{d}A\\&- \int_{\Gamma^{C,M}}\sigma_n^M\mathbf{n}^M\cdot\mathbf{v}^M\,\mathrm{d}A + \int_{\Gamma^{C,S}}\sigma_n^S\mathbf{n}^M\cdot\mathbf{v}^S\,\mathrm{d}A = 0 \quad \forall\mathbf{v}\in\mathbf{V}\end{aligned} \tag{23}$$

Let us point out that in the following text we will sometimes alternate σ_n and $\sigma_n(\mathbf{u})$ as the notation of the normal contact traction to preserve the notation as simple as possible.

$$\begin{aligned}&\int_\Omega \boldsymbol{\sigma}\left(\mathbf{u}\right) : \boldsymbol{\nabla}\mathbf{v}\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\mathbf{v}\,\mathrm{d}V - \int_{\Gamma^\sigma}\overline{\mathbf{t}}\cdot\mathbf{v}\,\mathrm{d}A\\&- \int_{\Gamma^{C,S}}\sigma_n^M\mathbf{n}^M\cdot\mathbf{v}^M\,\mathrm{d}A + \int_{\Gamma^{C,S}}\sigma_n^S\mathbf{n}^M\cdot\mathbf{v}^S\,\mathrm{d}A = 0 \quad \forall\mathbf{v}\in\mathbf{V}\end{aligned} \tag{24}$$

This equation could be further simplified using the notation of 'jump' of $\mathbf{v}\cdot\mathbf{n}^M$ defined as $[\mathbf{v}]\cdot\mathbf{n}^M = \mathbf{v}^S\cdot\mathbf{n}^M - \mathbf{v}^M\cdot\mathbf{n}^M$ and realizing that $\sigma_n^S = \sigma_n^M$ (which was derived from Newton's second law in (14))

$$\begin{aligned}&\int_\Omega \boldsymbol{\sigma}\left(\mathbf{u}\right) : \boldsymbol{\nabla}\mathbf{v}\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\mathbf{v}\,\mathrm{d}V - \int_{\Gamma^\sigma}\overline{\mathbf{t}}\cdot\mathbf{v}\,\mathrm{d}A\\&+ \int_{\Gamma^{C,S}}\sigma_n^S\left[\mathbf{v}\right]\cdot\mathbf{n}^M\,\mathrm{d}A = 0 \quad \forall\mathbf{v}\in\mathbf{V}\end{aligned} \tag{25}$$

The 'jump' notation will be particularly useful for the derivation of the Wriggers' approach to the Nitsche method. Also, we can notice that $[\mathbf{v}]\cdot\mathbf{n}^M$ is actually the variation of $g_n(\mathbf{u})$ which we will denote as δg_n[3] For this reason, we can write

$$\delta g_n = \mathbf{v}^S\cdot\mathbf{n}^M - \mathbf{v}^M\cdot\mathbf{n}^M = [\mathbf{v}]\cdot\mathbf{n}^M \tag{26}$$

So (25) can be rewritten as

$$\begin{aligned}&\int_\Omega \boldsymbol{\sigma}\left(\mathbf{u}\right) : \boldsymbol{\nabla}\mathbf{v}\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\mathbf{v}\,\mathrm{d}V - \int_{\Gamma^\sigma}\overline{\mathbf{t}}\cdot\mathbf{v}\,\mathrm{d}A\\&+ \int_{\Gamma^{C,S}}\sigma_n^S\delta g_n\,\mathrm{d}A = 0 \quad \forall\mathbf{v}\in\mathbf{V}\end{aligned} \tag{27}$$

So far, we did not take into consideration the Hertz–Signorini–Moreau conditions (17). To complete the weak formulation of the elastic frictionless contact problem, we need to somehow incorporate them into equation (27). This could be done by reformulating our problem into a variational inequality. We will not provide all the mathematical details but rather suggest the general idea, as presented in Yastrebov [57, Chapter 4.3.1, p.140]. If the Hertz–Signorini–Moreau conditions hold, it could be shown that $g_n = 0$ and $\delta g_n \geq 0$ on $\overline{\Gamma}^C$. At the same time the conditions state that $\sigma_n \leq 0$ on Γ^C, more precisely $\sigma_n \leq 0$ on $\overline{\Gamma}^C$ and $\sigma_n = 0$ on $\Gamma^C \setminus \overline{\Gamma}^C$. This leads us to the conclusion that

$$\int_{\Gamma^{C,S}}\sigma_n\delta g_n\,\mathrm{d}A \leq 0 \tag{28}$$

So it holds

$$\int_\Omega \boldsymbol{\sigma}\left(\mathbf{u}\right) : \boldsymbol{\nabla}\mathbf{v}\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\mathbf{v}\,\mathrm{d}V - \int_{\Gamma^\sigma}\overline{\mathbf{t}}\cdot\mathbf{v}\,\mathrm{d}A \geq 0 \tag{29}$$

It turns out that (27), with the properly defined function space for admissible solution, really constitutes the weak formulation of the elastic frictionless contact problem (16) [57, Chapter 4.3.1, p.140],(17).

$$\int_\Omega \boldsymbol{\sigma}\left(\mathbf{u}\right) : \boldsymbol{\nabla}\mathbf{v}\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\mathbf{v}\,\mathrm{d}V - \int_{\Gamma^\sigma}\overline{\mathbf{t}}\cdot\mathbf{v}\,\mathrm{d}A \geq 0 \quad \forall\mathbf{v}\in\mathbf{K} \tag{30}$$

[3]Because for every fixed $\mathbf{X}\in\Omega_0$, $g_n(\mathbf{u})$ is a functional depending on $\mathbf{u}$, we can calculate its variation δg_n defined as

$$\delta g_n = \frac{\mathrm{d}}{\mathrm{d}\alpha}g_n(\mathbf{u}+\alpha\mathbf{v})\Big|_{\alpha=0} = 0 \quad \forall\mathbf{v}\in\mathbf{V}$$

$$\mathbf{V} = \left\{ \mathbf{v} \in H^1 (\Omega)^d ; \mathbf{v} = \mathbf{0} \text{ on } \Gamma^D \right\} \tag{31}$$

$$\mathbf{K} = \left\{ \mathbf{v} \in \mathbf{V}; \delta g_n = \mathbf{v}^S \cdot \mathbf{n}^M - \mathbf{v}^M \cdot \mathbf{n}^M \geq 0 \text{ on } \overline{\Gamma}^{C,S} \right\} \tag{32}$$

As could be found in [31], the inequality (30) is a so-called variational inequality of the second kind, admits one unique solution (in [15] we can find this is ensured by Stampacchia's Theorem) and is a minimizer to a certain functional. For the details on this formulation, we recommend to see Duvaut and Lions [14] and Kikuchi and Oden [34]. We would like to remark that, as [15] mentions, $\mathbf{f}$ belongs to $L^2 (\Omega)^d$, $\overline{\mathbf{t}}$ belongs to $L^2 (\Gamma^\sigma)^d$ and $\overline{\mathbf{u}}$ belongs to $H^{\frac{3}{2}} \left(\Gamma^D\right)^d$.

3 Unconstrained variational methods for enforcing the contact boundary conditions

In the previous chapter, we formulated our originally strong elastic frictionless contact problem as a variational inequality. Despite the fact that this is mathematically correct, as Yastrebov mentions in [57, Chapter 4.4, p.144]: '...the variational inequality is hard to apply for contact with finite sliding and/or rotations. That is why, nowadays, most of the practical studies in contact mechanics are based on the so-called variational equalities, which are easy to introduce in a finite element framework and does not require totally new minimization techniques.' For this reason, we would prefer to obtain the formulation of our problem as a variational equality. And really, from the theory of optimization it is known that there are certain methods that allow us to find such a formulation. Generally, most of the currently used methods could be divided into two classes [31] - penalty methods and mixed methods. As Yastrebov mentions [57, Chapter 1.1.1, p.7] both of these are implemented in the currently leading finite element software packages ANSYS, ABAQUS and COMSOL. We will focus on a particular class — so-called unconstrained methods. This means that the whole problem is formulated in the form of one equation (variational in our case) and fulfilling of no other conditions is required. This is very convenient and efficient in terms of the implementation of the method, especially in FEniCS.

Penalty methods allow us to replace the variational inequality (30) with a variational equality, in which contact conditions are enforced by adding additional integral over the contact boundary penalising the non-physical penetration. Because this additional integral contains a non-linear function, the Macaulay bracket, for example, the variational problem is non-linear. From the mechanical point of view, this can also be understood as the approximation of contact tractions with some specially chosen function. Even though this class of methods is relatively simple to implement and has a clear mechanical interpretation, penalty methods are generally not consistent. The meaning of consistency will be specified later in the chapter on the mathematical properties of Nitsche-FPR. To at least suggest the sense, let us say that by consistency we mean that if there is a solution $\mathbf{u}$ to a strongly formulated contact problem (16), (17) and $\mathbf{u}$ is sufficiently regular then $\mathbf{u}$ is also solution of our weakly formulated method at a discrete level. However, for penalty methods, this is not generally true as a small penetration is allowed, and contact conditions are therefore not fulfilled exactly (on the level of the continuum formulation of the problem). These methods introduce the penalty parameter. The higher it is, the smaller penetration occurs, which motivates us to use as high penalty as possible. Nevertheless, this is not possible as high penalty values cause ill-conditioning of the resulting discrete problem, which could cause that iterative solvers like Newton's method may fail to converge. In the following chapters, one version of the penalty method will be presented.

Another class of methods are **mixed methods**, for example, the Lagrange multiplier method or the augmented Lagrangian method. Compared to penalty methods, where the only variable was $\mathbf{u}$, in these, a new unknown variable — a Lagrange multiplier — representing the normal contact traction is introduced. This again allows us to reformulate the variational inequality (30) as a variational equality. The formulation is again non-linear but consistent, opposite to the penalty methods. The solution of the weakly formulated problem is a saddle point of the corresponding Lagrangian. For this reason, the Babuška-Brezzi condition must be satisfied at the discrete level to ensure well-posedness. It means that the approximation functions for displacements and Lagrange multipliers must be chosen appropriately, which is not the case for some naive choices. As [31] mentions, to overcome this issue, Barbosa and Hughes [6] proposed a stabilised method which allows us to circumvent the discrete Babuška-Brezzi condition. The drawback of mixed methods is that introducing Lagrange multipliers raises the number of DOFs in a discrete problem. In this thesis, we will mention the augmented Lagrangian method. Another well-known method — the Lagrange multiplier method — will not be analysed

as it is not unconstrained.

Although we will present one version of the penalty method and the augmented Lagrangian method, the main focus of this thesis is on **Nitsche methods**, which represent sort of a 'third' class together with classic penalty and mixed methods. Derived by Joachim A. Nitsche in 1970 [44] for treating interface conditions between the meshes, Nitsche method possesses the beneficial properties of both penalty and mixed methods. The Nitsche methods do not require adding Lagrange multiplier unknowns, but contrary to the penalty methods, they are consistent. The absence of a Lagrange multiplier means that the system of linear equations of the discrete problem is not enlarged compared to penalty methods and we do not have to worry about the Babuška-Brezzi condition. In the following text, two versions of the Nitsche method will be mentioned. One by Wriggers and Zavarise [55, Chapter 5, p.106], [56] and the other, more recent, from 2013 by Chouly, Hild and Renard from [29], [30]. As the methods are not completely the same, the differences will be analysed.

All the presented methods will be considered for the case of bilateral contact, i.e. the contact of two elastic bodies [57, Chapter 1.1, p.5].

3.1 Illustration of the considered methods—a one-dimensional example

Before proceeding to the analysis of particular methods, we would like to provide the reader with a brief overview of all the methods applied to a simple one-dimensional example. This is purely for the comfort of the reader, so he/she could grasp the general idea of each method before we discuss its formulation in the three-dimensional continuum. We will analyse the following example with two degrees of freedom u_1 and u_2:

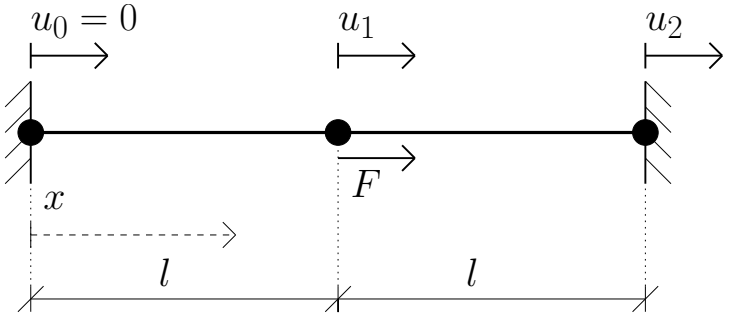


Figure 2: Geometry of the 1D example

As Figure 2 suggests, the structure is a beam of length $2l$ with a constant circular cross-section of diameter h. This is discretised by two linear 1D elements of length l. To enforce contact in node 2, force F is applied at node 1. Node 0 is fixed. We presume that there is a rigid obstacle next to the node 2, where a contact occurs. Obviously, it must hold $u_2 = 0$, but we do not prescribe this condition directly. Instead of that, we will enforce the contact boundary condition $u_2 \leq 0$ with various methods for enforcing contact the contact boundary conditions and we will present the gained results for every method. By these means, we illustrate the nature of every method and give the reader a notion what she/he could expect.

The gap function $g_n(\mathbf{u})$ and the contact traction $\sigma_n(\mathbf{u})$ could be now conveniently expressed as

$$g_n(\mathbf{u}) = -u_2 \tag{33}$$

$$\sigma_n(\mathbf{u}) = \frac{EA}{l}(u_2 - u_1) \tag{34}$$

where the area of beam cross-section A and Young's modulus E are considered to be constant over the whole structure.

The variations of $g_n(\mathbf{u})$ and $\sigma_n(\mathbf{u})$ are expressed as

$$\delta g_n = -\delta u_2 \tag{35}$$

$$\delta\sigma_n = \sigma_n(\delta\mathbf{u}) = \frac{EA}{l}(\delta u_2 - \delta u_1) \tag{36}$$

Compared to the notation in the other chapters, for the sake of clarity, the variation of $\mathbf{u}$ is denoted as $\delta\mathbf{u}$ instead of $\mathbf{v}$. Also, we would like to note that in the following weak forms, the average of stress - $\langle\sigma_n(\mathbf{u})\rangle_t$ in Nitsche-FPR or $\langle\sigma_n(\mathbf{u})\rangle$ in Nitsche-Wriggers, is replaced with mere σ_n. Formally, we can define the normal contact tractions on the active contact zone of the rigid plane to be equal to the ones on the active contact zone of the elastic body. This additional definition of the tractions on the rigid body is correct, because the displacement of the rigid body is zero in every material point, but the stress and tractions can be arbitrary. Equipped with these definitions of tractions, we can formally apply the average and jump operators. For all the methods, we will focus on the case when contact occurs.

For all the methods listed below, the terms in weak forms connected with elastic energy and external loads are discretised in the following way

$$\begin{aligned}&\int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\varepsilon}(\delta\mathbf{u})\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\delta\mathbf{u}\,\mathrm{d}V - \int_{\Gamma^\sigma} \overline{\mathbf{t}}\cdot\delta\mathbf{u}\,\mathrm{d}A \\ &= \frac{EA}{l}u_1\delta u_1 + \frac{EA}{l}(u_2-u_1)(\delta u_2-\delta u_1) - F\delta u_1\end{aligned} \tag{37}$$

Also, we would like to mention the following functions and their definitions, as they are defined later in the text when the particular methods are derived. We will utilize the 'ramp' function $\{x\}$ defined using the Macaulay bracket and the Heaviside function $H(x)$. These are defined as

$$\{x\} := \frac{x+|x|}{2} = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{38}$$

$$H(x) := \frac{1}{2}(1+\mathrm{sgn}(x)) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{39}$$

$$\mathrm{sgn}\,x := \begin{cases} 1, & x \geq 0 \\ -1, & x < 0 \end{cases} \tag{40}$$

Also, for a simpler notation, let us introduce the beam stiffness in tension/compression

$$k = \frac{EA}{l} \tag{41}$$

and the penalty parameter

$$P = \frac{\gamma E}{l}, \quad \gamma > 0 \tag{42}$$

3.1.1 The Nitsche-Wriggers method (113)

The weak form of the Nitsche-Wriggers method (113) reads

$$\begin{aligned}&\int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\varepsilon}(\delta\mathbf{u})\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\delta\mathbf{u}\,\mathrm{d}V - \int_{\Gamma^\sigma} \overline{\mathbf{t}}\cdot\delta\mathbf{u}\,\mathrm{d}A\\ &+ \int_{\Gamma^{C,S}} \sigma_n(\mathbf{u})H(-g_n(\mathbf{u}))\delta g_n(\delta\mathbf{u})\,\mathrm{d}A - \int_{\Gamma^{C,S}} \sigma_n(\delta\mathbf{u})\{-g_n(\mathbf{u})\}\,\mathrm{d}A\\ &- \frac{\gamma E}{h}\int_{\Gamma^{C,S}} \{-g_n(\mathbf{u})\}\delta g_n(\delta\mathbf{u})\,\mathrm{d}A = 0 \qquad \forall\delta\mathbf{u}\in\mathbf{V}\end{aligned} \tag{43}$$

Discretising this variational equation in a way described above, we obtain

$$\begin{aligned}&ku_1\delta u_1 + k(u_2-u_1)(\delta u_2-\delta u_1) - F\delta u_1\\ &- k\,(u_2-u_1)\,\delta u_2 - k\,(\delta u_2-\delta u_1)\,u_2\\ &+ Pu_2\delta u_2 = 0 \qquad \forall\delta u_1,\delta u_2\in\mathbb{R}\end{aligned} \tag{44}$$

This can be separated into two equations which read

$$\begin{aligned}ku_1\delta u_1 - k(u_2-u_1)\delta u_1 - F\delta u_1 + ku_2\delta u_1 = 0 \qquad \forall\delta u_1\in\mathbb{R}\\ k(u_2-u_1)\delta u_2 - k(u_2-u_1)\delta u_2 - ku_2\delta u_2 + Pu_2\delta u_2 = 0 \qquad \forall\delta u_2\in\mathbb{R}\end{aligned} \tag{45}$$

which means

$$\begin{aligned}ku_1 - k(u_2-u_1) - F + ku_2 = 0\\ k(u_2-u_1) - k(u_2-u_1) - ku_2 + Pu_2 = 0\end{aligned} \tag{46}$$

This is equivalent to the matrix equation

$$\begin{pmatrix}2k & 0\\ 0 & P-k\end{pmatrix}\begin{pmatrix}u_1\\ u_2\end{pmatrix} = \begin{pmatrix}F\\ 0\end{pmatrix} \tag{47}$$

The solution of this linear system is

$$\begin{aligned}u_1 &= \frac{F}{2k}\\ u_2 &= 0\end{aligned} \tag{48}$$

which is the physically correct solution that we expected. Therefore, we see that the contact condition $u_2 \le 0$ is fulfilled exactly and what is particularly important, irrespectively of the penalty value P. This differentiates the Nitsche-Wriggers method from the penalty method described in the following section, where the contact boundary condition is never fulfilled exactly for any finite penalty value.

Even though the terms integrated over the contact surface in the weak form of the Nitsche-Wriggers method could seem somewhat tricky, the method itself is very natural from the mechanical point of view. Firstly, if we forget about the possibility of contact, the linear system describing our problem reads

$$\begin{pmatrix}2k & -k\\ -k & k\end{pmatrix}\begin{pmatrix}u_1\\ u_2\end{pmatrix} = \begin{pmatrix}F\\ 0\end{pmatrix} \tag{49}$$

This is the discretised version of the equation

$$\int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\varepsilon}(\delta\mathbf{u})\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\delta\mathbf{u}\,\mathrm{d}V - \int_{\Gamma^\sigma} \overline{\mathbf{t}}\cdot\delta\mathbf{u}\,\mathrm{d}A = 0 \qquad \forall\delta\mathbf{u}\in\mathbf{V} \tag{50}$$

The penalty method, described in the following section, would approximately enforce the contact boundary conditions by adding the penalty P to the position $(2,2)$ in the stiffness matrix. This is equivalent to adding

a spring with stiffness P to the node 2. On the continuum level, this is achieved by adding the following zero-valued term to the weak form

$$-\frac{\gamma E}{h}\int_{\Gamma^{C,S}}\{-g_n(\mathbf{u})\}\delta g_n(\delta\mathbf{u})\,\mathrm{d}A \tag{51}$$

However, this approach, despite being very simple, ignores the contribution of the integrals over the potential contact zone and is therefore inconsistent. These missing terms are represented by the integral expressions present in the weak form of the Nitsche-Wriggers method which reads

$$\int_{\Gamma^{C,S}}\sigma_n(\mathbf{u})H(-g_n(\mathbf{u}))\delta g_n(\delta\mathbf{u})\,\mathrm{d}A \tag{52}$$

It turns out that in terms of this example, considering these missing contributions is equivalent with adding such terms to the second row of the stiffness matrix that the second row becomes zero-valued. The whole linear system then reads

$$\begin{pmatrix}2k & -k\\ 0 & 0\end{pmatrix}\begin{pmatrix}u_1\\ u_2\end{pmatrix}=\begin{pmatrix}F\\ 0\end{pmatrix} \tag{53}$$

However, u_2 could now be arbitrary. The Nitsche-Wriggers method now does the same as the penalty method - add the penalty P to the position $(2,2)$ in the stiffness matrix. We see that if the correct solution $u_2 = 0$ is obtained by this procedure, adding any value of P[4] does not change the solution. On the continuum level, this is achieved by adding the same integral term as in the penalty method, which reads

$$-\frac{\gamma E}{h}\int_{\Gamma^{C,S}}\{-g_n(\mathbf{u})\}\delta g_n(\delta\mathbf{u})\,\mathrm{d}A \tag{54}$$

In terms of our discrete problem we get

$$\begin{pmatrix}2k & -k\\ 0 & P\end{pmatrix}\begin{pmatrix}u_1\\ u_2\end{pmatrix}=\begin{pmatrix}F\\ 0\end{pmatrix} \tag{55}$$

We see that we obtain a unique solution, which is the same as (48). We can observe that the resulting stiffness matrix is not symmetric. This could be rectified by adding another zero valued term to the weak form, which reads

$$-\int_{\Gamma^{C,S}}\sigma_n(\delta\mathbf{u})\{-g_n(\mathbf{u})\}\,\mathrm{d}A \tag{56}$$

For our discrete problem, it means changing the second column by adding k to the position $(1,2)$ and $-k$ to the position $(2,2)$. This does not affect the solution, as we know that $u_2 = 0$. With this we obtain

$$\begin{pmatrix}2k & 0\\ 0 & P-k\end{pmatrix}\begin{pmatrix}u_1\\ u_2\end{pmatrix}=\begin{pmatrix}F\\ 0\end{pmatrix} \tag{57}$$

which is exactly the same as (47). We can observe now that if we choose $P = k$, the stiffness matrix becomes singular and the value of u_2 is arbitrary. In order to avoid this, we must choose penalty value such that

$$\begin{aligned}P &\neq k\\ \frac{E\gamma}{l} &\neq \frac{EA}{l}\end{aligned} \tag{58}$$

and so

$$\gamma \neq A \tag{59}$$

This indicates that there are certain limitations on the choice of γ. In the case of more complicated two-dimensional and three-dimensional problems it turns out that there is a certain minimal penalty value called C, where $C > 0$.

[4]We remind P is supposed to be positive.

3.1.2 The penalty method (157)

$$\begin{aligned}&\int_\Omega \boldsymbol{\sigma}\,(\mathbf{u}) : \boldsymbol{\nabla}\delta\mathbf{u}\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\delta\mathbf{u}\,\mathrm{d}V - \int_{\Gamma^\sigma} \overline{\mathbf{t}}\cdot\delta\mathbf{u}\,\mathrm{d}A\\ &- \frac{\gamma E}{h}\int_{\Gamma^{C,S}}\{-g_n(\mathbf{u})\}\delta g_n(\delta\mathbf{u})\,\mathrm{d}A = 0 \quad \forall\delta\mathbf{u}\in\mathbf{V}\end{aligned} \tag{60}$$

Discretising this form gives

$$\begin{aligned}&ku_1\delta u_1 + k(u_2-u_1)(\delta u_2-\delta u_1) - F\delta u_1\\ &+ Pu_2\delta u_2 = 0 \quad \forall\delta u_1,\delta u_2\in\mathbf{V}\end{aligned} \tag{61}$$

or in a matrix form

$$\begin{pmatrix}2k & -k\\ -k & P+k\end{pmatrix}\begin{pmatrix}u_1\\ u_2\end{pmatrix} = \begin{pmatrix}F\\ 0\end{pmatrix} \tag{62}$$

If we have a look at the weak form (60), we can see that it contains the standard elastic terms and one additional term which reads

$$-\frac{\gamma E}{h}\int_{\Gamma^{C,S}}\{-g_n(\mathbf{u})\}\delta g_n(\delta\mathbf{u})\,\mathrm{d}A \tag{63}$$

Let us firstly consider only the standard elastic terms in equation (60). That means considering the following problem

$$\int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\varepsilon}(\delta\mathbf{u})\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\delta\mathbf{u}\,\mathrm{d}V - \int_{\Gamma^\sigma}\overline{\mathbf{t}}\cdot\delta\mathbf{u}\,\mathrm{d}A = 0 \quad \forall\delta\mathbf{u}\in\mathbf{V} \tag{64}$$

If we discretise this weak form, we obtain

$$\begin{pmatrix}2k & -k\\ -k & k\end{pmatrix}\begin{pmatrix}u_1\\ u_2\end{pmatrix} = \begin{pmatrix}F\\ 0\end{pmatrix} \tag{65}$$

This formulation does not take into consideration the contact condition in node 2 at all. The penalty method introduces the contact to the equation by adding an additional term

$$-\frac{\gamma E}{h}\int_{\Gamma^{C,S}}\{-g_n(\mathbf{u})\}\delta g_n(\delta\mathbf{u})\,\mathrm{d}A \tag{66}$$

In terms of our discrete problem, this is equivalent to adding the penalty P to the position $(2,2)$ in the stiffness matrix. This could be also understood as adding a spring with stiffness P to the node 2 which resists the penetration. The great advantage of the penalty method is that it is very simple. We just need to find an elastic stiffness matrix and add some additional stiffness terms to it. However, a disadvantage of this method is apparent from the character of solution, which reads

$$\begin{aligned}u_1 &= \frac{F(k+P)}{k(k+2P)}\\ u_2 &= \frac{F}{k+2P}\end{aligned} \tag{67}$$

We can immediately spot that for any finite valued penalty parameter P, $u_2 \neq 0$. However, the higher the penalty parameter P is, the smaller is the error. The physically correct solution $u_2 = 0$ is obtained if $P\to\infty$ as

$$\begin{aligned}\lim_{P\to\infty} u_1 &= \frac{F}{2k}\\ \lim_{P\to\infty} u_2 &= 0\end{aligned} \tag{68}$$

In the case of our one-dimensional example we can of course choose the arbitrarily large (but finite) penalty parameter P and obtain very precise approximation of the physical solution with no trouble. However, for large-scale two-dimensional or three-dimensional problems, choosing too high penalty parameter causes ill-conditioning of the linear system. Therefore, the maximal precision of the penalty method is limited.

3.1.3 The Nitsche-FPR method(134)

$$\begin{aligned}
&\int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\nabla}\delta\mathbf{u}\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\delta\mathbf{u}\,\mathrm{d}V - \int_{\Gamma^\sigma}\overline{\mathbf{t}}\cdot\delta\mathbf{u}\,\mathrm{d}A \\
&+ \int_{\Gamma^{C,S}} \theta\frac{h}{\gamma E}\sigma_n(\mathbf{u})\sigma_n(\delta\mathbf{u})\,\mathrm{d}A \\
&- \int_{\Gamma^{C,S}} \frac{\gamma E}{h}\{-g_n(\mathbf{u}) - \frac{h}{\gamma E}\sigma_n(\mathbf{u})\}\left(\delta g_n(\delta\mathbf{u}) - \theta\frac{h}{\gamma E}\sigma_n(\delta\mathbf{u})\right)\mathrm{d}A = 0 \quad \forall\delta\mathbf{u}\in\mathbf{V}
\end{aligned} \tag{69}$$

Discretising the form above we get

$$\begin{aligned}
&ku_1\delta u_1 + k(u_2-u_1)(\delta u_2-\delta u_1) - F\delta u_1 \\
&+\theta\frac{k^2}{P}(u_2-u_1)(\delta u_2-\delta u_1) \\
&- P\left(u_2 - \frac{k}{P}(u_2-u_1)\right)\left(-\delta u_2 - \theta\frac{k}{P}(\delta u_2-\delta u_1)\right) = 0 \quad \forall\delta u_1,\delta u_2\in\mathbf{V}
\end{aligned} \tag{70}$$

The parameter θ could be chosen from $[-1,1]$, however, as [15] mentions, three choices are of a special interest, namely $\theta=-1$, $\theta=0$ and $\theta=1$. Then the matrix form of (70) reads

If $\theta=-1$

$$\begin{pmatrix} 2k & 0 \\ 0 & P-k \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} F \\ 0 \end{pmatrix} \tag{71}$$

If $\theta=1$

$$\begin{pmatrix} 2k & -2k \\ 0 & P+k \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} F \\ 0 \end{pmatrix} \tag{72}$$

If $\theta=0$

$$\begin{pmatrix} 2k & -k \\ 0 & P \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} F \\ 0 \end{pmatrix} \tag{73}$$

In all cases, we obtain the physically correct solution we required

$$\begin{aligned} u_1 &= \frac{F}{2k} \\ u_2 &= 0 \end{aligned} \tag{74}$$

The Nitsche-FPR method is very similar to the Nitsche-Wriggers method and for $\theta=-1$ we obtain exactly the same symmetric stiffness matrix as in the case of the Nitsche-Wriggers method (see (47)). For $\theta=1$ and $\theta=0$, the stiffness matrix is non-symmetric. Concerning our 2x2 matrix it does not matter, but generally, the solution of a large non-symmetric system could be much more demanding compared to its symmetric counterpart. On the other hand, the non-symmetric cases have one advantage in our case. This is that there is no risk of choosing P such that the matrix becomes singular, as could happen if $\theta=-1$ when the stiffness matrix is symmetric.

3.1.4 The augmented Lagrangian method (177), (178)

$$\begin{aligned} &\int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\varepsilon}(\delta\mathbf{u})\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\delta\mathbf{u}\,\mathrm{d}V - \int_{\Gamma^\sigma} \overline{\mathbf{t}}\cdot\delta\mathbf{u}\,\mathrm{d}A \\ - &\int_{\Gamma^C} \frac{\gamma E}{h}\left\{-g_n(\mathbf{u}) - \frac{h}{\gamma E}\lambda\right\}\delta g_n(\delta\mathbf{u})\,\mathrm{d}A = 0 \quad \forall\delta\mathbf{u}\in\mathbf{V} \end{aligned} \tag{75}$$

$$-\int_{\Gamma^C}\frac{h}{\gamma E}\left(\lambda\delta\lambda + \frac{\gamma E}{h}\left\{-g_n(\mathbf{u}) - \frac{h}{\gamma E}\lambda\right\}\delta\lambda\right)\,\mathrm{d}A = 0 \quad \forall\delta\lambda\in\mathbb{R} \tag{76}$$

Again, let us now discretise this weak form

$$\begin{aligned} &ku_1\delta u_1 + k(u_2-u_1)(\delta u_2-\delta u_1) - F\delta u_1 \\ &+ P\left(u_2 - \frac{1}{P}\lambda\right)\delta u_2 = 0 \quad \forall\delta u_1,\delta u_2\in\mathbf{V} \end{aligned} \tag{77}$$

$$-\frac{1}{P}\left(\lambda\delta\lambda + P\left(u_2-\frac{1}{P}\lambda\right)\delta\lambda\right) = 0 \quad \forall\delta\lambda\in\mathbb{R} \tag{78}$$

This is equivalent to the matrix equation

$$\begin{pmatrix} 2k & -k & 0 \\ -k & P+k & -1 \\ 0 & -1 & 0 \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \\ \lambda \end{pmatrix} = \begin{pmatrix} F \\ 0 \\ 0 \end{pmatrix} \tag{79}$$

The solution of this linear system reads

$$\begin{aligned} u_1 &= \frac{F}{2k} \\ u_2 &= 0 \\ \lambda &= -\frac{F}{2} \end{aligned} \tag{80}$$

which is exactly the physically precise solution we sought. We can observe that the Lagrange multiplier λ was introduced as an independent unknown. The physical meaning is generally the normal contact traction. In our one-dimensional case, it is the contact force. Despite the introduction of an additional unknown eventually provides us with the solution we wanted, the whole linear system in now enlarged. This could become a problem in the case of large-scale problems.

3.2 The Nitsche method

This method is named in honour of its original author, German mathematician and professor of applied mathematics at the University of Freiburg, Joachim A. Nitsche [2]. He was the first one to describe this method in his German-written paper [44] in 1970, where he focused on treating the Dirichlet boundary or interface conditions in a weak sense. Twelve years later, this approach was also investigated by Arnold [4]. As [31] reports, the Nitsche method is different from standard penalisation and mixed techniques. Firstly it is consistent[5] In this aspect, the Nitsche method differs from standard penalisation techniques, which generally lack consistency [34]. Also, the Nitsche method does not use any additional unknown (a Lagrange multiplier), and so the Babuška-Brezzi condition does not have to be fulfilled, contrarily to mixed methods.

Although Nitsche method was not originally applied to treating contact conditions, it gained considerable popularity for other types of boundary conditions, mostly for linear conditions on the boundary of a domain or at the interface between sub-domains [31]. As Wriggers and Zavarise mention [56], the Nitsche approach was investigated by Stenberg to weakly enforce Dirichlet boundary conditions [53] and by Becker, Hansbo and Stenberg [7], [8] within domain decomposition methods with non-matching grids. Some review of the application of the Nitsche method to interface problems could also be found in [24]. Augarde and co-authors [37] state that the Nitsche method was also used to deal with decomposed domains in isogeometric analysis [3], [48] and discontinuous elements in the discontinuous Galerkin method [22].

As [31] mentions, the first attempts to apply the Nitsche method to contact emerged nearly three decades after the discovery. These concerned a bilateral contact and were done by Hansbo, A., Hansbo, P. and co-authors [25], [26] and by Wriggers and Zavarise [56]. In 2013 the new formulation of the Nitsche method by Chouly, Hild and Renard emerged [29], [30]. The authors provide a very consistent and thorough mathematical analysis of this method for Signorini's problem. We remind this is a unilateral problem where a linear elastic body is in frictionless contact with a rigid foundation. This was generalized to the problem of the contact of two elastic bodies by Fabré, Renard and Pousin in [15].

Although the Nitsche method gained a considerable popularity as the weak way of imposing some types of boundary conditions (for example Dirichlet boundary conditions or the conditions at the interface between sub-domains of the mesh) in the finite element method, up to the recent past it was not much considered for imposing contact boundary conditions. This changed in 2013 due to the works of Chouly, Hild and Renard [29], [30], but still, to the knowledge of the authors of this thesis, the Nitsche method is only rarely considered when solving contact problems, as [31] confirms. However, from the theoretical point of view, it seems that there is no apparent reason why this remarkable method should not be investigated and used in engineering practice. Although Konyukhov and Izi [35, Chapter 6, p.83] comment that: '...the necessity to implement the whole structural finite element (because of p_n[6]) positions the Nitsche method among the very seldomly used methods in practical FE codes', this seems more as an implementation problem than the drawback of the method itself.

The beauty of the Nitsche method lies primarily in two facts. Firstly, it is consistent compared to penalty methods, and theoretically (on the level of the continuum problem) the solution does not depend on the value of the penalty parameter. Also, a considerably smaller penalty value could be used, so the ill-conditioning of the problem is avoided. Secondly, it does not require adding any additional unknown (a Lagrange multiplier), so the resulting system of equations solved in the iteration steps of the finite element discretised problem is not

[5] We would like to mention that by the consistency of the method we mean that if there is a solution $\mathbf{u}$ to the strongly formulated contact problem and $\mathbf{u}$ is sufficiently regular then $\mathbf{u}$ is also the solution to the weak problem formulated with our method.

[6] By p_n Konyukhov and Izi mean the normal contact traction, referred as σ_n in this thesis

enlarged, and the Babuška-Brezzi condition does not have to be fulfilled. These properties make the method interesting not only for small-scale but especially for large-scale problems where both the ill-conditioning and the enlarging of the solved system must be avoided as much as possible. Some numerical properties of the Nitsche method will be described in this chapter.

In this thesis, we will present two previously mentioned approaches to imposing contact boundary conditions with the Nitsche method. The first one is by Wriggers, described in [55, Chapter 5, p.106] and his paper with Zavarise [56]. Wriggers' approach follows, as he comments in [56], the original paper of Nitsche from 1970 [44]. The second one is based on the paper [15], which follows the modern approach of Chouly, Hild and Renard from [29], [30]. Interestingly, the weak form derived in [15] does not coincide with the one derived by Wriggers. To our knowledge, the mutual connection of both methods has not been described so far. In this chapter, we will try to provide some comments on this.

Before analysing the chosen Nitsche methods, we would like to mention that the Nitsche method could also be used for the weak enforcement of Dirichlet boundary conditions, which was one of its original applications. However, for the sake of simplicity, we will suppose that Dirichlet conditions are enforced in a standard way. For some remarks and literature on enforcing Dirichlet boundary conditions with the Nitsche method see [37].

3.2.1 Derivation of Wriggers' approach to the Nitsche method (Nitsche-Wriggers)

Though Wriggers and co-authors state the functional of elastic energy and the corresponding weak form for the elastic frictionless contact problem in [55, Chapter 5, p.106] or [56], they do not clearly state, how the mentioned could be consistently derived. The authors of [56] state that they took inspiration directly from the original paper of Nitsche (author of this method) [44]. Nevertheless, Nitsche's paper [44] does not consider the problem of enforcing contact boundary conditions. Wriggers and Zavarise are also not giving any details on how the method could be consistently derived when treating contact boundary conditions. Therefore, in this section, we will present an approach based on the classical understanding of the Nitsche method from [44] leading to the functional of elastic energy and the corresponding weak form presented by Wriggers and co-authors. For shorter notation, this version is referred to as Nitsche-Wriggers.

We will build on equation (25). However, instead of integrating the contact terms over the potential contact zone $\Gamma^{C,S}$ we will firstly integrate them only over the active contact zone $\overline{\Gamma}^{C,S}$. After the equation of the Nitsche-Wriggers method will be derived, the integral over $\overline{\Gamma}^{C,S}$ will be extended to the integral over $\Gamma^{C,S}$ using the Macaulay bracket formalism. The derivation starts with equation

$$\begin{aligned}&\int_{\Omega} \boldsymbol{\sigma}\left(\mathbf{u}\right) : \boldsymbol{\nabla}\mathbf{v}\,\mathrm{d}V - \int_{\Omega} \rho\overline{\mathbf{b}}\cdot\mathbf{v}\,\mathrm{d}V - \int_{\Gamma^{\sigma}} \overline{\mathbf{t}}\cdot\mathbf{v}\,\mathrm{d}A \\ &+ \int_{\overline{\Gamma}^{C,S}} \sigma_n^S\left[\mathbf{v}\right]\cdot\mathbf{n}^M\,\mathrm{d}A = 0 \qquad \forall\mathbf{v}\in\mathbf{V}\end{aligned} \tag{81}$$

We will utilize the following formalism adopted from the classical derivation of the Discontinuous Galerkin method. For easier notation, let us define the jump $[\bullet]$ and the average $\langle\bullet\rangle$ operators for any scalar-valued of vector-valued function $\mathbf{a}$ on Γ^C as

$$[\mathbf{a}] = \mathbf{a}^S - \mathbf{a}^M \tag{82}$$

$$\langle\mathbf{a}\rangle = \frac{1}{2}\left(\mathbf{a}^S + \mathbf{a}^M\right) \tag{83}$$

In the following text, we will utilize the expressions listed below

$$[\sigma_n\mathbf{v}]\cdot\mathbf{n}^M = \sigma_n^S\mathbf{v}^S\cdot\mathbf{n}^M - \sigma_n^M\mathbf{v}^M\cdot\mathbf{n}^M \tag{84}$$

$$[\mathbf{v}]\cdot\mathbf{n}^M = \mathbf{v}^S\cdot\mathbf{n}^M - \mathbf{v}^M\cdot\mathbf{n}^M \tag{85}$$

$$[\sigma_n] = \sigma_n^S - \sigma_n^M \tag{86}$$

and

$$\langle\sigma_n\mathbf{v}\rangle\cdot\mathbf{n}^M = \frac{1}{2}\left(\sigma_n^S\mathbf{v}^S\cdot\mathbf{n}^M + \sigma_n^M\mathbf{v}^M\cdot\mathbf{n}^M\right) \tag{87}$$

$$\langle\mathbf{v}\rangle\cdot\mathbf{n}^M = \frac{1}{2}\left(\mathbf{v}^S\cdot\mathbf{n}^M + \mathbf{v}^M\cdot\mathbf{n}^M\right) \tag{88}$$

$$\langle\sigma_n\rangle = \frac{1}{2}\left(\sigma_n^S + \sigma_n^M\right) \tag{89}$$

It can be verified that the for any scalar-valued or vector-valued functions $\mathbf{a}$ and $\mathbf{b}$ on Γ^C holds

$$[\mathbf{a}\cdot\mathbf{b}] = [\mathbf{a}]\langle\mathbf{b}\rangle + \langle\mathbf{a}\rangle[\mathbf{b}] \tag{90}$$

as

$$\begin{aligned}&[\mathbf{a}]\langle\mathbf{b}\rangle + \langle\mathbf{a}\rangle[\mathbf{b}] = \\ &= \left(\mathbf{a}^S - \mathbf{a}^M\right)\frac{1}{2}\left(\mathbf{b}^S + \mathbf{b}^M\right) + \frac{1}{2}\left(\mathbf{a}^S + \mathbf{a}^M\right)\left(\mathbf{b}^S - \mathbf{b}^M\right) \\ &= \frac{1}{2}\mathbf{a}^S\cdot\mathbf{b}^S + \frac{1}{2}\mathbf{a}^S\cdot\mathbf{b}^M - \frac{1}{2}\mathbf{a}^M\cdot\mathbf{b}^S - \frac{1}{2}\mathbf{a}^M\cdot\mathbf{b}^M + \\ &+ \frac{1}{2}\mathbf{a}^S\cdot\mathbf{b}^S - \frac{1}{2}\mathbf{a}^S\cdot\mathbf{b}^M + \frac{1}{2}\mathbf{a}^M\cdot\mathbf{b}^S - \frac{1}{2}\mathbf{a}^M\cdot\mathbf{b}^M \\ &= \mathbf{a}^S\cdot\mathbf{b}^S - \mathbf{a}^M\cdot\mathbf{b}^M\end{aligned} \tag{91}$$

For this reason, the following identity holds, and we will utilize it in the derivation of the Nitsche-Wriggers method

$$[\sigma_n \mathbf{v}] \cdot \mathbf{n}^M = [\sigma_n] \langle \mathbf{v} \rangle \cdot \mathbf{n}^M + \langle \sigma_n \rangle [\mathbf{v}] \cdot \mathbf{n}^M \tag{92}$$

Now we can substitute (92) to equation (81), as $\sigma_n^S [\mathbf{v}] \cdot \mathbf{n}^M = \sigma_n^S \mathbf{v}^S \cdot \mathbf{n}^M - \sigma_n^S \mathbf{v}^M \cdot \mathbf{n}^M = \sigma_n^S \mathbf{v}^S \cdot \mathbf{n}^M - \sigma_n^M \mathbf{v}^M \cdot \mathbf{n}^M = [\sigma_n \mathbf{v}] \cdot \mathbf{n}^M$ on $\overline{\Gamma}^C$

$$\begin{aligned} &\int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\nabla} \mathbf{v} \, \mathrm{d}V - \int_\Omega \rho \overline{\mathbf{b}} \cdot \mathbf{v} \, \mathrm{d}V - \int_{\Gamma^\sigma} \overline{\mathbf{t}} \cdot \mathbf{v} \, \mathrm{d}A \\ &+ \int_{\overline{\Gamma}^{C,S}} \left([\sigma_n] \langle \mathbf{v} \rangle \cdot \mathbf{n}^M + \langle \sigma_n \rangle [\mathbf{v}] \cdot \mathbf{n}^M \right) \mathrm{d}A = 0 \quad \forall \mathbf{v} \in \mathbf{V} \end{aligned} \tag{93}$$

However, from equation (14) we know that

$$[\sigma_n(\mathbf{u})] = \sigma_n^S(\mathbf{u}) - \sigma_n^M(\mathbf{u}) = 0 \text{ on } \overline{\Gamma}^C \tag{94}$$

and so

$$\begin{aligned} &\int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\nabla} \mathbf{v} \, \mathrm{d}V - \int_\Omega \rho \overline{\mathbf{b}} \cdot \mathbf{v} \, \mathrm{d}V - \int_{\Gamma^\sigma} \overline{\mathbf{t}} \cdot \mathbf{v} \, \mathrm{d}A \\ &+ \int_{\overline{\Gamma}^{C,S}} \langle \sigma_n(\mathbf{u}) \rangle [\mathbf{v}] \cdot \mathbf{n}^M \, \mathrm{d}A = 0 \quad \forall \mathbf{v} \in \mathbf{V} \end{aligned} \tag{95}$$

Although it could be tempting to discretise and numerically solve our problem based on the variational formulation given in (95), we cannot successfully do that. If we understand the following integral expression from (95) as a bilinear form $a(\mathbf{u}, \mathbf{v})$, it could be verified that $a(\mathbf{u}, \mathbf{v})$ is not symmetric and positive definite.

$$a(\mathbf{u}, \mathbf{v}) = \int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\nabla} \mathbf{v} \, \mathrm{d}V + \int_{\overline{\Gamma}^{C,S}} \langle \sigma_n(\mathbf{u}) \rangle [\mathbf{v}] \cdot \mathbf{n}^M \, \mathrm{d}A \tag{96}$$

For this reason, it would not be ensured that there exist a unique solution to our problem. The lack of symmetry also means that there would not be any corresponding energy functional to our problem, such that the variational equation of the Nitsche-Wriggers method would be its stationarity condition. This is somehow philosophically discomforting. Also, it is widely known that the numerical solution of a non-symmetric linear system is usually far more computationally demanding than the solution of its symmetric counterpart [20]. For these reasons, we will adopt the classical approach to the Nitsche method [44] and add some terms to equation (95). These additional terms will be zero-valued (in the Lebesgue integral sense) if being evaluated for the sought solution $\mathbf{u}$ (when $g_n(\mathbf{u}) = 0$), and therefore does not change the solution in any way. As suggested, the additional terms will also change the bilinear form $a(\mathbf{u}, \mathbf{v})$ in such a way that it will eventually become symmetric and positive-definite.

Nevertheless, it should be pointed out that we do not necessarily need a symmetric bilinear form to ensure the existence of a unique solution and so, non-symmetric versions of the Nitsche-Wriggers method could be constructed. Even though the numerical solution of the discretised problem could be more complicated, the non-symmetric versions of the Nitsche method could possess some interesting properties, as will be demonstrated in the following chapter on Nitsche-FPR method. However, we will investigate only the symmetric version of the Nitsche-Wriggers method in this thesis.

The bilinear part of

$$\int_{\overline{\Gamma}^{C,S}} \langle \sigma_n(\mathbf{v}) \rangle \left([\mathbf{u}] \cdot \mathbf{n}^M + [\mathbf{X}] \cdot \mathbf{n}^M \right) \mathrm{d}A = \int_{\overline{\Gamma}^{C,S}} \langle \sigma_n(\mathbf{v}) \rangle g_n(\mathbf{u}) \, \mathrm{d}A = 0 \tag{97}$$

will restore symmetry. When adding this integral to equation (95), the solution of the equation is not changed. This is because if $\mathbf{u}$ is the solution, then $g_n(\mathbf{u}) = 0$ on $\overline{\Gamma}^C$ and so (97) is zero. We remind here the definition of the gap function

$$g_n(\mathbf{u}) = \mathbf{X}^S \cdot \mathbf{n}^M + \mathbf{u}^S \cdot \mathbf{n}^M - \mathbf{X}^M \cdot \mathbf{n}^M - \mathbf{u}^M \cdot \mathbf{n}^M = [\mathbf{u}] \cdot \mathbf{n}^M + [\mathbf{X}] \cdot \mathbf{n}^M$$

Let us remark, that although we consider a positive sign in the expression (97), which is necessary to preserve the symmetry of our problem, a negative sign could also be considered and the resulting bilinear form is then non-symmetric.

Finally, the bilinear part of

$$\frac{\gamma E}{h}\int_{\overline{\Gamma}^{C,S}}\left([\mathbf{u}]\cdot\mathbf{n}^M+[\mathbf{X}]\cdot\mathbf{n}^M\right)[\mathbf{v}]\cdot\mathbf{n}^M\,\mathrm{d}A=\frac{\gamma E}{h}\int_{\overline{\Gamma}^{C,S}}g_n(\mathbf{u})\delta g_n(\mathbf{v})\,\mathrm{d}A=0 \tag{98}$$

secures the positive definiteness of the bilinear form $a(\mathbf{u},\mathbf{v})$, given the sufficiently large penalty parameter γ. E/h is the scaling parameter of the penalty term so that γ is dimensionless. E is Young's modulus and h is the characteristic size of the mesh elements (defined for example as the average of the largest and the smallest diameter of the sphere circumscribed to the elements). h is considered to be constant over the mesh. In equation (98) we utilize the fact that

$$g_n(\mathbf{u})=\left(\mathbf{x}^S-\mathbf{x}^M\right)\cdot\mathbf{n}^M=\left(\mathbf{X}^S+\mathbf{u}^S-\mathbf{X}^M-\mathbf{u}^M\right)\cdot\mathbf{n}^M \tag{99}$$

and

$$\delta g_n(\mathbf{v})=\mathbf{v}^S\cdot\mathbf{n}^M-\mathbf{v}^M\cdot\mathbf{n}^M=[\mathbf{v}]\cdot\mathbf{n}^M \tag{100}$$

Exactly as in the case of the additional symmetry term, adding (98) to (95) does not change the solution of the equation, because if $\mathbf{u}$ is the solution of the equation, then $g_n(\mathbf{u})=0$ on $\overline{\Gamma}^C$. Let us comment that the bilinear part of (98) is also symmetric to preserve the symmetry of the bilinear form $a(\mathbf{u},\mathbf{v})$. Finally, we can formulate our problem as

$$\begin{aligned}&\int_\Omega\boldsymbol{\sigma}(\mathbf{u}):\boldsymbol{\nabla}\mathbf{v}\,\mathrm{d}V-\int_\Omega\rho\overline{\mathbf{b}}\cdot\mathbf{v}\,\mathrm{d}V-\int_{\Gamma^\sigma}\overline{\mathbf{t}}\cdot\mathbf{v}\,\mathrm{d}A\\&+\int_{\overline{\Gamma}^{C,S}}\langle\sigma_n(\mathbf{u})\rangle[\mathbf{v}]\cdot\mathbf{n}^M\,\mathrm{d}A+\int_{\overline{\Gamma}^{C,S}}\langle\sigma_n(\mathbf{v})\rangle\left([\mathbf{u}]\cdot\mathbf{n}^M+[\mathbf{X}]\cdot\mathbf{n}^M\right)\mathrm{d}A\\&+\frac{\gamma E}{h}\int_{\overline{\Gamma}^{C,S}}\left([\mathbf{u}]\cdot\mathbf{n}^M+[\mathbf{X}]\cdot\mathbf{n}^M\right)[\mathbf{v}]\cdot\mathbf{n}^M\,\mathrm{d}A=0\quad\forall\mathbf{v}\in\mathbf{V}\end{aligned} \tag{101}$$

or equivalently

$$\begin{aligned}&\int_\Omega\boldsymbol{\sigma}(\mathbf{u}):\boldsymbol{\nabla}\mathbf{v}\,\mathrm{d}V-\int_\Omega\rho\overline{\mathbf{b}}\cdot\mathbf{v}\,\mathrm{d}V-\int_{\Gamma^\sigma}\overline{\mathbf{t}}\cdot\mathbf{v}\,\mathrm{d}A\\&+\int_{\overline{\Gamma}^{C,S}}\langle\sigma_n(\mathbf{u})\rangle\delta g_n(\mathbf{v})\,\mathrm{d}A+\int_{\overline{\Gamma}^{C,S}}\langle\sigma_n(\mathbf{v})\rangle g_n(\mathbf{u})\,\mathrm{d}A\\&+\frac{\gamma E}{h}\int_{\overline{\Gamma}^{C,S}}g_n(\mathbf{u})\delta g_n(\mathbf{v})\,\mathrm{d}A=0\quad\forall\mathbf{v}\in\mathbf{V}\end{aligned} \tag{102}$$

In the obtained variational equation (102), all contact terms are integrated over the active contact zone $\overline{\Gamma}^{C,S}$, where $g_n(\mathbf{u})=0$. However, we do not know $\overline{\Gamma}^{C,S}$ in forward. Therefore, we need to reformulate (102) in such a way, that the integral over $\overline{\Gamma}^{C,S}$ could be extended to the integral over $\Gamma^{C,S}$. We remind here that on $\Gamma^{C,S}$, we require the contact conditions (15) to hold, including $g_n(\mathbf{u})\geq 0$. It could be easily proved that the condition $g_n(\mathbf{u})\geq 0$ is equivalent to

$$g_n(\mathbf{u})=-\{-g_n(\mathbf{u})\} \tag{103}$$

where $\{\bullet\}$ is the Macaulay bracket defining the so called 'ramp' function as

$$\{x\}:=\frac{x+|x|}{2}=\begin{cases}x, & x\geq 0\\ 0, & x<0\end{cases} \tag{104}$$

The properties of the 'ramp' function listed below will be very useful in the following derivations

$$x\leq\{x\},\quad x\{x\}=\{x\}^2\quad\forall x\in\mathbb{R} \tag{105}$$

So in order to define our problem on $\Gamma^{C,S}$, we will substitute

$$\begin{aligned} g_n &\sim -\{-g_n(\mathbf{u})\} \\ \delta g_n &\sim H(-g_n(\mathbf{u}))\delta g_n(\mathbf{v}) \end{aligned} \tag{106}$$

where $H(\bullet)$ is the Heaviside function defined as

$$H(x) := \frac{1}{2}(1+\operatorname{sgn}(x)) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{107}$$

and $\operatorname{sgn}(\bullet)$ denotes the sign function defined as

$$\operatorname{sgn} x := \begin{cases} 1, & x \geq 0 \\ -1, & x < 0 \end{cases} \tag{108}$$

The relation between the 'ramp' function and the Heaviside function is

$$\{x\} = x \cdot H(x) \tag{109}$$

Also the property (110) below is very useful

$$\{x\} \cdot H(x) = \{x\} \tag{110}$$

because it allows us to simplify the following derivative as

$$\frac{d}{dx}\left(\frac{1}{2}\{x\}^2\right) = \{x\} \cdot H(x) = \{x\} \tag{111}$$

The substitution as stated in (106) will secure that only the negative gap will be penalized. Substituting (106) to (102) we obtain the formulation of our problem, which now reads

$$\begin{aligned} &\int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\nabla}\mathbf{v} \,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}} \cdot \mathbf{v} \,\mathrm{d}V - \int_{\Gamma^\sigma} \overline{\mathbf{t}} \cdot \mathbf{v} \,\mathrm{d}A \\ &+ \int_{\Gamma^{C,S}} \langle\sigma_n(\mathbf{u})\rangle H(-g_n(\mathbf{u}))\delta g_n(\mathbf{v}) \,\mathrm{d}A - \int_{\Gamma^{C,S}} \langle\sigma_n(\mathbf{v})\rangle\{-g_n(\mathbf{u})\} \,\mathrm{d}A \\ &- \frac{\gamma E}{h}\int_{\Gamma^{C,S}} \{-g_n(\mathbf{u})\}H(-g_n(\mathbf{u}))\delta g_n(\mathbf{v}) \,\mathrm{d}A = 0 \quad \forall \mathbf{v} \in \mathbf{V} \end{aligned} \tag{112}$$

And finally, we will utilize the fact, that $\boldsymbol{\sigma}(\mathbf{u})$ is a symmetric second-order tensor and so $\boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\nabla}\mathbf{v} = \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\nabla}_S\mathbf{v} = \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{v})$. Here $\boldsymbol{\varepsilon}(\mathbf{v})$ is a small deformation strain tensor. Also, because of (110), $H(-g_n(\mathbf{u}))$ in the last (penalty) member could be omitted. Considering these, we obtain

$$\begin{aligned} &\int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{v}) \,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}} \cdot \mathbf{v} \,\mathrm{d}V - \int_{\Gamma^\sigma} \overline{\mathbf{t}} \cdot \mathbf{v} \,\mathrm{d}A \\ &+ \int_{\Gamma^{C,S}} \langle\sigma_n(\mathbf{u})\rangle H(-g_n(\mathbf{u}))\delta g_n(\mathbf{v}) \,\mathrm{d}A - \int_{\Gamma^{C,S}} \langle\sigma_n(\mathbf{v})\rangle\{-g_n(\mathbf{u})\} \,\mathrm{d}A \\ &- \frac{\gamma E}{h}\int_{\Gamma^{C,S}} \{-g_n(\mathbf{u})\}\delta g_n(\mathbf{v}) \,\mathrm{d}A = 0 \quad \forall \mathbf{v} \in \mathbf{V} \end{aligned} \tag{113}$$

The previously derived weak formulation is equivalent to the first variation of the following functional of elastic energy

$$\begin{aligned} \Pi_W(\mathbf{u}) = &\frac{1}{2}\int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{u}) \,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}} \cdot \mathbf{u} \,\mathrm{d}V - \int_{\Gamma^\sigma} \overline{\mathbf{t}} \cdot \mathbf{u} \,\mathrm{d}A \\ &- \int_{\Gamma^{C,S}} \langle\sigma_n(\mathbf{u})\rangle\{-g_n(\mathbf{u})\} \,\mathrm{d}A + \frac{1}{2}\frac{\gamma E}{h}\int_{\Gamma^{C,S}} \{-g_n(\mathbf{u})\}^2 \,\mathrm{d}A \end{aligned} \tag{114}$$

We can verify, that the stationarity condition of the functional $\Pi_W(\mathbf{u})$ with respect to $\mathbf{u}$ yields (113). The variation of $\Pi_W(\mathbf{u})$ with respect to $\mathbf{u}$ is defined in a classic way as

$$\delta\Pi_W(\mathbf{u}, \mathbf{v}) = \left.\frac{\mathrm{d}}{\mathrm{d}\alpha}\Pi_W(\mathbf{u} + \alpha\mathbf{v})\right|_{\alpha=0} = 0 \quad \forall \mathbf{v} \in \mathbf{V} \tag{115}$$

We remind here relation (111) which states that

$$\frac{d}{dx}\left(\frac{1}{2}\{x\}^2\right) = \{x\} \cdot H(x) = \{x\}$$

Utilizing this we obtain

$$\begin{aligned}
&\delta\Pi_W(\mathbf{u},\mathbf{v}) = \int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{v})\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\mathbf{v}\,\mathrm{d}V - \int_{\Gamma^\sigma} \overline{\mathbf{t}}\cdot\mathbf{v}\,\mathrm{d}A \\
&+ \int_{\Gamma^{C,S}} \langle\sigma_n(\mathbf{u})\rangle H(-g_n(\mathbf{u}))\delta g_n(\mathbf{v})\,\mathrm{d}A - \int_{\Gamma^{C,S}} \langle\sigma_n(\mathbf{v})\rangle\{-g_n(\mathbf{u})\}\,\mathrm{d}A \\
&- \frac{\gamma E}{h}\int_{\Gamma^{C,S}} \{-g_n(\mathbf{u})\}\delta g_n(\mathbf{v})\,\mathrm{d}A = 0 \qquad \forall\mathbf{v}\in\mathbf{V}
\end{aligned} \tag{116}$$

For the son-symmetric version of the method (in sense the relevant bilinear form is non-symmetric) no such energy functional exists.

3.2.2 Derivation of the approach by Fabré, Renard and Pousin to the Nitsche method (Nitsche-FPR)

We will now present another derivation of the Nitsche method based on the paper [15]. This approach is in return based on papers by Chouly, Hild and Renard [29] and [30] published in 2013 and 2015 respectively. For the purpose of brevity, we will refer to this approach as to the Nitsche-FPR approach. The 'FPR' abbreviation goes for the initial letters of the surnames of the authors of the paper - Fabré, Renard and Pousin. We would like to point out in advance that the resulting weak form will not coincide with the previous version by Wriggers. The connection of both approaches will be investigated in an independent section.

Similarly to the previous chapter on the derivation of Wriggers' version of the Nitsche method, we will start from equation (25), which reads

$$\begin{aligned} &\int_{\Omega} \boldsymbol{\sigma}\left(\mathbf{u}\right) : \boldsymbol{\nabla}\mathbf{v}\,\mathrm{d}V - \int_{\Omega} \rho\overline{\mathbf{b}}\cdot\mathbf{v}\,\mathrm{d}V - \int_{\Gamma^{\sigma}} \overline{\mathbf{t}}\cdot\mathbf{v}\,\mathrm{d}A \\ &+ \int_{\Gamma^{C,S}} \sigma_n^S\left[\mathbf{v}\right]\cdot\mathbf{n}^M\,\mathrm{d}A = 0 \quad \forall\mathbf{v}\in\mathbf{V} \end{aligned} \tag{117}$$

Now, in the paper [15] is derived that for an arbitrary $\gamma \in \mathbb{R}^+$ the contact conditions (17) on Γ^C can be equivalently rewritten as

$$\sigma_n^S = -\frac{1}{\gamma}\{-g_n - \gamma\sigma_n^S\} \tag{118}$$

We remind that g_n denotes the gap function, which reads

$$g_n(\mathbf{u}) = \mathbf{X}^S\cdot\mathbf{n}^M + \mathbf{u}^S\cdot\mathbf{n}^M - \mathbf{X}^M\cdot\mathbf{n}^M - \mathbf{u}^M\cdot\mathbf{n}^M = [\mathbf{u}]\cdot\mathbf{n}^M + [\mathbf{X}]\cdot\mathbf{n}^M \tag{119}$$

and

$$\sigma_n(\mathbf{u}) = \mathbf{n}^M\cdot\boldsymbol{\sigma}\left(\mathbf{u}\right)\cdot\mathbf{n}^M = \mathbf{n}^S\cdot\boldsymbol{\sigma}\left(\mathbf{u}\right)\cdot\mathbf{n}^S \tag{120}$$

We will now verify, that the contact conditions (17), which read

$$\begin{aligned} g_n &\geq 0 \quad \text{(i)} \\ \sigma_n &\leq 0 \quad \text{(ii)} \\ \sigma_n g_n &= 0 \quad \text{(iii)} \end{aligned} \tag{121}$$

are equivalent to the equation (118), given as

$$\sigma_n^S = -\frac{1}{\gamma}\{-g_n - \gamma\sigma_n^S\} \tag{122}$$

We would like to remark that 118 is an example of the so-called non-linear complementarity (NCP) function. As written in Remark 5 in [31], NCP allows to reformulate complementarity conditions ((17) in our case) using a single non-linear relationship (see [19] for more details). The formulation we have chosen is not the only one possible though it is possibly one of the simplest ones.

We will now prove the equivalence of the contact conditions (17) and equation (118). This proof was adopted from [29, p. 1297 - 1298]. We will write just σ_n instead of σ_n^S to work with less intriguing notation.

Firstly, we will prove the implication (17) $\Rightarrow$ (118). So, let us have such a regular vector field u on Ω that conditions (17)(i)–(iii) hold. We immediately see, that from the condition (17)(ii), we get that either $\sigma_n < 0$ or

$\sigma_n = 0$ holds. Let us consider the case $\sigma_n < 0$ first. In this case (17)(iii) implies that $g_n = 0$. By substituting this into (118), we get

$$-\frac{1}{\gamma}\{-g_n - \gamma\sigma_n\} = -\frac{1}{\gamma}\{-\gamma\sigma_n\} = \sigma_n \tag{123}$$

So we see, that equation (118) is fulfilled.

Now, let us focus on the case $\sigma_n = 0$. The condition (17)(i) can also be expressed as $\{-g_n\} = 0$.

$$-\frac{1}{\gamma}\{-g_n - \gamma\sigma_n\} = -\frac{1}{\gamma}\{-g_n\} = 0 = \sigma_n \tag{124}$$

So, we see, that in this case (118) holds again and therefore (17) $\Rightarrow$ (118).

We will focus now on the converse implication. So we have u such that (118) holds. This implies $\sigma_n \leq 0$. This immediately implies that (17)(ii) holds. Again, let us consider the cases $\sigma_n < 0$ and $\sigma_n = 0$ separately. Firstly, if $\sigma_n = 0$, then (118) can be rewritten as

$$0 = -\frac{1}{\gamma}\{-g_n\} \tag{125}$$

which is equivalent to condition (17)(i). Also, because we have considered $\sigma_n = 0$, we get $\sigma_n g_n = 0$ and so (17)(iii) holds too.

Finally, if $\sigma_n < 0$, then from From (118) we get $\{-g_n - \gamma\sigma_n\} > 0$ and so

$$-\gamma\sigma_n = \{-g_n - \gamma\sigma_n\} = -g_n - \gamma\sigma_n \tag{126}$$

Because this means $-g_n = 0$, (17)(i) and (17)(iii) are satisfied. This implies (17) $\Leftarrow$ (118)

[29, p. 1297 - 1298] also mentions that condition (118) is still equivalent to (17)(i)- (iii) on Γ_C when γ is a positive function defined on Γ_C instead of a positive constant.

Now, exactly as in [15, p. 23], let θ be a fixed parameter and $\theta \in [-1, 1]$. As will be shown later, θ determines the symmetry properties. We can formally write $[\mathbf{v}] \cdot \mathbf{n}^M = ([\mathbf{v}] \cdot \mathbf{n}^M - \theta\gamma\sigma_n(\mathbf{v}) + \theta\gamma\sigma_n(\mathbf{v}))$ and substitute this relation to (117), which yields

$$\begin{aligned}&\int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\nabla}\mathbf{v}\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}} \cdot \mathbf{v}\,\mathrm{d}V - \int_{\Gamma^\sigma} \overline{\mathbf{t}} \cdot \mathbf{v}\,\mathrm{d}A \\ &+ \int_{\Gamma^{C,S}} \theta\gamma\sigma_n^S(\mathbf{u})\sigma_n^S(\mathbf{v})\,\mathrm{d}A + \int_{\Gamma^{C,S}} \sigma_n^S(\mathbf{u})([\mathbf{v}] \cdot \mathbf{n}^M - \theta\gamma\sigma_n^S(\mathbf{v}))\,\mathrm{d}A = 0 \quad \forall \mathbf{v} \in \mathbf{V}\end{aligned} \tag{127}$$

We will utilize relation (118), which was derived in the beginning of this section, and substitute $\sigma_n^S(\mathbf{u}) = -\frac{1}{\gamma}\{-g_n(\mathbf{u}) - \gamma\sigma_n^S(\mathbf{u})\}$ in the last integral. From this we obtain

$$\begin{aligned}&\int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\nabla}\mathbf{v}\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}} \cdot \mathbf{v}\,\mathrm{d}V - \int_{\Gamma^\sigma} \overline{\mathbf{t}} \cdot \mathbf{v}\,\mathrm{d}A \\ &+ \int_{\Gamma^{C,S}} \theta\gamma\sigma_n^S(\mathbf{u})\sigma_n^S(\mathbf{v})\,\mathrm{d}A \\ &- \int_{\Gamma^{C,S}} \frac{1}{\gamma}\{-g_n(\mathbf{u}) - \gamma\sigma_n^S(\mathbf{u})\}([\mathbf{v}] \cdot \mathbf{n}^M - \theta\gamma\sigma_n^S(\mathbf{v}))\,\mathrm{d}A = 0 \quad \forall \mathbf{v} \in \mathbf{V}\end{aligned} \tag{128}$$

To ensure the stability of the method, [15, p. 23] introduces a 'stabilised formulation for elements having a small contribution'. This is based on replacing $\sigma_n^S(\mathbf{u})$ by the convex combination of $\sigma_n^S(\mathbf{u})$ and $\sigma_n^M(\mathbf{u})$. As we remember the relation (14), which reads

$$\sigma_n^S(\mathbf{u}) = \sigma_n^M(\mathbf{u}). \tag{129}$$

we can write

$$\sigma_n^S(\mathbf{u}) = t\sigma_n^M(\mathbf{u}) - t\sigma_n^S(\mathbf{u}) + \sigma_n^S(\mathbf{u}). \tag{130}$$

For this reason, in [15, p. 23], the authors replace $\sigma_n^S(\mathbf{u})$ by $\langle\sigma_n(\mathbf{u})\rangle_t$ defined as

$$\langle\sigma_n(\mathbf{u})\rangle_t = t\sigma_n^M(\mathbf{u}) + (1-t)\sigma_n^S(\mathbf{u}) \tag{131}$$

where $t \in [0,1]$ which may (in the context of the finite element method) differ from element to element of the discretised problem. Authors of [15] note that a similar approach has been developed in [1], where an optimal choice of the fixed parameter $t \in [0,1]$ is proposed. Let us also note that the notation $\langle\bullet\rangle_t$ is intentionally chosen to remind the definition of the average of contact stress (92) from the derivation of the Wriggers' approach to the Nitsche method in the previous section.

By substituting (131) to (128) and considering $[\mathbf{v}]\cdot\mathbf{n}^M = \delta g_n(\mathbf{v})$, the Nitsche method is formulated in terms of [15] and the resulting equation reads

$$\begin{aligned}
&\int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\nabla}\mathbf{v}\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\mathbf{v}\,\mathrm{d}V - \int_{\Gamma^\sigma} \overline{\mathbf{t}}\cdot\mathbf{v}\,\mathrm{d}A \\
&+ \int_{\Gamma^{C,S}} \theta\gamma\langle\sigma_n(\mathbf{u})\rangle_t\langle\sigma_n(\mathbf{v})\rangle_t\,\mathrm{d}A \\
&- \int_{\Gamma^{C,S}} \frac{1}{\gamma}\{-g_n(\mathbf{u}) - \gamma\langle\sigma_n(\mathbf{u})\rangle_t\}(\delta g_n(\mathbf{v}) - \theta\gamma\langle\sigma_n(\mathbf{v})\rangle_t)\,\mathrm{d}A = 0 \quad \forall\mathbf{v}\in\mathbf{V}
\end{aligned} \tag{132}$$

The paper [15] mentions that although θ could be chosen in $[-1,1]$, three choices are of special interest. Firstly $\theta = -1$, which (as the authors mention) leads to the symmetric method, which was proposed and analysed in [29]. Secondly, $\theta = 0$ establishes a non-symmetric version proposed in [30] and finally for $\theta = 1$, another non-symmetric version is obtained.

One could remember that in the chapter on the derivation of Nitsche-Wriggers, the parameter γ was also used. In this paragraph, let us denote this parameter as γ and the penalty parameter from this chapter as γ_{FPR}. It turns out that if

$$\frac{1}{\gamma_{FPR}} = \frac{\gamma E}{h} \tag{133}$$

then the weak forms of Nitsche-Wriggers and Nitsche-FPR coincide on the part of the potential contact zone where contact is activated according to both methods (in the sense that the expressions in Macaulay brackets are positive in both weak forms). We will not verify this here because the comparison of both methods will be discussed in more detail in the following section. Same as in the previous chapter on Wriggers' version of the Nitsche method, E/h is the scaling parameter of the penalty term, so that γ is dimensionless. E is Young's modulus and h is the characteristic size of the mesh elements (e.g. h the average of the largest and smallest diameter of the sphere circumscribed to the elements. h is considered to be constant over the mesh). To sum up, the fact that the weak forms of both methods coincide on $\overline{\Gamma}^C$ if (133) holds, motivates us to substitute γ_{FPR} for $h/(\gamma E)$ in (132). This yields

$$\begin{aligned}
&\int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\nabla}\mathbf{v}\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\mathbf{v}\,\mathrm{d}V - \int_{\Gamma^\sigma} \overline{\mathbf{t}}\cdot\mathbf{v}\,\mathrm{d}A \\
&+ \int_{\Gamma^{C,S}} \theta\frac{h}{\gamma E}\langle\sigma_n(\mathbf{u})\rangle_t\langle\sigma_n(\mathbf{v})\rangle_t\,\mathrm{d}A \\
&- \int_{\Gamma^{C,S}} \frac{\gamma E}{h}\left\{-g_n(\mathbf{u}) - \frac{h}{\gamma E}\langle\sigma_n(\mathbf{u})\rangle_t\right\}\left(\delta g_n(\mathbf{v}) - \theta\frac{h}{\gamma E}\langle\sigma_n(\mathbf{v})\rangle_t\right)\mathrm{d}A = 0 \quad \forall\mathbf{v}\in\mathbf{V}
\end{aligned} \tag{134}$$

Finally, as in the Nitsche-Wriggers case, for $\theta = -1$ when the relevant bilinear form is symmetric, such energy functional could be found that its variation is the equation (134).

$$\begin{aligned}\Pi_{FPR,SYM}(\mathbf{u}) =& \frac{1}{2}\int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{u})\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\mathbf{u}\,\mathrm{d}V - \int_{\Gamma^\sigma}\overline{\mathbf{t}}\cdot\mathbf{u}\,\mathrm{d}A \\ &- \frac{h}{2\gamma E}\int_{\Gamma^{C,S}}\langle\sigma_n(\mathbf{u})\rangle_t^2\,\mathrm{d}A \\ &+ \frac{\gamma E}{2h}\int_{\Gamma^{C,S}}\left\{-g_n(\mathbf{u}) - \frac{h}{\gamma E}\sigma_n(\mathbf{u})\right\}^2\,\mathrm{d}A\end{aligned} \tag{135}$$

It could be verified that the first variation of $\Pi(\mathbf{u})$ with respect to $\mathbf{u}$ gives the left-hand side of (134). Let us remind useful relation (111) which reads

$$\frac{d}{dx}\left(\frac{1}{2}\{x\}^2\right) = \{x\}\cdot H(x) = \{x\}$$

Defining the stationarity condition of $\Pi_{FPR,SYM}(\mathbf{u})$ as

$$\delta\Pi_{FPR,SYM}(\mathbf{u},\mathbf{v}) = \left.\frac{\mathrm{d}}{\mathrm{d}\alpha}\Pi_{FPR,SYM}(\mathbf{u}+\alpha\mathbf{v})\right|_{\alpha=0} = 0 \qquad \forall\mathbf{v}\in\mathbf{V} \tag{136}$$

we obtain

$$\begin{aligned}&\delta\Pi_{FPR,SYM}(\mathbf{u},\mathbf{v}) = \int_\Omega \boldsymbol{\sigma}\,(\mathbf{u}) : \boldsymbol{\nabla}\mathbf{v}\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\mathbf{v}\,\mathrm{d}V - \int_{\Gamma^\sigma}\overline{\mathbf{t}}\cdot\mathbf{v}\,\mathrm{d}A \\ &+ \int_{\Gamma^{C,S}}\theta\frac{h}{\gamma E}\langle\sigma_n(\mathbf{u})\rangle_t\langle\sigma_n(\mathbf{v})\rangle_t\,\mathrm{d}A \\ &- \int_{\Gamma^{C,S}}\frac{\gamma E}{h}\left\{-g_n(\mathbf{u}) - \frac{h}{\gamma E}\langle\sigma_n(\mathbf{u})\rangle_t\right\}\left(\delta g_n(\mathbf{v}) - \theta\frac{h}{\gamma E}\langle\sigma_n(\mathbf{v})\rangle_t\right)\mathrm{d}A = 0 \qquad \forall\mathbf{v}\in\mathbf{V}\end{aligned} \tag{137}$$

3.2.3 Some remarks on the mathematical properties of Nitsche-FPR

In this section, we would like to mention some interesting properties of the Nitsche-FPR method. Due to an extensive effort of many authors (let us mention Chouly, Hild, Renard, Fabré, Pousin and Mlika [29], [30], [31]), the Nitsche-FPR method has been thoroughly mathematically analysed. The following text is quite brief and is mostly based on an excellent overview [31], which we recommend to the attention of an interested reader. For a briefer notation, let us define the function space $\mathbf{U}^s$ as $\mathbf{U}^s = \{\mathbf{u}\in H^s(\Omega)^d\,;\mathbf{u} = \overline{\mathbf{u}} \text{ on } \Gamma^D\}$[7], where d is the dimension of the problem, which is either 2 or 3.

At the beginning of this chapter, we suggested that compared to penalty approaches, the Nitsche method is consistent and on the theoretical continuum level, contact conditions are fulfilled exactly. We did not prove this for Nitsche-Wriggers, though in the case of Nitsche-FPR there exists a proof of this - see theorem 3.4. in [15] or Lemma 2 [31]. These consider the Nitsche-FPR method to be consistent in the following sense. Let us suppose that $\mathbf{u}$ is the solution to the strong problem of the contact of two elastic bodies formulated by equations (16) and contact conditions (17). Let us also suppose that $\mathbf{u}$ is sufficiently regular. According to [15], this could be typically $\mathbf{u}\in\mathbf{U}^{2+\nu}$, for $\nu > 0$. If these conditions are fulfilled, then $\mathbf{u}$ is also a solution to the Nitsche-FPR problem (135). We recommend the interested reader to see theorem 3.4. for more details on the proof.

Also, as the authors of [15] state in theorem 3.5, if $\mathbf{u}\in\mathbf{U}^2$ is a solution to the weak problem (135) of the Nitsche-FPR method then $\mathbf{u}$ is a solution to strong problem (16), (17).

[7] Here, $H^s(\Omega)^d$ is the Sobolev space, i.e. the space of functions the derivatives of which up to the order s belong to the Lebesgue space $L^2(\Omega)^d$.

Moreover, according to Theorem 1 [31], under some requirements on the penalty parameter γ (which include the value of parameter θ) the discretely formulated problem 135 of the Nitsche-FPR method is well-posed and so admits one unique solution.

Now we will cite a few results on the convergence of Nitsche-FPR from [31]. Firstly, if the solution $\mathbf{u}$ has a regularity $\mathbf{U}^{\frac{3}{2}+\nu}, 0 < \nu \leq k - 1/2$, where $k = 1, 2$ is the polynomial degree of the Lagrange finite elements, then an optimal convergence in H^1-norm of order $O\left(h^{\frac{1}{2}+\nu}\right)$ was proved in [31]. The proof applies to both two-dimensional and three-dimensional problems. From the mathematical point of view, it is quite interesting that the authors achieved this without any additional assumption on the contact/friction zone, such as an increased regularity of the contact stress. The authors of [31] mention that it has been quite challenging to establish the same proof for other mixed or penalty methods without any additional assumptions and it was firstly done only recently in 2015.

As mentioned, the Nitsche-FPR problem (135) contains the parameter θ, which decides the symmetric or non-symmetric nature of the method. According to [31], if $\theta = -1$, the symmetric version of the method is obtained, and there exists a corresponding energy functional described in the previous text. When $\theta \neq -1$, the symmetry is lost. This also means that the resulting linear system solved during Newton iterations of the finite element problem is not symmetric and some efficient solvers that require symmetric stiffness matrix cannot be applied. However, non-symmetric variants of the method can also have certain advantages. If $\theta = 0$, the number of terms is reduced, which makes the method easier to implement [31]. In the case of some complicated problems, this could be very convenient. Also, it was observed in [31] that this version of Nitsche-FPR requires less Newton iterations to converge. Moreover, authors of [31] report that even the range of penalty parameter for which the method converged was wider than for the symmetric option. For $\theta = 1$ one obtains another non-symmetric variant. In this case, interestingly, the discrete version of the problem (135) remains well-posed and converges optimally for every positive penalty value [31].

The paper [31] also provides some details on a priori and a posteriori estimates of the error of the Nitsche-FPR method in relevant norms.

3.2.4 Connection between Nitsche-Wriggers and Nitsche-FPR methods

So far, we have described two versions of the Nitsche method for enforcing the contact boundary conditions of two elastic bodies in the case of small deformation, which can be found in the current literature — Nitsche-Wriggers (described in book [55, Chapter 5, p.106] and paper [56]) and Nitsche-FPR (described in [15]). It should also be noted that an approach analogical to the Nitsche-FPR method is currently dominant in the mechanical and applied mathematical journal papers. We will discuss and demonstrate the behaviour of both approaches and their strengths and weaknesses in the following chapters on a group of testing numerical examples.

However, it is quite confusing that even though both methods are presented as the Nitsche methods, the resulting weak forms are not the same. Let us compare the weak form of the Nitsche-Wriggers method (113)

$$\begin{aligned}
&\int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{v}) \,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}} \cdot \mathbf{v} \,\mathrm{d}V - \int_{\Gamma^\sigma} \overline{\mathbf{t}} \cdot \mathbf{v} \,\mathrm{d}A \\
&+ \int_{\Gamma^{C,S}} \langle \sigma_n(\mathbf{u}) \rangle H(-g_n(\mathbf{u})) \delta g_n(\mathbf{v}) \,\mathrm{d}A - \int_{\Gamma^{C,S}} \langle \sigma_n(\mathbf{v}) \rangle \{-g_n(\mathbf{u})\} \,\mathrm{d}A \\
&- \frac{\gamma E}{h} \int_{\Gamma^{C,S}} \{-g_n(\mathbf{u})\} H(-g_n(\mathbf{u})) \delta g_n(\mathbf{v}) \,\mathrm{d}A = 0 \quad \forall \mathbf{v} \in \mathbf{V}
\end{aligned} \tag{138}$$

and the weak form of the Nitsche-FPR method (134)

$$\begin{aligned}
&\int_\Omega \boldsymbol{\sigma}\left(\mathbf{u}\right) : \boldsymbol{\nabla}\mathbf{v}\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\mathbf{v}\,\mathrm{d}V - \int_{\Gamma^\sigma}\overline{\mathbf{t}}\cdot\mathbf{v}\,\mathrm{d}A \\
&+ \int_{\Gamma^{C,S}} \theta\frac{h}{\gamma E}\langle\sigma_n(\mathbf{u})\rangle_t\langle\sigma_n(\mathbf{v})\rangle_t\,\mathrm{d}A \\
&- \int_{\Gamma^{C,S}} \frac{\gamma E}{h}\left\{-g_n(\mathbf{u}) - \frac{h}{\gamma E}\langle\sigma_n(\mathbf{u})\rangle_t\right\}\left(\delta g_n(\mathbf{v}) - \theta\frac{h}{\gamma E}\langle\sigma_n(\mathbf{v})\rangle_t\right)\mathrm{d}A = 0 \qquad \forall\mathbf{v}\in\mathbf{V}
\end{aligned} \tag{139}$$

Observing both equations we see that the integral over the boundary contains different expressions. Immediately, the question raises whether both methods are, in fact, equivalent or whether there is some difference between them.

In order to enlighten this issue, let us focus on the symmetric version of both methods. Equation (138) refers to the symmetric version of the Nitsche-Wriggers method and choosing $\theta = -1$ in (139) we ensure the same for Nitsche-FPR. In this case, there exist energy functionals for both methods, (114) for Nitsche-Wriggers and (135) for Nitsche-FPR, which read

$$\begin{aligned}
\Pi_W(\mathbf{u}) =& \frac{1}{2}\int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{u})\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\mathbf{u}\,\mathrm{d}V - \int_{\Gamma^\sigma}\overline{\mathbf{t}}\cdot\mathbf{u}\,\mathrm{d}A \\
&- \int_{\Gamma^{C,S}}\langle\sigma_n(\mathbf{u})\rangle\{-g_n(\mathbf{u})\}\,\mathrm{d}A + \frac{1}{2}\frac{\gamma E}{h}\int_{\Gamma^{C,S}}\{-g_n(\mathbf{u})\}^2\,\mathrm{d}A
\end{aligned} \tag{140}$$

$$\begin{aligned}
\Pi_{FPR}(\mathbf{u}) =& \frac{1}{2}\int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{u})\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\mathbf{u}\,\mathrm{d}V - \int_{\Gamma^\sigma}\overline{\mathbf{t}}\cdot\mathbf{u}\,\mathrm{d}A \\
&- \frac{h}{2\gamma E}\int_{\Gamma^{C,S}}\langle\sigma_n(\mathbf{u})\rangle_t^2\,\mathrm{d}A \\
&+ \frac{\gamma E}{2h}\int_{\Gamma^{C,S}}\left\{-g_n(\mathbf{u}) - \frac{h}{\gamma E}\sigma_n(\mathbf{u})\right\}^2\mathrm{d}A
\end{aligned} \tag{141}$$

Firstly, to simplify the notation, let us write just σ_n instead of $\langle\sigma_n(\mathbf{u})\rangle$ or $\langle\sigma_n(\mathbf{u})\rangle_t$, $\delta\sigma_n$ instead of $\langle\sigma_n(\mathbf{v})\rangle$ or $\langle\sigma_n(\mathbf{v})\rangle_t$. Although this notation is introduced here for the comfort of the reader, it is actually correct in the case of unilateral contact. This means the contact of one elastic body (which is considered to be 'slave' body) and rigid obstacle (e.g. rigid plane). In the case of unilateral contact, the normal contact tractions on the active contact zone of the rigid obstacle can be defined to be equal to the ones on the active contact zone of the elastic body. This definition of tractions on the rigid body is correct, because the displacement of the rigid body is zero in every material point, but the stress and tractions can be arbitrary. Equipped with these definitions of tractions, we can formally apply the average and jump operators and we can write σ_n instead of $\langle\sigma_n(\mathbf{u})\rangle$ or $\langle\sigma_n(\mathbf{u})\rangle_t$, $\delta\sigma_n$ instead of $\langle\sigma_n(\mathbf{v})\rangle$ or $\langle\sigma_n(\mathbf{v})\rangle_t$.

As we see that the expressions which are integrated over volume are the same for both methods, we will restrict our attention only to the integral over the potential contact zone. This integral will be further denoted as $\Pi_{W,C}$ for Nitsche-Wriggers and $\Pi_{FPR,C}$ for Nitsche-FPR. With the suggested notation, we can rewrite energy functionals for both Nitsche methods (140) and (141) as

$$\Pi_{FPR,C} = \int_{\Gamma^{C,S}}\left(-\frac{h}{2\gamma E}\sigma_n^2 + \frac{\gamma E}{2h}\left\{-g_n - \frac{h}{\gamma E}\sigma_n\right\}^2\right)\mathrm{d}A \tag{142}$$

$$\Pi_{W,C} = \int_{\Gamma^{C,S}}\left(-\sigma_n\{-g_n\} + \frac{\gamma E}{2h}\{-g_n\}^2\right)\mathrm{d}A \tag{143}$$

Now let us have the given displacement field $\mathbf{u}$. Then there exists a sub-domain $\tilde{\Gamma}^C, \quad \tilde{\Gamma}^C \subset \Gamma^C$, where contact is activated according to both methods (in the sense that expressions in Macaulay brackets are positive in both

(142) and (143)). Then we get

$$\Pi_{FPR,C}^{\tilde{\Gamma}^{C,S}} = \int_{\tilde{\Gamma}^{C,S}} \left(-\frac{h}{2\gamma E}\sigma_n^2 + \frac{\gamma E}{2h}\left(-g_n - \gamma\sigma_n\right)^2 \right) \mathrm{d}A \tag{144}$$

$$= \int_{\tilde{\Gamma}^{C,S}} \left(\frac{\gamma E}{2h} g_n^2 + g_n\sigma_n \right) \mathrm{d}A$$

$$\Pi_{W,C}^{\tilde{\Gamma}^{C,S}} = \int_{\tilde{\Gamma}^{C,S}} \left(\sigma_n g_n + \frac{\gamma E}{2h} g_n^2 \right) \mathrm{d}A \tag{145}$$

Now, we finally see that added terms integrated over the contact boundary used by the Nitsche-Wriggers method and by the symmetric version of the Nitsche-FPR method coincide on $\tilde{\Gamma}^{C,S}$. However, the two formulations differ in the criterion that decides whether contact occurs. Wriggers' version uses the sign of the gap as the contact indicator, while the FPR version uses the sign of $g_n + \frac{h}{\gamma E}\sigma_n$, i.e., of the gap plus some multiple of the traction normal to the contact surface. Both criteria are not equivalent, and it turns out that the different choice of a contact indicator also influences the differences in the convergence of both methods in terms of the finite element numerical solution. Specifically, during the numerical testing, it turned out that contrary to the Nitsche-FPR method, Wriggers' contact indicator could be quite sensitive to a small numerical error. This results from the fact that in Nitsche-Wriggers, both the expressions $-\{-g_n(\mathbf{u})\}$ and its variation $H(-g_n(\mathbf{u}))\delta g_n(\mathbf{v})$ appear. The latter is the discontinuous function for states for which $g_n = 0$. If $\mathrm{sgn}(g_n(\mathbf{u})) = 1$ holds in a certain point of the potential contact zone then there is no contact in this point and $H(-g_n(\mathbf{u}))\delta g_n(\mathbf{v}) = 0$. On the contrary, if $\mathrm{sgn}(g_n(\mathbf{u})) = -1$, there is a contact in the considered point of the potential contact zone and $H(-g_n(\mathbf{u}))\delta g_n(\mathbf{v}) = \delta g_n(\mathbf{v})$. The problem is that during the numerical computation, even in the case of a very precise solution, we obtain some small numerical error in displacement values. So as the theoretical value of gap function is supposed to be zero on the active contact zone, it could be numerically approximated either by a small positive or negative value. If the value is negative, $H(-g_n(\mathbf{u}))\delta g_n(\mathbf{v}) = \delta g_n(\mathbf{v})$, which is correct but for the case of small positive value $H(-g_n(\mathbf{u}))\delta g_n(\mathbf{v}) = 0$, as if the point was not in contact at all, which is wrong.

This discrepancy could be solved by the artificial replacement of $H(-g_n(\mathbf{u}))\delta g_n(\mathbf{v})$ with $H(\varepsilon_g - g_n(\mathbf{u}))\delta g_n(\mathbf{v})$. The shifted Heaviside function ensures that even if the numerical error causes the originally non-positive gap function value to be positive but smaller than ε_g, the incriminated point is still considered to be in contact. It is complicated to give some general rules for choosing the value of ε_g as it depends on many factors. In the numerical examples presented in the further chapters, we chose $\varepsilon_g = 10^{-9}h$. Parameter h is the characteristic size of the mesh from the finite element discretisation, which we also utilise in the scaling parameter of penalty values. However, the interested reader should bear in mind that this choice is not unique and should be made only after taking into consideration the solved problem, chosen linear solver and generally the characteristic size of the expected numerical error. In Nitsche-FPR, the choice of contact indicator is more stable in these terms, as there is a positive multiple of the normal traction added to the gap. Thus, even if the zero gap is approximated by a small positive value, there is still a chance that the traction is negative and so we get a negative result, which is correctly interpreted as activated contact.

So far, it has been demonstrated that the symmetric versions of Nitsche-Wriggers and Nitsche-FPR coincide on $\tilde{\Gamma}^{C,S}$. However, this definitely does not hold for the inactive part of the potential contact zone, denoted as $\Gamma^{C,S} \setminus \tilde{\Gamma}^{C,S}$. Here, the expression in Macaulay brackets is negative and the whole Macaulay bracket is therefore equal to zero. Let us remark that this region could be different for Nitsche-Wriggers and Nitsche-FPR as the contact indicators of both methods are different as well. On $\Gamma^{C,S} \setminus \tilde{\Gamma}^{C,S}$, the contributions of Nitsche-Wriggers

and Nitsche-FPR to the energy functional are different and they read

$$\begin{aligned}\Pi_{FPR,C}^{\Gamma^{C,S}\setminus\tilde{\Gamma}^{C,S}} &= \int_{\Gamma^{C,S}\setminus\tilde{\Gamma}^{C,S}} \left(-\frac{h}{2\gamma E}\sigma_n^2 + \frac{\gamma E}{2h}\left\{-g_n - \frac{h}{\gamma E}\sigma_n\right\}^2\right) \mathrm{d}A \\ &= \int_{\Gamma^{C,S}\setminus\tilde{\Gamma}^{C,S}} \left(-\frac{h}{2\gamma E}\sigma_n^2\right) \mathrm{d}A\end{aligned} \tag{146}$$

$$\begin{aligned}\Pi_{W,C}^{\Gamma^{C,S}\setminus\tilde{\Gamma}^{C,S}} &= \int_{\Gamma^{C}\setminus\tilde{\Gamma}^{C,S}} \left(-\sigma_n\{-g_n\} + \frac{\gamma E}{2h}\{-g_n\}^2\right) \mathrm{d}A \\ &= 0\end{aligned} \tag{147}$$

Therefore, on the inactive part of the potential contact zone, Nitsche-Wriggers does not modify the elastic energy functional in any way, as we would intuitively expect. In contrast to this, in Nitsche-FPR, there is a negative semidefinite functional added to the elastic energy. That leads to a certain reduction of stiffness. However, this reduction does not have any effect on the solution of the problem, because if there is no contact, then $\sigma_n = 0$ and the whole integral over the inactive contact zone vanishes. However, numerical approximation in the finite element method may lead to certain errors and so to differences between the solution predicted by Nitsche-Wriggers and Nitsche-FPR. Also, unless the active contact zone is known a priori and is discretised in such a way that all the faces (in 3D) or edges (in 2D) discretising the active contact zone are not in contact along their entire length, there will always be only partially activated segments near the boundary of the active contact zone. More precisely, there exist such segments of the active contact zone where the contact indicator changes sign along the segment. When calculating the integral over the potential contact zone numerically, we consider only the contribution of 'active' integration points. In other words, only the contribution of integration points located at the active portion of the segment is considered. Because both methods — Nitsche-Wriggers and Nitsche-FPR — differ in the definition of the contact indicator, they could possibly consider slightly different portions of the partially activated segments to be in contact. Consequently, each method could possibly evaluate the integral over these partially activated segments in the different number of surface integration points. This contributes to the differences in the numerical solution predicted by Nitsche-Wriggers and Nitsche-FPR.

What came to us as a surprise was that for the numerical examples we tested, the Nitsche-Wriggers method was not able to converge (in the sense of some limited number of iterations) in most of the cases. On the contrary, the Nitsche-FPR method mostly did not suffer from convergence difficulties, except for the cases when the provided penalty value was not large enough. In some specific cases, the convergence of Nitsche-Wriggers improved after we introduced the shift of the Heaviside function described above in this section. However, for most of the tested problems, this adjustment did not help. To clarify this peculiar behaviour, we investigated a simple two-dimensional example with only two degrees of freedom, which is described in chapter 4.2. It turned out that the reason for the convergence problems is that the function which we obtain as the left-hand side of the discretised weak form of Nitsche-Wriggers is not continuous with respect to two unknown DOFs. This causes that the convergence of Newton's method, which we use to linearise our non-linear problem in every iteration step, is not guaranteed. In our experience, Newton's method usually fails to converge in the close proximity of the solution. One such example is documented in 4.2.4. This suggests that generally, the discretised energy functional is not a C^1-differentiable function of the unknown DOFs which causes the problems with the convergence of Newton's method. On the contrary, the function obtained as the left-hand side of the discretised weak form of the Nitsche-FPR method is continuous with respect to the unknown DOFs. For the symmetric variant ($\theta = -1$) this means that the discretised energy functional is the C^1-differentiable function. This can be seen even intuitively, because the weak form of Nitsche-FPR does not include the term with the Heaviside function (as is the case of Nitsche-Wriggers), which is responsible for the discontinuity of the discretised weak form of the Nitsche-Wriggers method.

3.3 Penalty method and its connection to Nitsche-Wriggers

This well-known method for the variational enforcement of the contact boundary conditions has a relatively long history. Augarde and co-authors [37] mention that Babuška ([5]) was the first one who used penalty method for enforcing the Dirichlet boundary conditions in terms of the finite element method. The idea of the method is elegant and simple — to increase the magnitude of energy accordingly to the mutual penetration of bodies. To derive the variational form of the method, let us remember the equation (25) which we used to derive Nitsche methods, which reads

$$\begin{aligned}&\int_{\Omega} \boldsymbol{\sigma}\left(\mathbf{u}\right) : \boldsymbol{\nabla}\mathbf{v}\,\mathrm{d}V - \int_{\Omega} \rho\overline{\mathbf{b}}\cdot\mathbf{v}\,\mathrm{d}V - \int_{\Gamma^{\sigma}} \overline{\mathbf{t}}\cdot\mathbf{v}\,\mathrm{d}A \\ &+ \int_{\Gamma^{C,S}} \sigma_n^S\left[\mathbf{v}\right]\cdot\mathbf{n}^M\,\mathrm{d}A = 0 \quad \forall\mathbf{v}\in\mathbf{V}\end{aligned} \tag{148}$$

In this derivation we will follow a way suggested in Yastrebov's book [57, Chapter 4.5.1, p.145-146] and we will adopt some of his notation as well. As in the case of every other method, our goal is to satisfy the following Hertz-Signorini-Moreau conditions defined by (17)

$$\begin{aligned} g_n &\geq 0 \quad \text{(i)} \\ \sigma_n &\leq 0 \quad \text{(ii)} \\ \sigma_n g_n &= 0 \quad \text{(iii)} \end{aligned} \tag{149}$$

Trying to achieve this goal, let us define the normal contact traction $\sigma_n(\mathbf{u})$, supposing that it is a continuous function of the penetration, as

$$\sigma_n \approx P(\{-g_n\}) = \begin{cases} 0, & g_n > 0 \\ P(-g_n), & g_n \leq 0 \end{cases} \tag{150}$$

According to [57, Chapter 4.5.1, p.145], $P(x)$ is demanded to be a non-positive, continuous and strictly monotonically decreasing function. Also, it is required that

$$P(0) = 0 \tag{151}$$

$$P(x) \overset{x\to\infty}{\longrightarrow} -\infty \tag{152}$$

These assumptions are only logical considering the function $P(x)$ is approximating contact traction and is dependent on the gap function. We would like to point out that $P(x)$ does not have to be linear in x to fulfil the previously mentioned properties though we will prefer to stick to the further described linear option. Having $P(x)$ fulfilling the previous conditions, we see that if $g_n > 0$, then $\sigma_n = 0$ and so $\sigma_n g_n = 0$. In this case, contact conditions are fulfilled exactly. However, when $g_n < 0$ then $\sigma_n = P(\{-g_n\}) < 0$ and so $\sigma_n g_n \neq 0$. Therefore, there could not be any non-zero contact traction without non-zero penetration. This suggests that using penalty method, contact conditions are never fulfilled exactly, as a contact surface is never fully restricted from the penetration but rather resists it with certain stiffness governed by function $\dfrac{\partial P(\{-g_n\})}{\partial g_n}$. This behaviour is often compared to a model where the surface is not impenetrable but is modelled by springs with a zero initial length and stiffness given by $P(\{-g_n\})$. These springs are supposed to deform in the direction normal to the surface (as we consider the case of normal contact). This can be also understood, as if certain constitutive equation for the behaviour of the 'surface layer' was introduced. In terms of the detailed micro-mechanical analysis of solids, this is actually a quite correct approach, as the 'surface layer' structure is very complex containing for example layers of contaminants, absorbed gas, oxide films etc. and it is reasonable to describe its stiffness by specially constructed function. The interested reader should see [55, Chapter 5.1.2, p.72-75] for more detail and some relevant literature recommendations.

Though $P(\{-g_n\})$ is subjected to the aforementioned requirements, the possible choice of such a function is definitely not unique. As mentioned previously, some constitutive equations for 'surface layer' could serve us as an inspiration. In this thesis, we will focus on a choice

$$P(\{-g_n\}) = -\frac{\gamma E}{h}\{-g_n\} \tag{153}$$

However, we remind once more that it is possible to choose a non-linear function as well and this choice could lead to a better performance of the method (see [57, Chapter 4.5.3, p.165-167] for more details, some examples and some particularly interesting choices). There are two reasons why we defined $P(\{-g_n\})$ as (153) in this thesis. Firstly, $P(\{-g_n\})$ defined in this way is a linear function of g_n, which is simple to implement. Secondly, this very choice has a connection with the Nitsche-Wriggers method as will be shown later. Supposing $\sigma_n \approx P(\{-g_n\})$ we can substitute (153) for σ_n^S to (25) ((25) is reminded in the beginning of this chapter) and we obtain

$$\begin{aligned}&\int_\Omega \boldsymbol{\sigma}\left(\mathbf{u}\right) : \boldsymbol{\nabla}\mathbf{v}\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\mathbf{v}\,\mathrm{d}V - \int_{\Gamma^\sigma} \overline{\mathbf{t}}\cdot\mathbf{v}\,\mathrm{d}A\\ &-\frac{\gamma E}{h}\int_{\Gamma^{C,S}} \{-g_n(\mathbf{u})\}\left[\mathbf{v}\right]\cdot\mathbf{n}^M\,\mathrm{d}A = 0 \quad \forall\mathbf{v}\in\mathbf{V}\end{aligned} \tag{154}$$

Reminding that

$$g_n(\mathbf{u}) = \left(\mathbf{x}^S - \mathbf{x}^M\right)\cdot\mathbf{n}^M = \left(\mathbf{X}^S + \mathbf{u}^S - \mathbf{X}^M - \mathbf{u}^M\right)\cdot\mathbf{n}^M \tag{155}$$

$$\delta g_n(\mathbf{v}) = \mathbf{v}^S\cdot\mathbf{n}^M - \mathbf{v}^M\cdot\mathbf{n}^M = \left[\mathbf{v}\right]\cdot\mathbf{n}^M \tag{156}$$

we can replace $\left[\mathbf{v}\right]\cdot\mathbf{n}^M$ in (154) with δg_n which yields

$$\begin{aligned}&\int_\Omega \boldsymbol{\sigma}\left(\mathbf{u}\right) : \boldsymbol{\nabla}\mathbf{v}\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\mathbf{v}\,\mathrm{d}V - \int_{\Gamma^\sigma} \overline{\mathbf{t}}\cdot\mathbf{v}\,\mathrm{d}A\\ &-\frac{\gamma E}{h}\int_{\Gamma^{C,S}} \{-g_n(\mathbf{u})\}\delta g_n(\mathbf{v})\,\mathrm{d}A = 0 \quad \forall\mathbf{v}\in\mathbf{V}\end{aligned} \tag{157}$$

which completes our formulation of the penalty method. Here, $\frac{\gamma E}{h}$ is the scaled penalty parameter chosen in exactly the same way as for the Nitsche methods and the augmented Lagrangian method.

Though the method is very easy to understand and implement, it suffers from certain drawbacks. As reported in Wriggers' book [55, Chapter 6.3.2, p.118], contact conditions are fulfilled exactly (on the level of the continuum problem) only if $\gamma \to \infty$ (see Luenberger [38] for more detail). In numerical computation, only finite value of γ could be considered. For that reason, to fulfil contact conditions with sufficiently small error, we are forced to choose as high value of γ as possible. However, this leads to the ill-conditioning of the system of equations obtained when solving the problem numerically with the finite element method [55, Chapter 6.3.2, p.119].

Looking at (157), we could see that the equation is very similar to the one for the Nitsche-Wriggers equation. Let us remind here the equation of Nitsche-Wriggers method (113)

$$\begin{aligned}&\int_\Omega \boldsymbol{\sigma}\left(\mathbf{u}\right) : \boldsymbol{\nabla}\mathbf{v}\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\mathbf{v}\,\mathrm{d}V - \int_{\Gamma^\sigma} \overline{\mathbf{t}}\cdot\mathbf{v}\,\mathrm{d}A\\ &+\int_{\Gamma^{C,S}} \langle\sigma_n(\mathbf{u})\rangle H(-g_n(\mathbf{u}))\delta g_n(\mathbf{v})\,\mathrm{d}A - \int_{\Gamma^{C,S}} \langle\sigma_n(\mathbf{v})\rangle\{-g_n(\mathbf{u})\}\,\mathrm{d}A\\ &-\frac{\gamma E}{h}\int_{\Gamma^{C,S}} \{-g_n(\mathbf{u})\}\delta g_n(\mathbf{v})\,\mathrm{d}A = 0 \quad \forall\mathbf{v}\in\mathbf{V}\end{aligned} \tag{158}$$

If we eliminate some terms in (158), namely the integral

$$\int_{\Gamma^{C,S}} \langle\sigma_n(\mathbf{u})\rangle H(-g_n(\mathbf{u}))\delta g_n(\mathbf{v})\,\mathrm{d}A \tag{159}$$

and

$$\int_{\Gamma^{C,S}} \langle \sigma_n(\mathbf{v}) \rangle \{-g_n(\mathbf{u})\} \, \mathrm{d}A \tag{160}$$

we get

$$\begin{aligned} &\int_\Omega \boldsymbol{\sigma}\left(\mathbf{u}\right) : \boldsymbol{\nabla}\mathbf{v} \, \mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}} \cdot \mathbf{v} \, \mathrm{d}V - \int_{\Gamma^\sigma} \overline{\mathbf{t}} \cdot \mathbf{v} \, \mathrm{d}A \\ &- \frac{\gamma E}{h} \int_{\Gamma^{C,S}} \{-g_n(\mathbf{u})\} \delta g_n(\mathbf{v}) \, \mathrm{d}A = 0 \quad \forall \mathbf{v} \in \mathbf{V} \end{aligned} \tag{161}$$

which is exactly the same as the weak form of the penalty method (157). Therefore, we see that penalty method could be understood as an inconsistent version of the Nitsche-Wriggers method where certain contributions to the integral over the potential contact zone are missing. Although the chosen formulation of penalty method leads to certain disadvantages (such as the ill-conditioning of the linear system solved in every iteration step), it has its vast advantages in its simplicity and the fact that there exists an energy functional which, if discretised, is C^1-differentiable function with respect to the unknown DOFs, and reads

$$\begin{aligned} \Pi_{PEN}(\mathbf{u}) = &\frac{1}{2} \int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{u}) \, \mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}} \cdot \mathbf{u} \, \mathrm{d}V - \int_{\Gamma^\sigma} \overline{\mathbf{t}} \cdot \mathbf{u} \, \mathrm{d}A \\ &+ \frac{1}{2} \frac{\gamma E}{h} \int_{\Gamma^{C,S}} \{-g_n(\mathbf{u})\}^2 \, \mathrm{d}A \end{aligned} \tag{162}$$

Variation of this functional defined as

$$\delta\Pi_{PEN}(\mathbf{u}, \mathbf{v}) = \frac{\mathrm{d}}{\mathrm{d}\alpha} \Pi_{PEN}(\mathbf{u} + \alpha\mathbf{v}) \bigg|_{\alpha=0} = 0 \quad \forall \mathbf{v} \in \mathbf{V} \tag{163}$$

yields again equation (157). Again, we remind the following relation, which is very useful when deriving the variation

$$\frac{d}{dx} \left(\frac{1}{2} \{x\}^2 \right) = \{x\} \cdot H(x) = \{x\} \tag{164}$$

The numerical examples we conducted (documented in the following chapters) suggest that if the computation converges, then the norm of the error (either in displacement or stress) is mostly smaller for Nitsche-Wriggers than for the described penalty method. However, the penalty method proved itself to be very stable compared to Nitsche-Wriggers. The details of this behaviour will be given in the following chapters.

3.4 Augmented Lagrangian and its connection to Nitsche-FPR

As Yastrebov writes in his book ([57, Chapter 1.1.1, p.7]), the augmented Lagrangian method was firstly proposed by Arrow and Solow in 1958 (in a very raw version and not in connection with contact problems). In 1969 this was followed by Hestens [28] and Powell [45] who both worked out the method independently. It should be noted that both authors were concerned with enforcing the equality constraint. This was generalised to an inequality constraint by Rockafellar ([46] and [47]) in 1973. In this thesis, we will follow Rockafellar's formulation, as stated in Yasterbov's book [57, Chapter 4.7.2, p.179-180]. Concerning contact mechanics, Yastrebov states ([57, Chapter 1.1.1, p.7]) that first applications of the augmented Lagrangian method to the frictionless contact problem could be traced to Glowinski and Le Tallec [21] and Wriggers, Simo and Taylor [40].

Despite its relatively long history, the augmented Lagrangian method still remains one of the most popular methods for solving contact problems today as it possesses many convenient properties. In order to understand them, it is very useful to compare them with the Lagrange multiplier method. The methods are related as they both introduce another unknown variable λ called a Lagrange multiplier. Concerning contact problems, the physical meaning of a Lagrange multiplier is the normal contact traction on the contact part of the boundary where this additional unknown variable is defined. In the Lagrange multiplier method, introducing additional unknown allows the contact boundary conditions to be fulfilled exactly (in a theoretical sense on the continuum level) without adding penalty term at all. This sounds very tempting compared to the penalty method, where the contact boundary conditions are enforced exactly only for the infinite penalty value. This forces us to choose high penalty values to keep the error small and results in the possible ill-conditioning of the solved system. However, the price for the absence of the penalty value in the Lagrange multiplier method is that the corresponding energy functional (if discretised) is not sufficiently smooth function (with respect to the unknown DOFs), namely not C^1-differentiable ([55, Chapter 6.3.8, p.126]). This is very inconvenient in terms of Newton's method, which we will use for the linearisation of the stationarity conditions of the functional. The idea of the augmented Lagrangian method is to overcome these difficulties by adding a certain penalty term to other standard terms in the Lagrange multiplier method. It turns out that if this is done in a right way, the obtained energy functional (if discretised) is not only C^1-differentiable function with respect to the unknown DOFs, but the method fulfils the contact conditions exactly (on the continuum level).

Moreover, the Lagrange multiplier method introduces another unknown variable λ, but as it has a physical meaning of contact traction, it is still restricted by contact conditions (17). For this reason, in the Lagrange multiplier method, the additional condition $\lambda \leq 0$ is required when solving the contact problem. In terms of the optimization theory, this kind of problem is referred to as a constrained one. This poses a certain drawback, as we would very much prefer to have a fully unconstrained problem as it makes the method rather easier to implement in any finite element software. As Yastrebov comments ([57, Chapter 1.1.1, p.7]), the reason is that because of inequality constraints, the formulation has to be considered in combination with a so-called active set strategy (see for instance [38] for more detail). We note that an active set strategy is a check and update of active and passive constraints. In FEniCS, this is not impossible, but it is rather complicated. For this reason, we would prefer to have some unconstrained version of the Lagrange multiplier method. It happens that augmented Lagrangian is precisely such a method, converting an originally constrained problem into an unconstrained saddle point problem often called the min-max problem. In the context of mechanics, this means that during the numerical solution of the discrete finite element problem, we seek for the minimum with respect to displacement variables and at the same time for the maximum with respect to the Lagrange multiplier.

There are more approaches to solving the augmented Lagrangian saddle point problem. One of them is based on a consecutive independent updating of a primal variable (in this context displacements) and a dual variable

(here the Lagrange multiplier representing contact traction). During the iteration steps, firstly Lagrange multipliers are updated using a certain algebraic formula. After that, for the fixed value of Lagrange multiplier, minimization problem for primal variable (displacement in our case) is solved using standard techniques. Yastrebov mentions this idea has been originally proposed by Powell [45] and is still employed today under the name Uzawa's algorithm (named in honour of Japanese economist Hirofumi Uzawa). Fletcher [18] has developed another numerical approach for solving saddle point problem. Opposite to the previous technique where at each iteration step, the fixed value of a Lagrange multiplier is considered, Fletcher solves the saddle point problem as a whole and during the iteration steps updates both variables simultaneously.

It is also important to mention that a suitable combination of function spaces of approximation (shape) functions (for displacements and a Lagrange multiplier) for the finite element approximation must be chosen to achieve stability. This does not necessarily have to be the case for some naive choices. Generally, the approximation function spaces must be chosen to fulfil the Babuška-Brezzi condition (see [10] for more details on the Babuška-Brezzi condition). It should be mentioned that Barbosa and Hughes [6] proposed a stabilized method which allows us to circumvent the discrete Babuška-Brezzi condition.

Although the augmented Lagrangian method possesses many above mentioned qualities, there are also certain drawbacks. Firstly, the number of DOFs of the discrete finite element problem is significantly higher than in the case of penalty or Nitsche methods, as there are additional DOFs of Lagrange multipliers. This could lead to a higher computational time of the iterations during the solution. Yastrebov reports this could be reduced by nesting Lagrange multipliers in the local update procedure using Uzawa's algorithm ([57, Chapter 4.7.1, p.173]). On the other hand, Uzawa's algorithm is reported to produce only linear convergence rate ([57, Chapter 4.7.1, p.173]). In this thesis, we will focus on Rockafellar's approach (see [57, Chapter 4.7.2, p.180]) for its connection with the Nitsche-FPR method, and this one is reported to suffer from ill-conditioning for high penalty coefficients [57, Chapter 4.7.3, p.187]. This could again motivate using Uzawa's algorithm though we will prefer continuous updating of both variables as it is more natural to FEniCS.

After briefly discussing the history and the general background of the method, let us proceed to its variational formulation. In the case of the linear elastic frictionless contact problem there exists an energy functional $\Pi_{AL}(\mathbf{u}, \lambda)$. We will adopt it in a form presented by Yastrebov in [57, Chapter 4.7.2, p.197-180], which in turn is based on the work of Rockafellar ([46] and [47]). This provides us with a fully unconstrained problem. If the obtained energy functional is discretised, it is a C^1-differentiable function with respect to the unknown DOFs [55, Chapter 6.3.8, p.126]. The functional of energy of the augmented Lagrangian method is the so called saddle point energy functional and reads

$$\Pi_{AL}(\mathbf{u}, \lambda) = \frac{1}{2}\int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{u})\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\mathbf{u}\,\mathrm{d}V - \int_{\Gamma^\sigma}\overline{\mathbf{t}}\cdot\mathbf{u}\,\mathrm{d}A \tag{165}$$

$$- \int_{\Gamma^C}\frac{h}{2\gamma E}\left(\lambda^2 - \left\{-\lambda - \frac{\gamma E}{h}g_n(\mathbf{u})\right\}^2\right)\mathrm{d}A \tag{166}$$

Here $\mathbf{u} \in \mathbf{U}^1$ (same as in the previous methods)[8] and $\lambda \in W$, where

$$W = L^2\left(\Gamma^C\right) \tag{167}$$

Variable λ represents the normal contact traction, as is discussed in more detail further in this chapter and is independent of $\mathbf{u}$. Based on the nomenclature from the optimization theory, λ is also called Lagrange multiplier.

[8] $\mathbf{U}^s = \{\mathbf{u} \in H^s(\Omega)^d\,; \mathbf{u} = \overline{\mathbf{u}} \text{ on } \Gamma^D\}$

Now, let us write the stationarity conditions of $\Pi_{AL}(\mathbf{u},\lambda)$, which could be expressed as

$$\delta\Pi_{AL}(\mathbf{u},\lambda,\mathbf{v}) = \left.\frac{\mathrm{d}}{\mathrm{d}\alpha}\Pi_{AL}(\mathbf{u}+\alpha\mathbf{v},\lambda)\right|_{\alpha=0} = 0 \quad \forall\mathbf{v}\in\mathbf{V} \tag{168}$$

$$\delta\Pi_{AL}(\mathbf{u},\lambda,\delta\lambda) = \left.\frac{\mathrm{d}}{\mathrm{d}\beta}\Pi_{AL}(\mathbf{u},\lambda+\beta\delta\lambda)\right|_{\beta=0} = 0 \quad \forall\delta\lambda\in W \tag{169}$$

Let us remind the relation (111) for the derivative of the 'ramp' function (defined using Macaulay brackets) and the definitions of the gap function and its variation

$$\frac{d}{dx}\left(\frac{1}{2}\{x\}^2\right) = \{x\}\cdot H(x) = \{x\} \tag{170}$$

$$g_n(\mathbf{u}) = \left(\mathbf{x}^S - \mathbf{x}^M\right)\cdot\mathbf{n}^M = \left(\mathbf{X}^S + \mathbf{u}^S - \mathbf{X}^M - \mathbf{u}^M\right)\cdot\mathbf{n}^M \tag{171}$$

$$\delta g_n(\mathbf{v}) = \mathbf{v}^S\cdot\mathbf{n}^M - \mathbf{v}^M\cdot\mathbf{n}^M = [\mathbf{v}]\cdot\mathbf{n}^M \tag{172}$$

With these, we can write

$$\begin{aligned}\delta\Pi_{AL}(\mathbf{u},\lambda,\mathbf{v}) &= \int_\Omega \boldsymbol{\sigma}(\mathbf{u}):\boldsymbol{\varepsilon}(\mathbf{v})\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\mathbf{v}\,\mathrm{d}V - \int_{\Gamma^\sigma}\overline{\mathbf{t}}\cdot\mathbf{v}\,\mathrm{d}A \\ &- \int_{\Gamma^C}\left\{-\lambda - \frac{\gamma E}{h}g_n(\mathbf{u})\right\}\delta g_n(\mathbf{v})\,\mathrm{d}A\end{aligned} \tag{173}$$

$$\begin{aligned}\delta\Pi_{AL}(\mathbf{u},\lambda,\delta\lambda) &= \\ &- \int_{\Gamma^C}\frac{h}{\gamma E}\left(\lambda\delta\lambda + \left\{-\lambda - \frac{\gamma E}{h}g_n(\mathbf{u})\right\}\delta\lambda\right)\mathrm{d}A\end{aligned} \tag{174}$$

The previous equations represent the variations of $\Pi_{AL}(\mathbf{u},\lambda)$ with respect to $\mathbf{u}$ (173) and λ (174). Requiring these variations to be zero $\forall\mathbf{v}\in\mathbf{V}$ and $\forall\delta\lambda\in W$ we obtain the stationarity conditions of $\Pi_{AL}(\mathbf{u},\lambda)$, which can be written as

$$\begin{aligned}&\int_\Omega \boldsymbol{\sigma}(\mathbf{u}):\boldsymbol{\varepsilon}(\mathbf{v})\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\mathbf{v}\,\mathrm{d}V - \int_{\Gamma^\sigma}\overline{\mathbf{t}}\cdot\mathbf{v}\,\mathrm{d}A \\ &- \int_{\Gamma^C}\left\{-\lambda - \frac{\gamma E}{h}g_n(\mathbf{u})\right\}\delta g_n(\mathbf{v})\,\mathrm{d}A = 0 \quad \forall\mathbf{v}\in\mathbf{V}\end{aligned} \tag{175}$$

$$-\int_{\Gamma^C}\frac{h}{\gamma E}\left(\lambda\delta\lambda + \left\{-\lambda - \frac{\gamma E}{h}g_n(\mathbf{u})\right\}\delta\lambda\right)\mathrm{d}A = 0 \quad \forall\delta\lambda\in W \tag{176}$$

Factoring the positive fraction $h/(\gamma E)$ out of Macaulay brackets, we can easily write

$$\begin{aligned}&\int_\Omega \boldsymbol{\sigma}(\mathbf{u}):\boldsymbol{\varepsilon}(\mathbf{v})\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\mathbf{v}\,\mathrm{d}V - \int_{\Gamma^\sigma}\overline{\mathbf{t}}\cdot\mathbf{v}\,\mathrm{d}A \\ &- \int_{\Gamma^C}\frac{\gamma E}{h}\left\{-g_n(\mathbf{u}) - \frac{h}{\gamma E}\lambda\right\}\delta g_n(\mathbf{v})\,\mathrm{d}A = 0 \quad \forall\mathbf{v}\in\mathbf{V}\end{aligned} \tag{177}$$

$$-\int_{\Gamma^C}\frac{h}{\gamma E}\left(\lambda\delta\lambda + \frac{\gamma E}{h}\left\{-g_n(\mathbf{u}) - \frac{h}{\gamma E}\lambda\right\}\delta\lambda\right)\mathrm{d}A = 0 \quad \forall\delta\lambda\in W \tag{178}$$

With equations (177) and (178), the weak form of the augmented Lagrangian method is formulated and could be further discretised and solved by the finite element method. However, the question remains whether this method, which is nowadays standard for solving contact problems, is somehow connected to the previously discussed Nitsche methods. In [31], the authors argue that augmented Lagrangian is really connected with Nitsche-FPR method (134) for $\theta = 0$. Inspired by their reasoning, we will now discuss this connection in more detail.

Looking at the equation (177), we can recognize that if we formally replace λ with $\langle\sigma_n(\mathbf{u})\rangle_t$, we get exactly the equation of Nitsche-FPR method (134) for $\theta = 0$. For the purpose of clarity, let us remind the equation of Nitsche-FPR method (134)

$$\begin{aligned}&\int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\nabla}\mathbf{v}\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\mathbf{v}\,\mathrm{d}V - \int_{\Gamma^\sigma}\overline{\mathbf{t}}\cdot\mathbf{v}\,\mathrm{d}A\\ &+\int_{\Gamma^{C,S}}\theta\frac{h}{\gamma E}\langle\sigma_n(\mathbf{u})\rangle_t\langle\sigma_n(\mathbf{v})\rangle_t\,\mathrm{d}A\\ &-\int_{\Gamma^{C,S}}\frac{\gamma E}{h}\left\{-g_n(\mathbf{u})-\frac{h}{\gamma E}\langle\sigma_n(\mathbf{u})\rangle_t\right\}\left(\delta g_n(\mathbf{v})-\theta\frac{h}{\gamma E}\langle\sigma_n(\mathbf{v})\rangle_t\right)\mathrm{d}A = 0 \quad \forall\mathbf{v}\in\mathbf{V}\end{aligned} \tag{179}$$

For $\theta = 0$ we get

$$\begin{aligned}&\int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\nabla}\mathbf{v}\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\mathbf{v}\,\mathrm{d}V - \int_{\Gamma^\sigma}\overline{\mathbf{t}}\cdot\mathbf{v}\,\mathrm{d}A\\ &-\int_{\Gamma^{C,S}}\frac{\gamma E}{h}\left\{-g_n(\mathbf{u})-\frac{h}{\gamma E}\langle\sigma_n(\mathbf{u})\rangle_t\right\}\delta g_n\,\mathrm{d}A = 0 \quad \forall\mathbf{v}\in\mathbf{V}\end{aligned} \tag{180}$$

which is equation (177) in which λ is replaced with $\langle\sigma_n(\mathbf{u})\rangle_t$. From this we see that there is a connection between the augmented Lagrangian method and the Nitsche-FPR method for $\theta = 0$. Moreover, we can easily observe that equation (178), the second stationarity condition of functional $\Pi_{AL}(\mathbf{u},\lambda)$, is actually the weak formulation of condition (118), in which $\sigma_n(\mathbf{u})$ is replaced with λ. We can verify this, remembering the condition (118)

$$\sigma_n(\mathbf{u}) = -\frac{\gamma E}{h}\left\{-g_n(\mathbf{u})-\frac{h}{\gamma E}\sigma_n(\mathbf{u})\right\} \tag{181}$$

Here, compared to (118) we consider the penalty value $\gamma_{FPR} = h/(\gamma E)$ (as we generally do in a Nitsche-FPR method). If we consider the general case of two elastic bodies in contact, $\sigma_n(\mathbf{u})$ can be equivalently replaced with $\langle\sigma_n(\mathbf{u})\rangle_t$ defined in (131). As mentioned in the chapter on Nitsche-FPR, this ensures the stability of the method. If we perform this substitution in (181) (i.e. replace λ with $\langle\sigma_n(\mathbf{u})\rangle_t$), we get

$$\langle\sigma_n(\mathbf{u})\rangle_t = -\frac{\gamma E}{h}\left\{-g_n(\mathbf{u})-\frac{h}{\gamma E}\langle\sigma_n(\mathbf{u})\rangle_t\right\} \tag{182}$$

Now, as equation (178) holds $\forall\delta\lambda\in W$, we can write

$$\lambda = -\frac{\gamma E}{h}\left\{-g_n(\mathbf{u})-\frac{h}{\gamma E}\lambda\right\} \tag{183}$$

Finally substituting λ for $\langle\sigma_n(\mathbf{u})\rangle_t$ results in equation (182).

Summarizing our results, the augmented Lagrangian method can be understood as the Nitsche-FPR method with $\theta = 0$, where the normal contact traction $\sigma_n(\mathbf{u})$ is replaced with an unknown variable $\lambda\in W$ independent of $\mathbf{u}$ (in this context called a Lagrange multiplier). However, the value of the Lagrange multiplier (representing the normal contact traction) cannot acquire arbitrary values and is restricted by contact conditions (17) which are equivalently expressed as (118). To fulfil this, the weak form of (183) is introduced, formulated as (178). The relation between both methods is currently the matter of scientific research, as [31] mentions. Some further results on the connection of the augmented Lagrangian and Nitsche methods can be found in [12], [13] and [23]. Also, it turns out that the mentioned stabilized method by Barbosa and Hughes [6], which allow us to circumvent the discrete Babuška-Brezzi condition, has a very strong connection with the Nitsche-FPR method [31].

4 Numerical simulations

4.1 Objectives of the simulations and the FEniCS platform

In this chapter, we will present various numerical simulations conducted in FEniCS, which is an open-source computing platform (for Windows, Linux or Mac) for solving partial differential equations [16]. When formulating a problem in FEniCS, firstly the weak form of the solved problem is provided. It is declared using Unified Form Language (UFL) [54], which is also part of FEniCS. This language was purposefully created to closely remind 'handwritten' formulation so for someone who is familiar with this mathematical notation, the formulation of problem in FEniCS is very convenient and natural. Then, the user is allowed to choose approximation spaces for finite element approximation functions and define the boundary conditions. Based on these, FEniCS generates a finite element code and solves the resulting discrete problem. FeniCS is equipped with a high-level Python and C++ interfaces and especially the Python interface makes FEniCS very understandable even to inexperienced programmers. Another advantage of the FEniCS platform is that it is designed for parallel processing [16] and running scripts in parallel is very simple. This allows the user to relatively easily prepare his parallel code on his computer and then run this code in large-scale on high-performance computers.

From the mentioned properties of FEniCS we see that it is very convenient for our purposes. That is because firstly, we investigate various methods based on variational equalities in this thesis, and these can be naturally implemented in FEniCS, as this requires the user only to provide the weak formulation of the problem in the UFL language. Secondly, we need to repeatedly solve the same physical problem using different weak forms, which is also easy to do in FEniCS.

The purpose of the conducted simulations is to compare the accuracy and the computational efficiency of the investigated methods for enforcing the contact boundary conditions. Namely the penalty method, the augmented Lagrangian method, the Nitsche-Wriggers method and the Nitsche-FPR method (for $\theta = -1$, $\theta = 0$ and $\theta = 1$). The weak formulations of all these methods were implemented in FEniCS. Although the methods were formulated for the case of contact of two elastic bodies, in numerical examples, we will restrict ourselves only to the contact of one elastic body with a rigid plane (this type of problem is often called the Signorini problem [57, Chapter 1.1, p.5]). The reason for this is that the contact detection is much easier in this case and does not require any additional sophisticated algorithms, which importantly influence the efficiency and overall performance of the implemented codes. Also, the same algorithm for contact detection could interact differently with various methods for enforcing the contact boundary conditions. As the investigation of these interactions is not an objective of this thesis, we avoid these complications by choosing such a problem where the contact detection is trivial. Because we formulated our methods for the case of contact of two elastic bodies, the question arises how to deal with the jump and average operators we used. In order to avoid the possible inconsistency in the notation, we define the normal contact tractions on the active contact zone of the rigid obstacle to be equal to the ones on the corresponding active contact zone of the elastic body. Formally, this is correct, because the displacement of the rigid body is zero in every material point, but the stress and tractions can be arbitrary. Equipped with these definitions of tractions, we can formally apply the average and jump operators and we can write σ_n instead of $\langle\sigma_n(\mathbf{u})\rangle$ or $\langle\sigma_n(\mathbf{u})\rangle_t$, $\delta\sigma_n$ instead of $\langle\sigma_n(\mathbf{v})\rangle$ or $\langle\sigma_n(\mathbf{v})\rangle_t$.

Concerning the augmented Lagrangian method, we also utilized the multiphenics library. This package allows defining Lagrange multipliers only on the potential contact zone. Without multiphenics, FEniCS allows defining the Lagrange multipliers field only on the whole body. Of course, some 'hacks' could be utilized to circumvent this problem, but multiphenics offers a much cleaner way of implementing the augmented Lagrangian

method in FEniCS. We would like to point out that the multiphenics library does not include a class calling Newton's method non-linear solver in FEniCS, so we implemented it ourselves.

Because the solved weak formulations are generally non-linear (as the active contact zone $\overline{\Gamma}^C$ is not known in forward), we applied Newton's method. The solved weak formulation was therefore linearized in every iteration step. In terms of contact mechanic, every step of Newton's method could be understood as the guess of unknown variables (which always include displacements), i.e. the guess of the active contact zone. Because the problem is otherwise linear (we suppose small deformation and linear elastic body), sparse LU decomposition (Gaussian elimination) [52] is used to solve the problem in every Newton's method iteration step. In order to simplify the solved problem, integration is performed over the original instead of the current configuration, contrarily to what we supposed in theoretical derivations. However, with the supposed small penetrations and resulting deformations, the resulting error is very small and therefore negligible. Concerning integration, FEniCS uses Gauss–Legendre–Jacobi rule mapped onto simplices for polynomials of degree higher than six. For integrating polynomials of a degree less than or equal to six, a few other rules are implemented [36]. Without any additional programming, the user is allowed to set the degree of polynomials that would be integrated exactly. If one does not set it, the degree is chosen automatically by FEniCS. We keep the default setting for integration over volume as the volume integrals contain very standard terms in our case, but for integration over the surface, we implemented our own scheme (based on [39] which we used with the courtesy of Dr. Jeremy Bleyer). This scheme uses a classic three-point Gaussian quadrature mapped onto simplices (unit segment line in 2D and canonical right 45° unit triangle in 3D)

4.2 Analytical solution of the 2D benchmark problem — Nitsche-Wriggers

From the numerical examples, which will be presented in the following sections, we will see that the Nitsche-Wriggers method often has serious troubles to converge in FEniCS. In order to understand this behaviour, we investigated a simple 2D example, the discretised geometry of which is depicted in Figure 3. The problem consists of two linear elements only, which allow us to solve the problem analytically (the Wolfram Mathematica software was also utilized for this purpose). Comparing the analytically derived functional and its derivatives with its numerical counterpart, we will demonstrate the differences and suggest the reasons why Nitsche-Wriggers suffers from convergence problems in FEniCS.

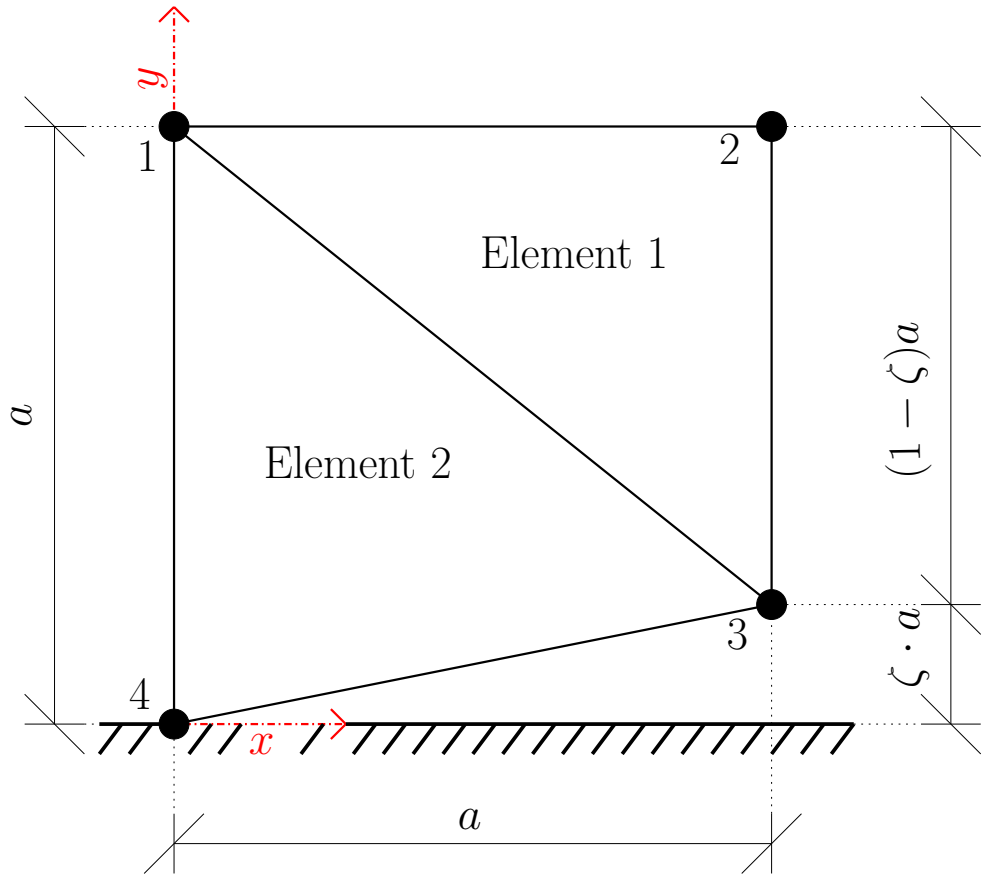


Figure 3: Geometry of the 2D benchmark example

Contact will be enforced with prescribed vertical displacements $v_1 = -C$, $v_2 = -C$ where $C > 0$. In this example, we will denote all vertical displacements (in y direction) by v - not to be confused with $\mathbf{v}$ - by which we understood test functions corresponding to displacement $\mathbf{u}$ in other chapters. As we consider small deformations and Poisson ratio is $\nu = 0$, all horizontal displacements (denoted by $\mathbf{u}$) are zero-valued. In order to avoid numerical error of lateral displacements in the numerical solution, we directly prescribe them to be zero-valued.

Now, let us remind the energy functional of the Nitsche-Wriggers method (114).

$$\begin{aligned}\Pi_W(\mathbf{u}) =& \frac{1}{2}\int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{u})\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\mathbf{u}\,\mathrm{d}V - \int_{\Gamma^\sigma}\overline{\mathbf{t}}\cdot\mathbf{u}\,\mathrm{d}A \\ &+ \int_{\Gamma^{C,S}} \langle\sigma_n(\mathbf{u})\rangle\left(-\{-g_n(\mathbf{u})\}\right)\,\mathrm{d}A + \frac{1}{2}\frac{\gamma E}{h}\int_{\Gamma^{C,S}}\left(-\{-g_n(\mathbf{u})\}\right)^2\,\mathrm{d}A\end{aligned} \tag{184}$$

We can separate the energy functional to four parts, as indicated below

$$\Pi_W(\mathbf{u}) = \Pi_{W,INT}(\mathbf{u}) + \Pi_{W,EXT}(\mathbf{u}) + \Pi_{W,NIT}(\mathbf{u}) + \Pi_{W,PEN}(\mathbf{u}) \tag{185}$$

where

$$\Pi_{W,INT}(\mathbf{u}) = \frac{1}{2}\int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{u})\,\mathrm{d}V \tag{186}$$

is an 'internal' elastic energy

$$\Pi_{W,EXT}(\mathbf{u}) = -\int_\Omega \rho\overline{\mathbf{b}}\cdot\mathbf{u}\,\mathrm{d}V - \int_{\Gamma^\sigma}\overline{\mathbf{t}}\cdot\mathbf{u}\,\mathrm{d}A \tag{187}$$

is the energy of external forces

$$\Pi_{W,NIT}(\mathbf{u}) = \int_{\Gamma^{C,S}} \langle\sigma_n(\mathbf{u})\rangle\left(-\{-g_n(\mathbf{u})\}\right)\,\mathrm{d}A \tag{188}$$

contains a term that secures symmetry of the corresponding bilinear form of the stationarity equation of the Nitsche-Wriggers energy functional and finally

$$\Pi_{W,PEN}(\mathbf{u}) = \frac{1}{2}\frac{\gamma E}{h}\int_{\Gamma^{C,S}} \left(-\{-g_n(\mathbf{u})\}\right)^2\,\mathrm{d}A \tag{189}$$

is the energy of the penalty term, providing positive definiteness to the corresponding bilinear form of stationarity equation of the Nitsche-Wriggers energy functional. For our discrete problem, these functional are converted into functions and these read

$$\begin{aligned}\Pi_{W,INT}(\mathbf{u}) =& \frac{1}{8}E\left(-(\zeta-1)(v_1-v_2)^2 - \frac{2(v_2-v_3)^2}{\zeta-1} + 2(v_1-v_4)^2\right)\\ &+ \frac{1}{8}E\left((v_3-\zeta v_1+(\zeta-1)v_4)^2\right)\end{aligned} \tag{190}$$

As we consider only prescribed displacements $v_1 = -C$, $v_2 = -C$ where $C > 0$ and no body load or surface load, it holds

$$\Pi_{W,EXT}(\mathbf{u}) = 0 \tag{191}$$

$$\Pi_{W,NIT}(\mathbf{u}) = E(-v_1+v_4)\begin{cases} \frac{1}{2}(-\zeta a - v_3 - v_4) & \zeta a + v_3 \le 0 \wedge v_4 \le 0\\ -\frac{(\zeta a+v_3)^2}{2(\zeta a+v_3-v_4)} & \zeta a + v_3 < 0 \wedge v_4 > 0\\ \frac{v_4^2}{2(\zeta a+v_3-v_4)} & \zeta a + v_3 > 0 \wedge v_4 < 0\\ 0 & \text{Elsewhere}\end{cases} \tag{192}$$

$$\Pi_{W,PEN}(\mathbf{u}) = E\frac{\gamma E}{h}\begin{cases} \frac{(\zeta a+v_3)^3}{3(\zeta a+v_3-v_4)} & \zeta a + v_3 < 0 \wedge v_4 > 0\\ -\frac{v_4^3}{3(\zeta a+v_3-v_4)} & \zeta a + v_3 > 0 \wedge v_4 < 0\\ \frac{1}{3}\left((\zeta a+v_3)^2 + (\zeta a+v_3)v_4 + v_4^2\right) & \zeta a + v_3 \le 0 \wedge v_4 \le 0\\ 0 & \text{Elsewhere}\end{cases} \tag{193}$$

where $\dfrac{\gamma E}{h}$ is the scaled penalty value.

In the following examples, we will consider Young's modulus $E = 30000$ and Poisson ratio $\nu = 0$. Also, we choose $a = 10$, $\zeta = 0.05$ so $\zeta \cdot a = 0.5$, $u_1 = u_2 = u_3 = u_4 = 0$ and $v_1 = v_2 = -1$. To reduce the notation, we will often omit the measurement units, but we presume the length to be in [mm], force in [N] and thus stress in [MPa]. We see that the theoretical solution of this problem is $\mathbf{u} = (v_3, v_4) = (-\zeta\cdot a, 0) = (-0.5, 0)$. In two-dimensional elasticity, we could consider either plane strain or plane stress but this is not important in our case, as $\nu = 0$ and both choices therefore coincide. Concerning scaled penalty value $\gamma E/h$ we will consider $h = a = 10$. Three values of the penalty parameter γ will be considered: $\gamma = 0$, $\gamma = 10$ and $\gamma = 40$.

Equipped with these, we can now plot the energy as the function of v_3 and v_4. Firstly, an analytical solution will be presented for a couple of values of the penalty parameter to give the reader an idea of how the Nitsche-Wriggers method works. Then, some of these results will be compared with the numerical solution from FEniCS, which will help us to understand the convergence difficulties that Nitsche-Wriggers suffers from in FEniCS. Also, we will show the results of the Nitsche-FPR method for some penalty parameter values so that the reader could compare this with Nitsche-Wriggers.

The absolute and relative tolerance of Newton's solver are set as $tol_{abs} = 10^{-7}$ and $tol_{rel} = 10^{-7}$. These tolerances limit the value of residuum (which could be interpreted as an unbalanced force) in every iteration of Newton's method. The maximum number of Newton's method iterations is 100.

4.2.1 Analytical solution for $\gamma = 0$

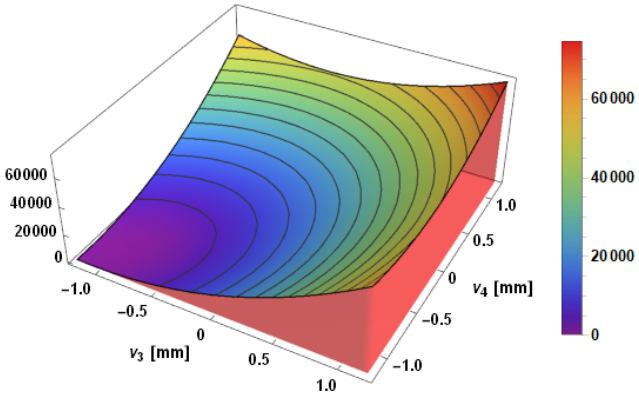


Figure 4: Elastic energy $\Pi_{W,INT}(v_3, v_4)$ [Nmm]

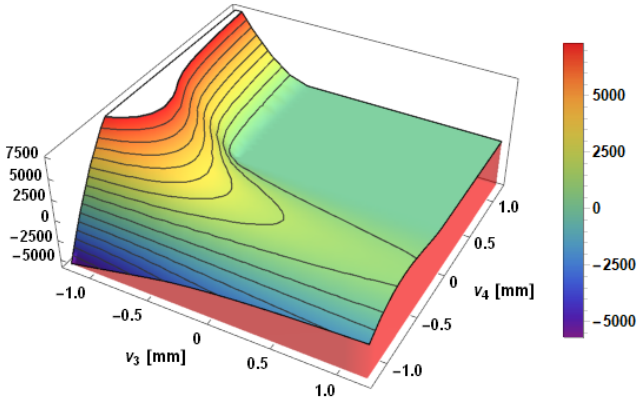


Figure 5: Nitsche energy $\Pi_{W,NIT}(v_3, v_4)$ [Nmm]

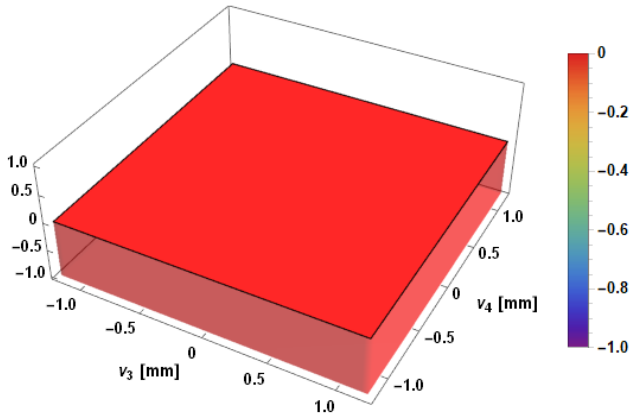


Figure 6: Penalty energy $\Pi_{W,PEN}(v_3, v_4)$ [Nmm]

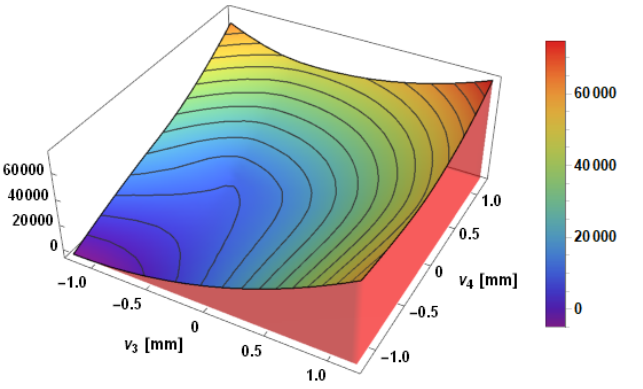


Figure 7: Energy functional $\Pi_W(v_3, v_4) = \Pi_{W,INT}(v_3, v_4) + \Pi_{W,NIT}(v_3, v_4) + \Pi_{W,PEN}(v_3, v_4)$ [Nmm]

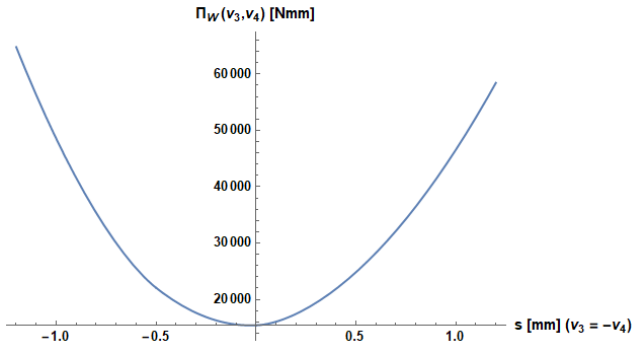


Figure 8: Slice of $\Pi_W(v_3, v_4)$, where $\Pi_W(v_3, v_4) = \Pi_W(v_3, -v_3)$

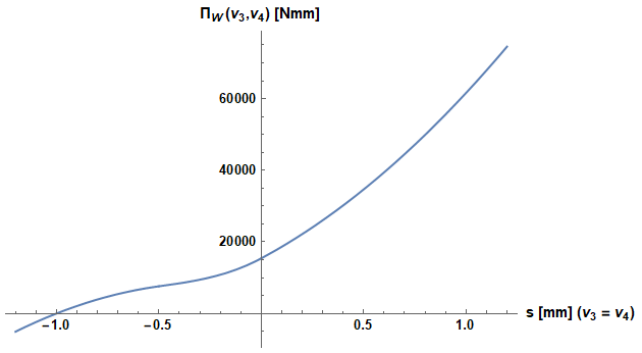


Figure 9: Slice of $\Pi_W(v_3, v_4)$, where $\Pi_W(v_3, v_4) = \Pi_W(v_3, v_3)$

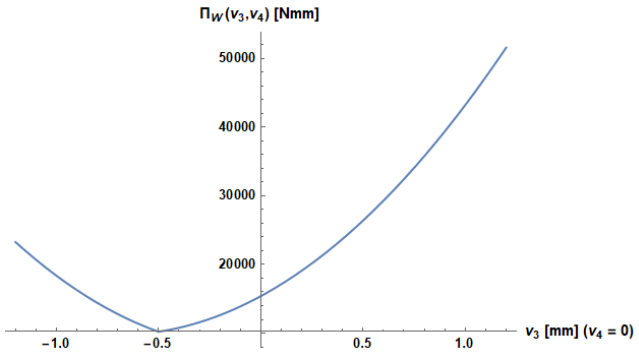


Figure 10: Slice of $\Pi_W(v_3, v_4)$, where $\Pi_W(v_3, v_4) = \Pi_W(v_3, 0)$

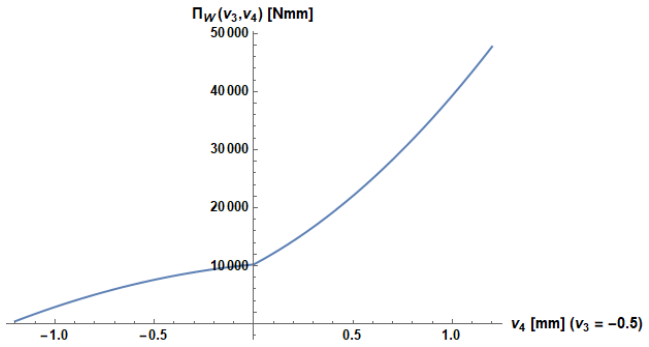


Figure 11: Slice of $\Pi_W(v_3, v_4)$, where $\Pi_W(v_3, v_4) = \Pi_W(-0.5, v_4)$

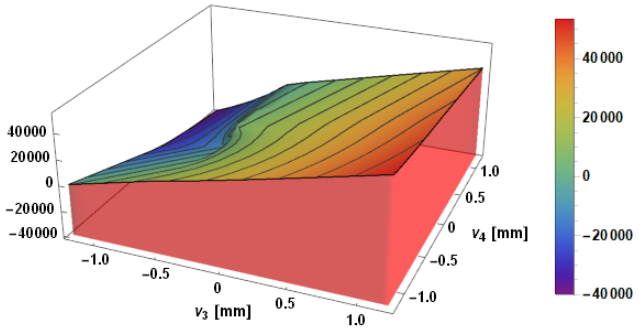


Figure 12: Derivative $\partial_{v_3}\Pi_W(v_3, v_4)$ [N]

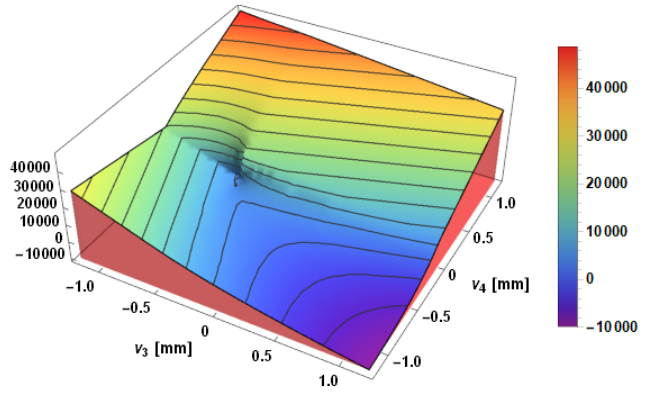


Figure 13: Derivative $\partial_{v_4}\Pi_W(v_3, v_4)$ [N]

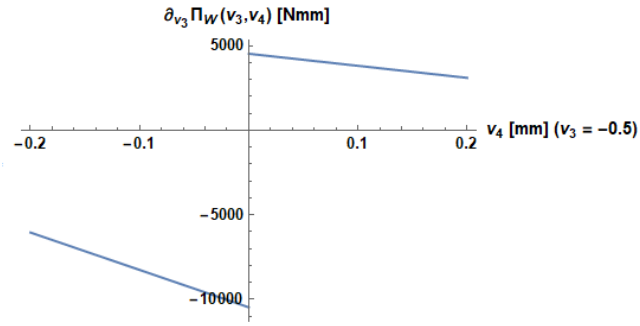


Figure 14: Slice of derivative $\partial_{v_3}\Pi_W(v_3, v_4)$, where $\Pi_W(v_3, v_4) = \Pi_W(-0.5, v_4)$

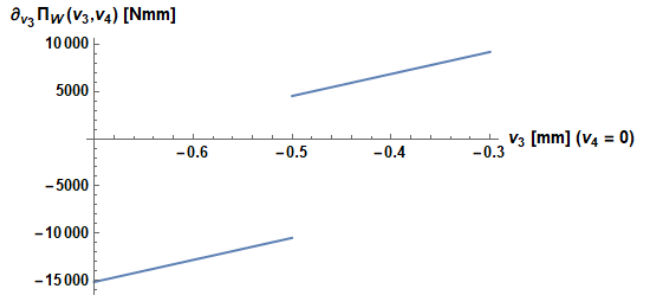


Figure 15: Slice of derivative $\partial_{v_3}\Pi_W(v_3, v_4)$, where $\Pi_W(v_3, v_4) = \Pi_W(v_3, 0)$

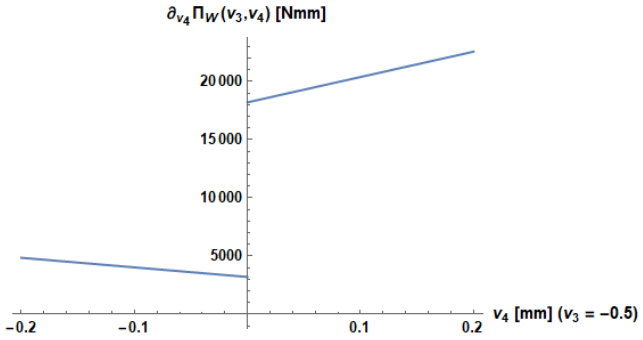


Figure 16: Slice of derivative $\partial_{v_4}\Pi_W(v_3, v_4)$, where $\Pi_W(v_3, v_4) = \Pi_W(-0.5, v_4)$

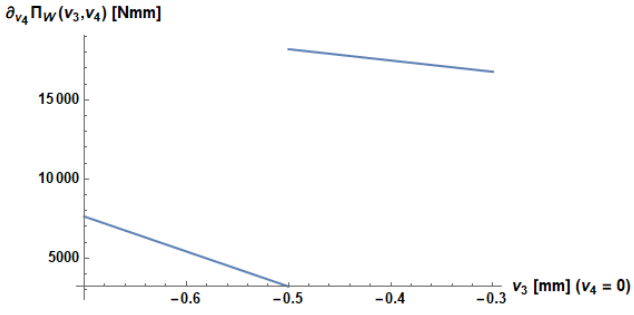


Figure 17: Slice of derivative $\partial_{v_4}\Pi_W(v_3, v_4)$, where $\Pi_W(v_3, v_4) = \Pi_W(v_3, 0)$

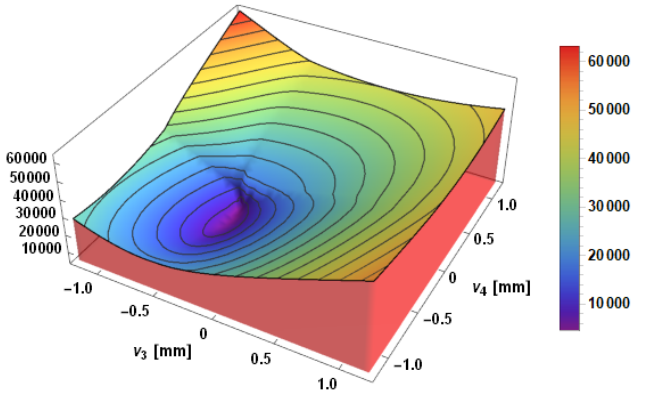


Figure 18: Euclidean norm of energy gradient $\|\boldsymbol{\nabla}\Pi_W(v_3, v_4)\|$ [N]

4.2.2 Analytical solution for $\gamma = 10$

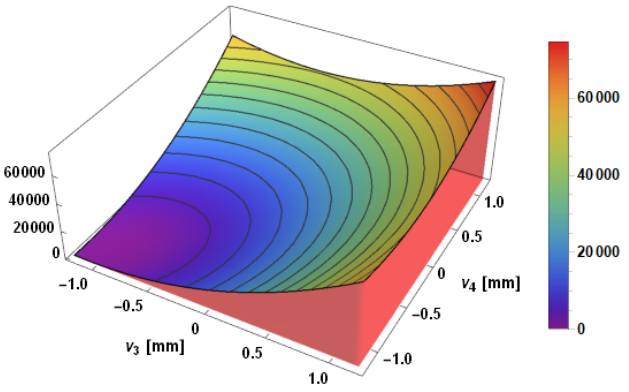


Figure 19: Elastic energy $\Pi_{W,INT}(v_3, v_4)$ [Nmm]

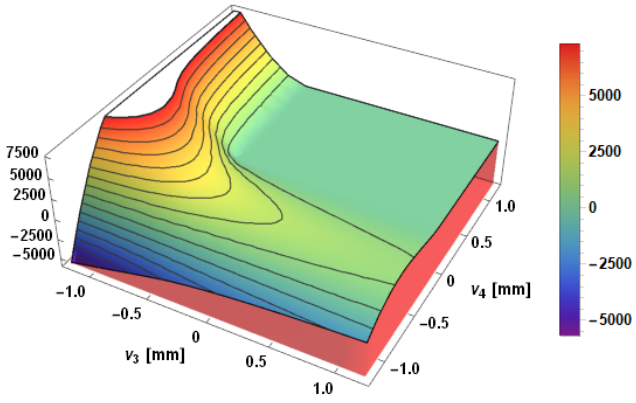


Figure 20: Nitsche energy $\Pi_{W,NIT}(v_3, v_4)$ [Nmm]

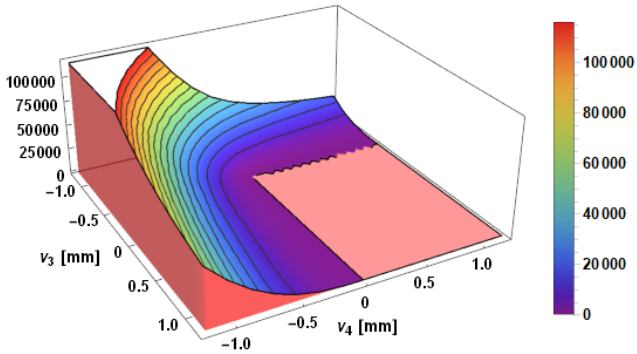


Figure 21: Penalty energy $\Pi_{W,PEN}(v_3, v_4)$ [Nmm]

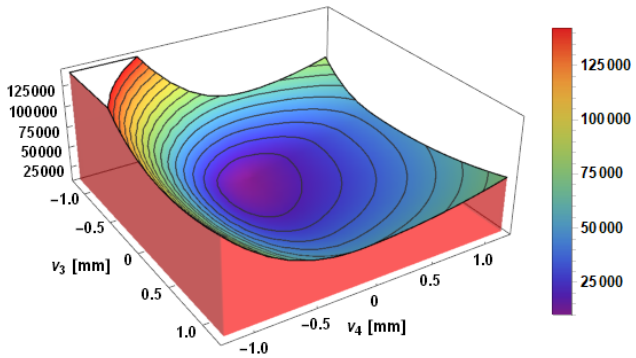


Figure 22: Energy functional $\Pi_W(v_3, v_4) = \Pi_{W,INT}(v_3, v_4) + \Pi_{W,NIT}(v_3, v_4) + \Pi_{W,PEN}(v_3, v_4)$ [Nmm]

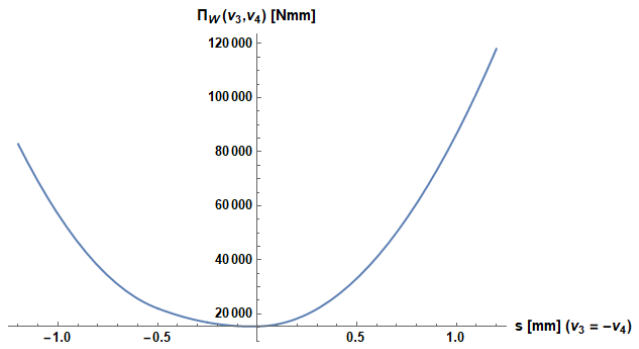


Figure 23: Slice of $\Pi_W(v_3, v_4)$, where $\Pi_W(v_3, v_4) = \Pi_W(v_3, -v_3)$

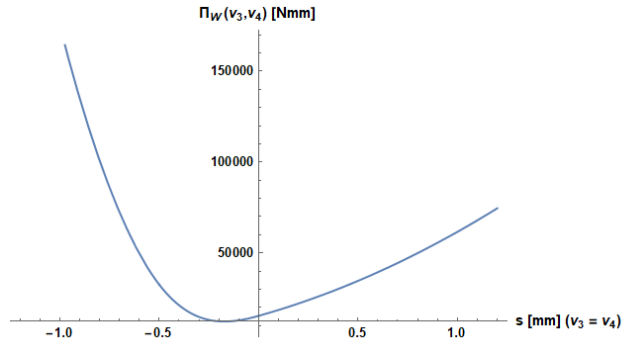


Figure 24: Slice of $\Pi_W(v_3, v_4)$, where $\Pi_W(v_3, v_4) = \Pi_W(v_3, v_3)$

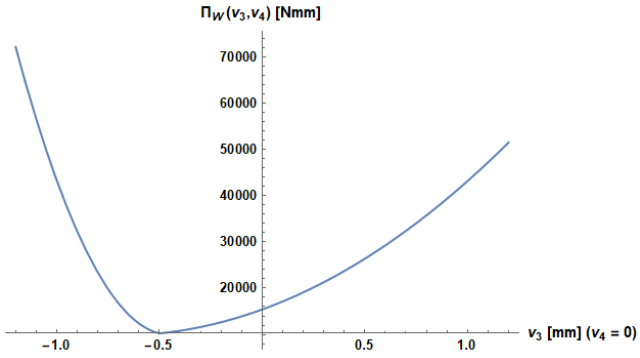


Figure 25: Slice of $\Pi_W(v_3, v_4)$, where $\Pi_W(v_3, v_4) = \Pi_W(v_3, 0)$

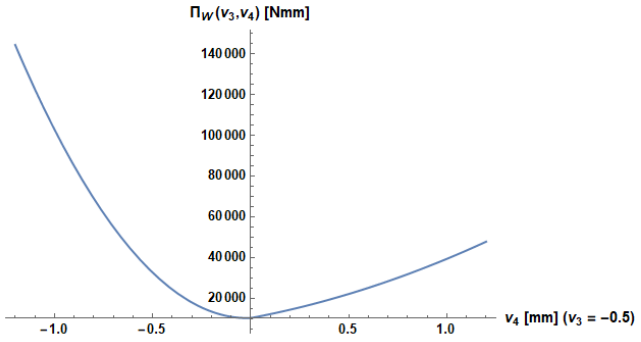


Figure 26: Slice of $\Pi_W(v_3, v_4)$, where $\Pi_W(v_3, v_4) = \Pi_W(-0.5, v_4)$

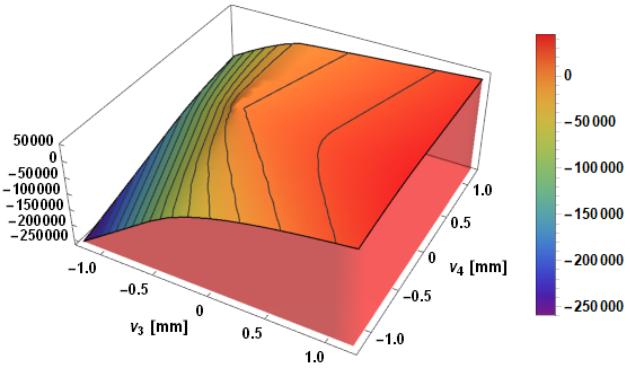


Figure 27: Derivative $\partial_{v_3}\Pi_W(v_3, v_4)$ [N]

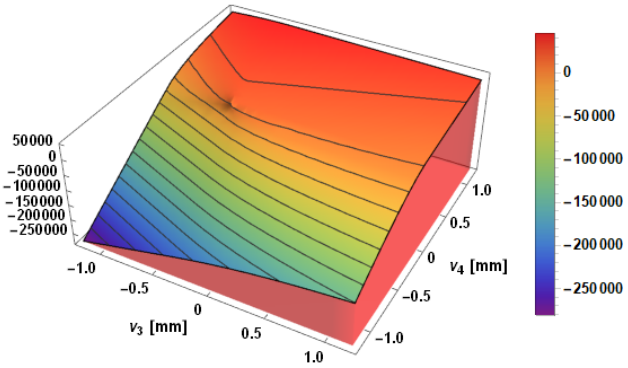


Figure 28: Derivative $\partial_{v_4}\Pi_W(v_3, v_4)$ [N]

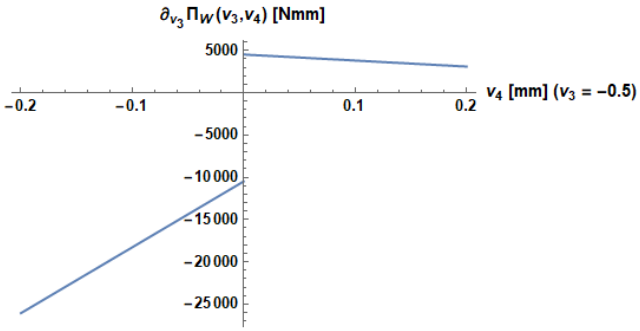


Figure 29: Slice of derivative $\partial_{v_3}\Pi_W(v_3, v_4)$, where $\Pi_W(v_3, v_4) = \Pi_W(-0.5, v_4)$

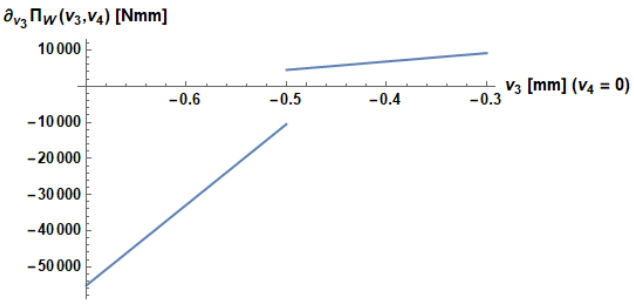


Figure 30: Slice of derivative $\partial_{v_3}\Pi_W(v_3, v_4)$, where $\Pi_W(v_3, v_4) = \Pi_W(v_3, 0)$

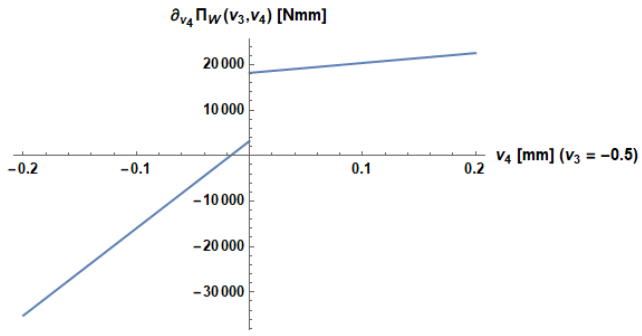


Figure 31: Slice of derivative $\partial_{v_4}\Pi_W(v_3, v_4)$, where $\Pi_W(v_3, v_4) = \Pi_W(-0.5, v_4)$

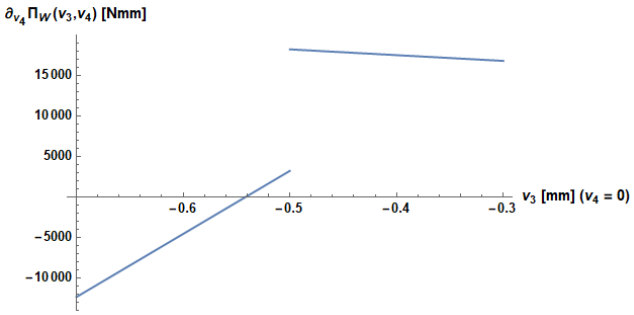


Figure 32: Slice of derivative $\partial_{v_4}\Pi_W(v_3, v_4)$, where $\Pi_W(v_3, v_4) = \Pi_W(v_3, 0)$

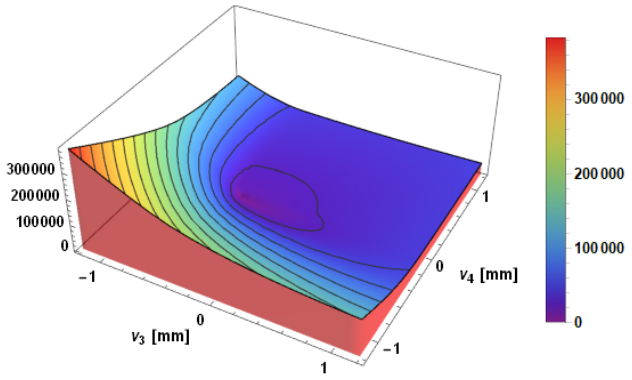


Figure 33: Euclidean norm of energy gradient $\|\boldsymbol{\nabla}\Pi_W(v_3, v_4)\|$ [N]

4.2.3 Analytical solution for $\gamma = 40$

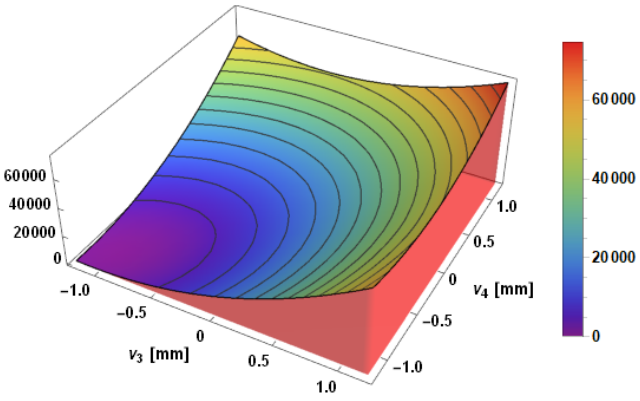


Figure 34: Elastic energy $\Pi_{W,INT}(v_3, v_4)$ [Nmm]

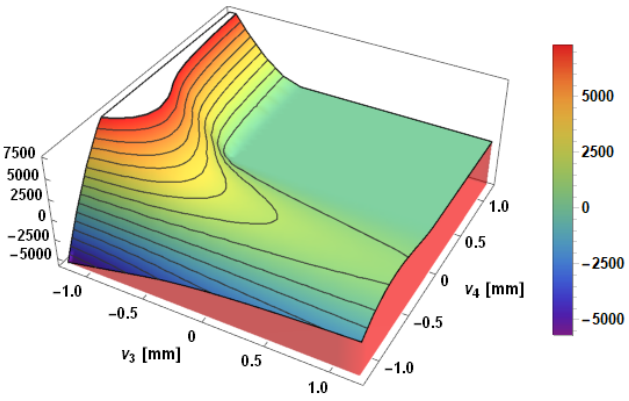


Figure 35: Nitsche energy $\Pi_{W,NIT}(v_3, v_4)$ [Nmm]

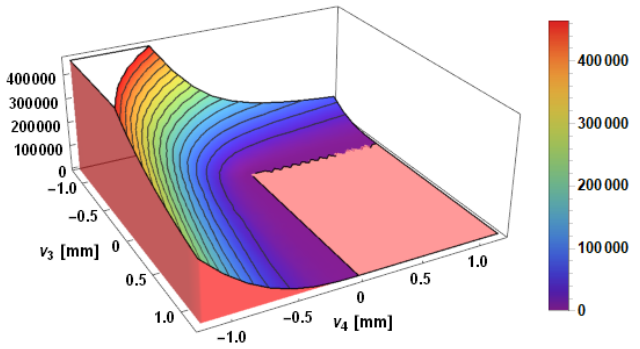


Figure 36: Penalty energy $\Pi_{W,PEN}(v_3, v_4)$ [Nmm]

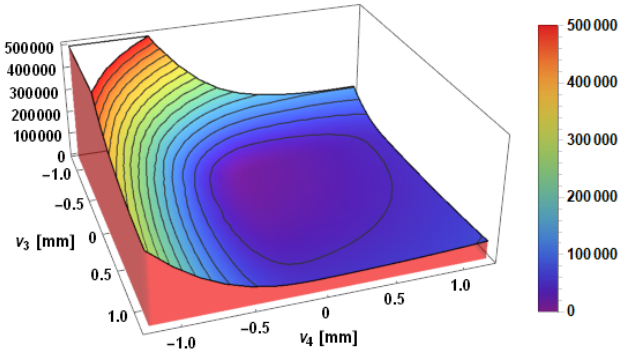


Figure 37: Energy functional $\Pi_W(v_3, v_4) = \Pi_{W,INT}(v_3, v_4) + \Pi_{W,NIT}(v_3, v_4) + \Pi_{W,PEN}(v_3, v_4)$ [Nmm]

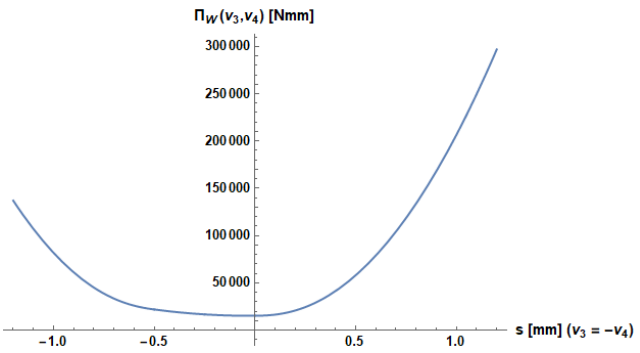


Figure 38: Slice of $\Pi_W(v_3, v_4)$, where $\Pi_W(v_3, v_4) = \Pi_W(v_3, -v_3)$

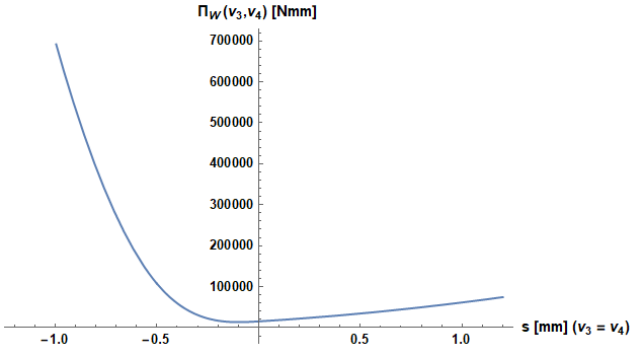


Figure 39: Slice of $\Pi_W(v_3, v_4)$, where $\Pi_W(v_3, v_4) = \Pi_W(v_3, v_3)$

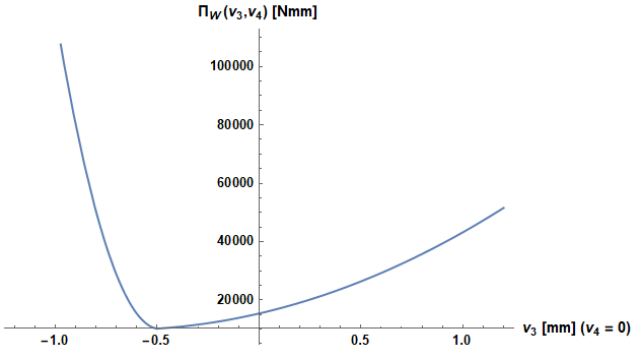


Figure 40: Slice of $\Pi_W(v_3, v_4)$, where $\Pi_W(v_3, v_4) = \Pi_W(v_3, 0)$

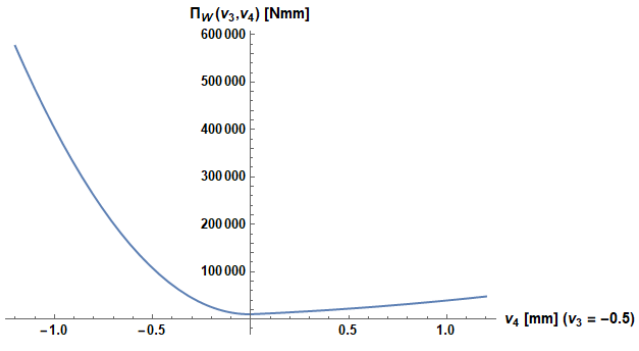


Figure 41: Slice of $\Pi_W(v_3, v_4)$, where $\Pi_W(v_3, v_4) = \Pi_W(-0.5, v_4)$

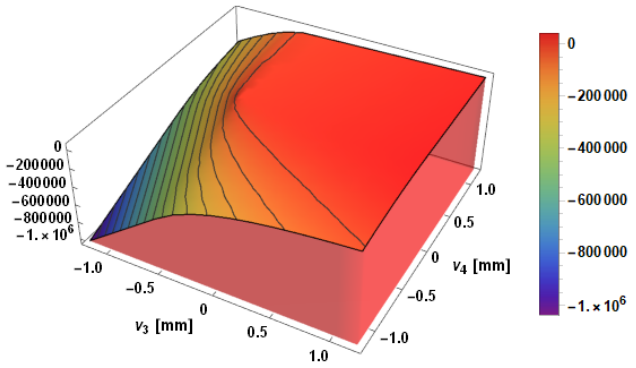


Figure 42: Derivative $\partial_{v_3}\Pi_W(v_3, v_4)$ [N]

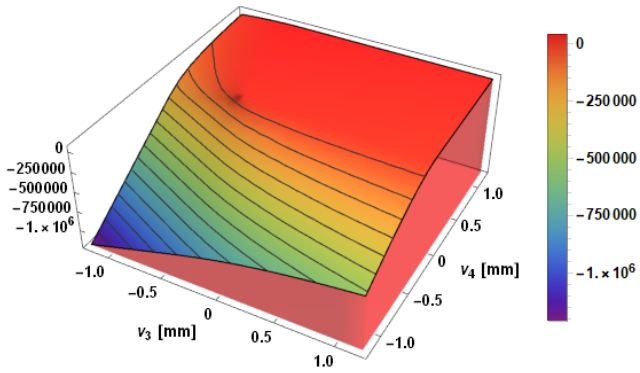


Figure 43: Derivative $\partial_{v_4}\Pi_W(v_3, v_4)$ [N]

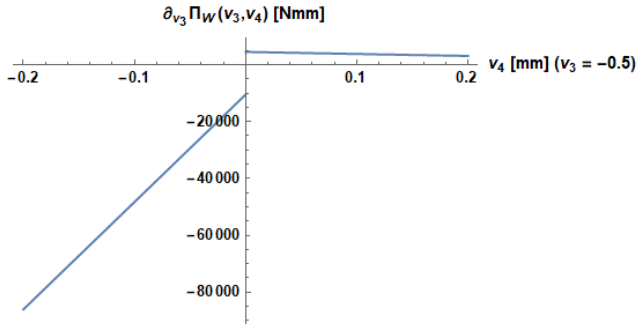


Figure 44: Slice of derivative $\partial_{v_3}\Pi_W(v_3, v_4)$, where $\Pi_W(v_3, v_4) = \Pi_W(-0.5, v_4)$

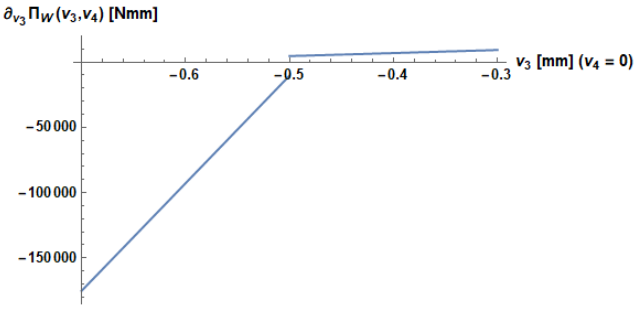


Figure 45: Slice of derivative $\partial_{v_3}\Pi_W(v_3, v_4)$, where $\Pi_W(v_3, v_4) = \Pi_W(v_3, 0)$

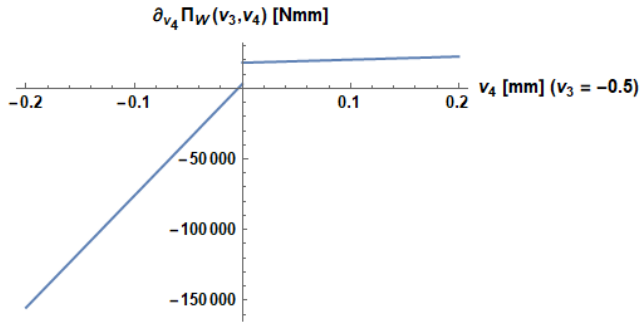


Figure 46: Slice of derivative $\partial_{v_4}\Pi_W(v_3, v_4)$, where $\Pi_W(v_3, v_4) = \Pi_W(-0.5, v_4)$

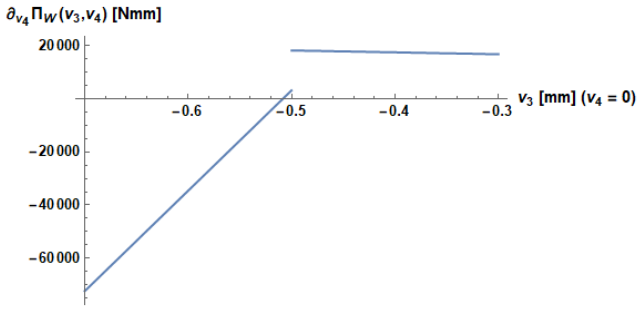


Figure 47: Slice of derivative $\partial_{v_4}\Pi_W(v_3, v_4)$, where $\Pi_W(v_3, v_4) = \Pi_W(v_3, 0)$

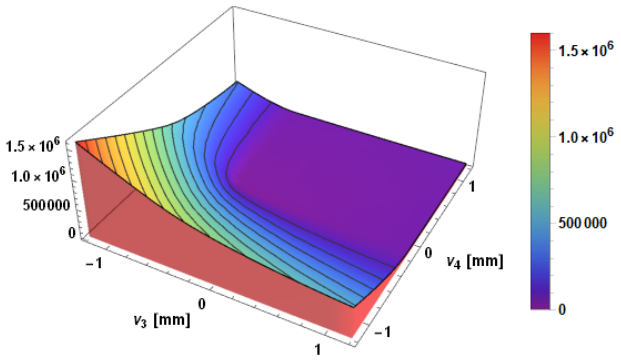


Figure 48: Euclidean norm of energy gradient $\|\nabla\Pi_W(v_3, v_4)\|$ [N]

4.2.4 Comparison of analytical solution and FEniCS solution for $\gamma = 10$

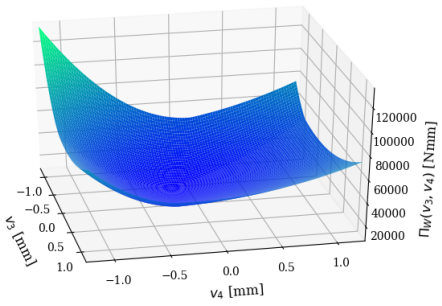


Figure 49: Energy functional $\Pi_W(v_3, v_4)$ [Nmm] — FEniCS

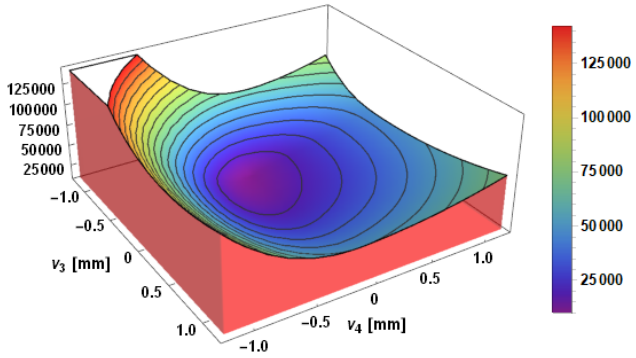


Figure 50: Energy functional $\Pi_W(v_3, v_4)$ [Nmm] — analytical

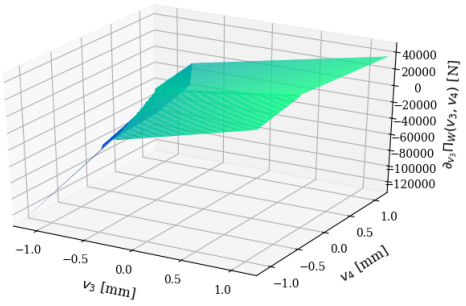


Figure 51: Derivative $\partial_{v_3}\Pi_W(v_3, v_4)$ [N] — FEniCS

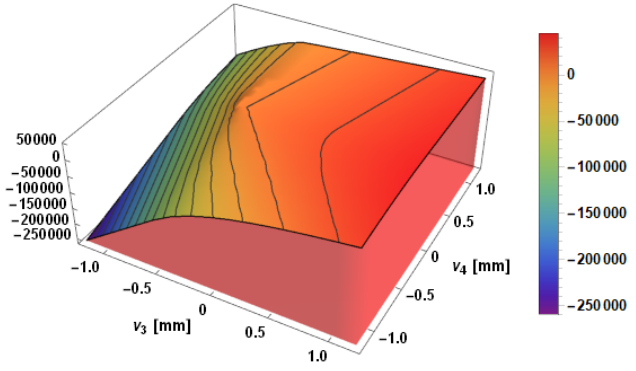


Figure 52: Derivative $\partial_{v_3}\Pi_W(v_3, v_4)$ [N] — analytical

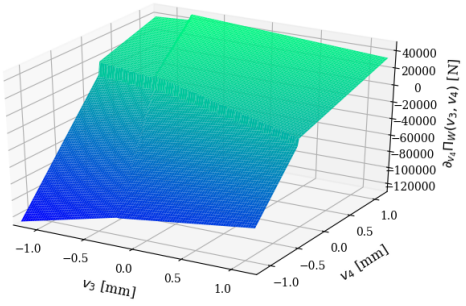


Figure 53: Derivative $\partial_{v_4}\Pi_W(v_3, v_4)$ [N] — FEniCS

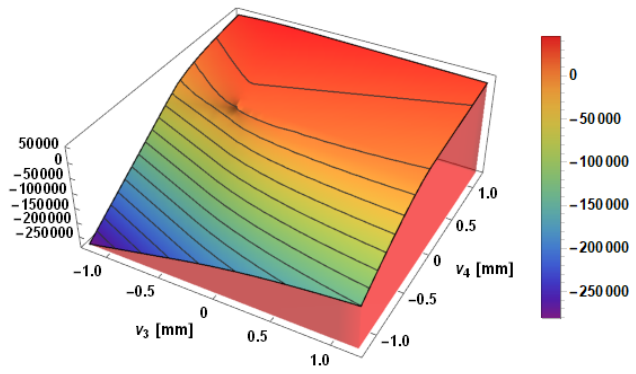


Figure 54: Derivative $\partial_{v_4}\Pi_W(v_3, v_4)$ [N] — analytical

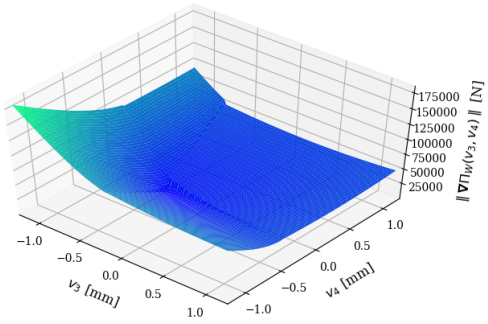


Figure 55: Euclidean norm of energy gradient $\|\nabla\Pi_W(v_3, v_4)\|$ [N] — FEniCS

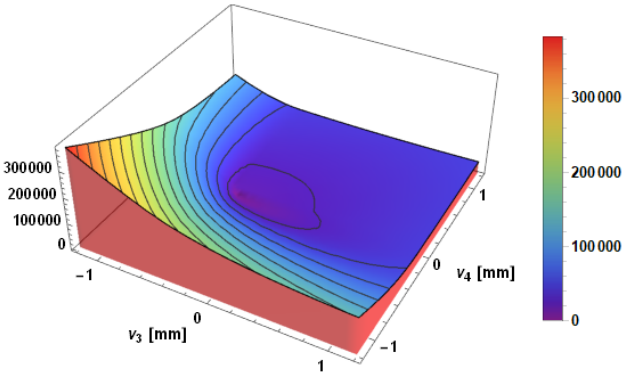


Figure 56: Euclidean norm of energy gradient $\|\nabla\Pi_W(v_3, v_4)\|$ [N] — analytical

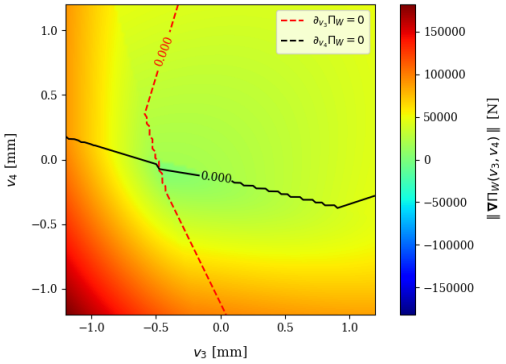


Figure 57: Isolines of zero derivatives, Euclidean norm of energy gradient in the background — FEniCS

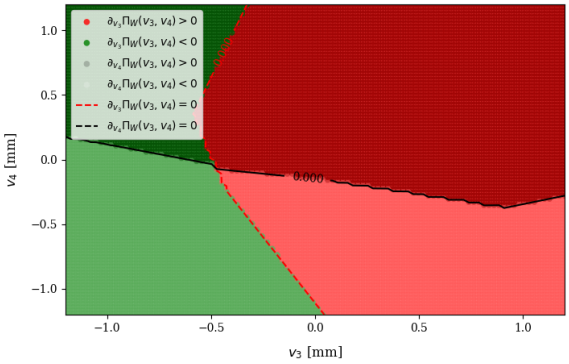


Figure 58: Isolines of zero derivatives, positive and negative regions of both derivatives (distinguished red-green for ∂_{v_3}, darker-lighter ∂_{v_4}) — FEniCS

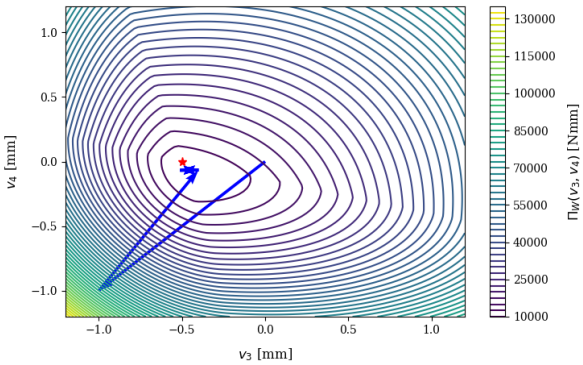


Figure 59: Iterations of Newton's method with contour plot of $\Pi_W(v_3, v_4)$ [N] — FEniCS, the star denotes the theoretical solution $(v_3, v_4) = (-0.5, 0)$

In Figure (59) we can see that after the second iteration, the arrows representing Newton iterations run back and forth. More precisely, the solution $\mathbf{u} = (v_3, v_4)$ oscillates between two values and fails to converge in 100 iterations. Below we list the values of unknown displacement $\mathbf{u} = (v_3, v_4)$ and Jacobian matrix $\mathbf{J} = \begin{pmatrix} \partial^2\Pi_W/\partial v_3^2 & \partial^2\Pi_W/\partial v_3\partial v_4 \\ \partial^2\Pi_W/\partial v_3\partial v_4 & \partial^2\Pi_W/\partial v_4^2 \end{pmatrix}$ in the first seven Newton iterations. We would like to point out that if $v_3 < -0.5$, then node 3 is in contact and so the whole bottom edge is in contact (as node 4 is always in contact). For this reason, the stiffnesses in the corresponding Jacobian matrix are higher. On the contrary, if $v_3 > -0.5$, then node 3 is not in contact and the bottom edge is partially in contact and partially not in contact. The lengths of the active and the inactive part of the bottom edge depend on the displacements v_3 and v_4 and the positions of the integration points on the edge. Therefore, if $v_3 > -0.5$, then the stiffnesses in the corresponding Jacobian matrix are smaller than if $v_3 < -0.5$.

- Iteration 1:

$$\mathbf{u} = (v_3, v_4) = (0.00000, 0.00000) \tag{194}$$

$$\mathbf{J} = \begin{pmatrix} 23289.5 & -7125 \\ -7125 & 21768.8 \end{pmatrix} \tag{195}$$

- Iteration 2:

$$(v_3, v_4) = (-1.00000, -1.00000) \tag{196}$$

$$\mathbf{J} = \begin{pmatrix} 123413 & 27918.9 \\ 27918.9 & 91857.4 \end{pmatrix} \tag{197}$$

- Iteration 3:

$$\mathbf{u} = (v_3, v_4) = (-0.40022, -0.06503) \tag{198}$$

$$\mathbf{J} = \begin{pmatrix} 29996.3 & 14732 \\ 14732 & 91497.8 \end{pmatrix} \tag{199}$$

- Iteration 4:

$$\mathbf{u} = (v_3, v_4) = (-0.51165, -0.06765) \tag{200}$$

$$\mathbf{J} = \begin{pmatrix} 123413 & 27918.9 \\ 27918.9 & 91857.4 \end{pmatrix} \tag{201}$$

- Iteration 5:

$$\mathbf{u} = (v_3, v_4) = (-0.40022, -0.06503) \tag{202}$$

$$\mathbf{J} = \begin{pmatrix} 29996.3 & 14732 \\ 14732 & 91497.8 \end{pmatrix} \tag{203}$$

- Iteration 6:

$$\mathbf{u} = (v_3, v_4) = (-0.51165, -0.06765) \tag{204}$$

$$\mathbf{J} = \begin{pmatrix} 123413 & 27918.9 \\ 27918.9 & 91857.4 \end{pmatrix} \tag{205}$$

- Iteration 7:

$$\mathbf{u} = (v_3, v_4) = (-0.40022, -0.06503) \tag{206}$$

$$\mathbf{J} = \begin{pmatrix} 29996.3 & 14732 \\ 14732 & 91497.8 \end{pmatrix} \tag{207}$$

Although this is just one example with specifically chosen numerical parameters, we encountered this behaviour of Nitsche-Wriggers in most of the numerical simulations we carried out. Very rarely, the method converged in the maximum number of iterations, and one such example will be presented in the following chapter. Naturally, the question arises, what is the reason for this peculiar behaviour. To answer this question, we studied this example with two DOFs. The fact that we have just two DOFs allows us to plot the discrete energy functional and its derivatives as the functions of both unknown displacements.

In Figure 59 and from the above stated Newton iterations (194)–(207) we can see that for some reason, the Nitsche-Wriggers method fails to converge in the vicinity of the theoretical solution. The reason is clear if we look at the Figures 29, 30, 31 and 32 depicting the slices of $\partial_{v_3}\Pi_W(\mathbf{u})$ and $\partial_{v_4}\Pi_W(\mathbf{u})$. In these figures we can clearly see that the derivatives are discontinuous with the discontinuity located near the theoretical solution $(v_3, v_4) = (-0.5, 0)$. This means that our discrete Nitsche-Wriggers energy functional is not C^1-differentiable with respect to v_3 and v_4 which is one of the conditions required so that the convergence of Newton's method could be guaranteed. For this reason, Newton's method fails to converge.

We would like to remark that in addition to the discontinuity in the theoretical solution, we can observe some additional lines of discontinuities in the figures 51 and 53. These figures depict the derivatives $\partial_{v_3}\Pi_W(\mathbf{u})$ and $\partial_{v_4}\Pi_W(\mathbf{u})$ which were evaluated numerically in FEniCS (contrarily to the analytical computations, in which case the integration was performed analytically). The additional lines of discontinuities are caused by the numerical integration. In this case, we use the three-point Gaussian quadrature to evaluate the integrals over the potential contact zones. This means that there are two quadrature points located on the inclined bottom edge, which is in contact. For certain linear combinations of displacements v_3, v_4 it could happen that the gap of one of the quadrature points is zero. Its contribution is therefore neglected, which causes the line of discontinuity in the derivatives values. The linear combinations of v_3, v_4 are geometrically describing a line. Because we have two integration points, we can observe two lines of discontinuities in the derivatives. This is particularly conspicuous in Figure 53.

4.2.5 Comparison of analytical solution and FEniCS solution for $\gamma = 40$

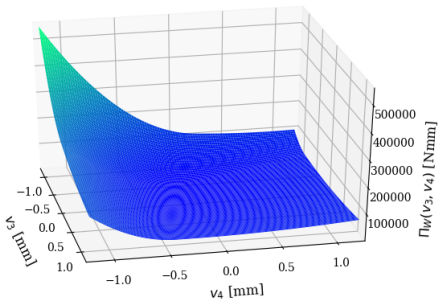


Figure 60: Energy functional $\Pi_W(v_3, v_4)$ [Nmm] — FEniCS

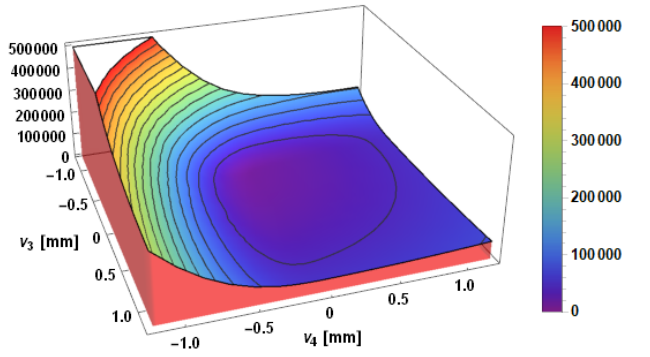


Figure 61: Energy functional $\Pi_W(v_3, v_4)$ [Nmm] — analytical

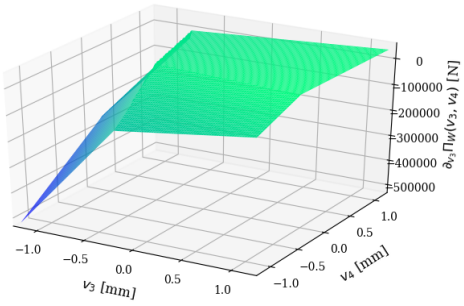


Figure 62: Derivative $\partial_{v_3}\Pi_W(v_3, v_4)$ [N] — FEniCS

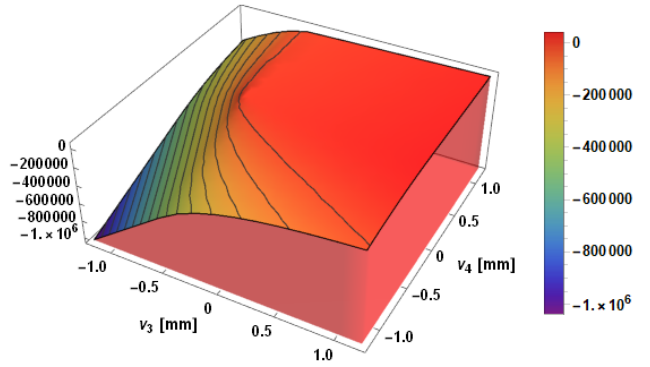


Figure 63: Derivative $\partial_{v_3}\Pi_W(v_3, v_4)$ [N] — analytical

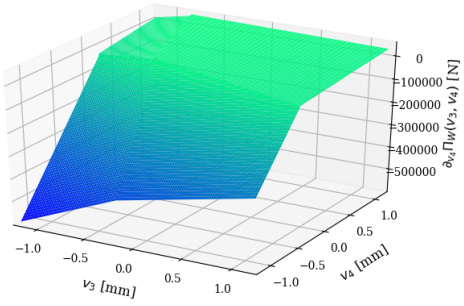


Figure 64: Derivative $\partial_{v_4}\Pi_W(v_3, v_4)$ [N] — FEniCS

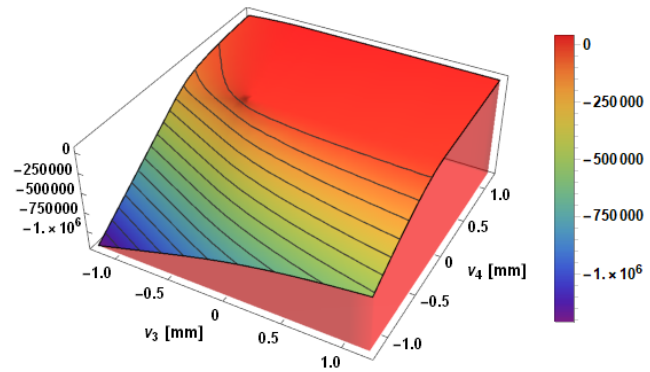


Figure 65: Derivative $\partial_{v_4}\Pi_W(v_3, v_4)$ [N] — analytical

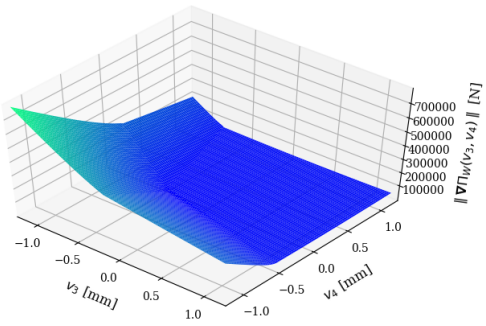


Figure 66: Euclidean norm of energy gradient $\|\nabla\Pi_W(v_3, v_4)\|$ [N] — FEniCS

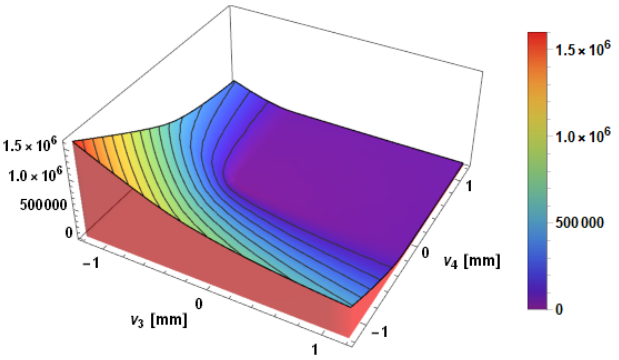


Figure 67: Euclidean norm of energy gradient $\|\nabla\Pi_W(v_3, v_4)\|$ [N] — analytical

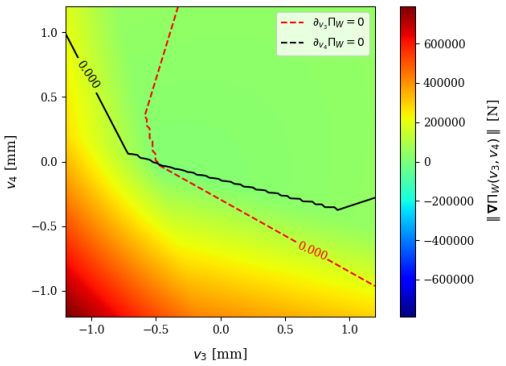


Figure 68: Isolines of zero derivatives, Euclidean norm of energy gradient in the background — FEniCS

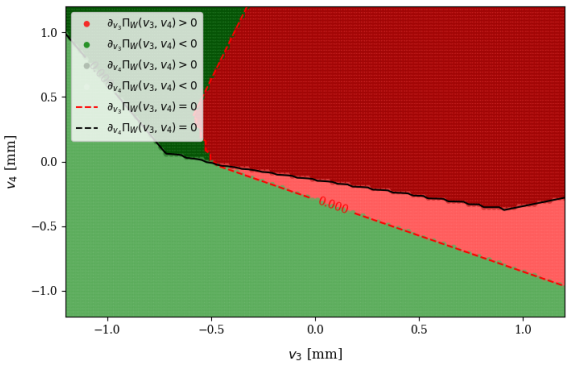


Figure 69: Isolines of zero derivatives, positive and negative regions of both derivatives (distinguished red-green for ∂_{v_3}, darker-lighter ∂_{v_4}) — FEniCS

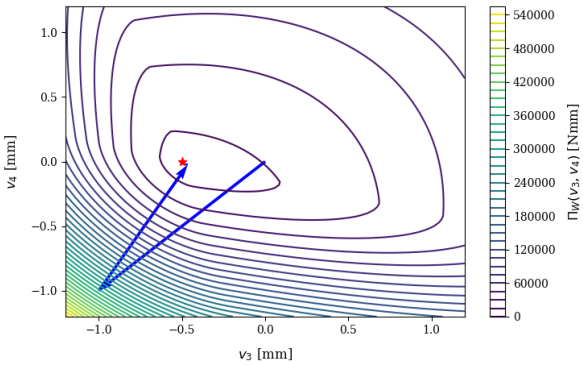


Figure 70: Iterations of Newton's method with contour plot of $\Pi_W(v_3, v_4)$ [N] — FEniCS, the star denotes the theoretical solution $(v_3, v_4) = (-0.5, 0)$

As was mentioned in the previous section, we encountered severe problems with the convergence of Nitsche-Wriggers method in FEniCS. These are caused by the fact that the Nitsche-Wriggers energy functional, if discretised, is not C^1-differentiable function with respect to the unknown DOFs, as we demonstrated by means of an example in the previous section 4.2.4. However, in some rare cases the method converged. One such example is the problem of a quadrilateral pressed against a rigid plane, which we analyse in this chapter with penalty parameter value $\gamma = 40$. This could also be seen from the graphical representation of Newton's method iterations 70. We remind that the very same problem only with penalty parameter value $\gamma = 10$ did not converge in FEniCS, as is described in the previous section 4.2.4. It may seem that the non-convergence of the previous example was caused only by the insufficient value of penalty parameter (which sometimes really could be the reason). Unfortunately, this is not the case here as was explained in the previous section 4.2.4. Below we present how displacement $\mathbf{u} = (v_3, v_4)$ and Jacobian matrix $\mathbf{J} = \begin{pmatrix} \partial^2\Pi_W/\partial v_3^2 & \partial^2\Pi_W/\partial v_3\partial v_4 \\ \partial^2\Pi_W/\partial v_3\partial v_4 & \partial^2\Pi_W/\partial v_4^2 \end{pmatrix}$ change during the Newton iterations:

- Iteration 1:

$$\mathbf{u} = (v_3, v_4) = (0.00000, 0.00000) \tag{208}$$

$$\mathbf{J} = \begin{pmatrix} 23289.5 & -7125 \\ -7125 & 21768.8 \end{pmatrix} \tag{209}$$

- Iteration 2:

$$\mathbf{u} = (v_3, v_4) = (-1.00000, -1.00000) \tag{210}$$

$$\mathbf{J} = \begin{pmatrix} 423783 & 178106 \\ 178106 & 392237 \end{pmatrix} \tag{211}$$

- Iteration 3:

$$\mathbf{u} = (v_3, v_4) = (-0.46516, -0.02395) \tag{212}$$

$$\mathbf{J} = \begin{pmatrix} 50116.6 & 89824.5 \\ 89824.5 & 371755 \end{pmatrix} \tag{213}$$

- Iteration 4:

$$\mathbf{u} = (v_3, v_4) = (-0.49328, -0.02395) \tag{214}$$

$$\mathbf{J} = \begin{pmatrix} 50116.6 & 89824.5 \\ 89824.5 & 371755 \end{pmatrix} \tag{215}$$

With the fourth Newton iteration, the discrete problem converges and the obtained solution is relatively close to the theoretical one, which is $\mathbf{u} = (v_3, v_4) = (-0.5, 0)$.

4.2.6 FEniCS solution for $\gamma = 10$ — Nitsche-FPR

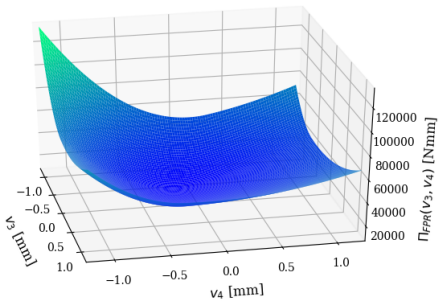


Figure 71: Energy functional $\Pi_{FPR}(v_3, v_4)$ [Nmm] - FEniCS

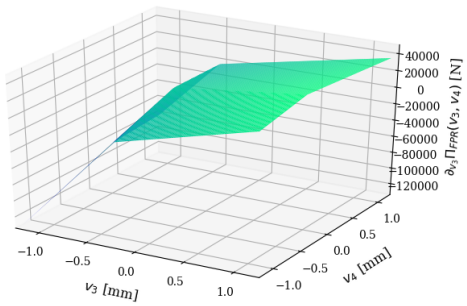


Figure 72: Derivative $\partial_{v_3}\Pi_{FPR}(v_3, v_4)$ [N] — FEniCS

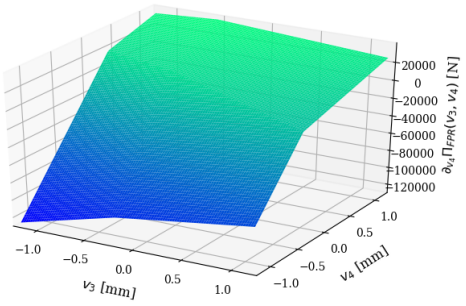


Figure 73: Derivative $\partial_{v_4}\Pi_{FPR}(v_3, v_4)$ [N] — FEniCS

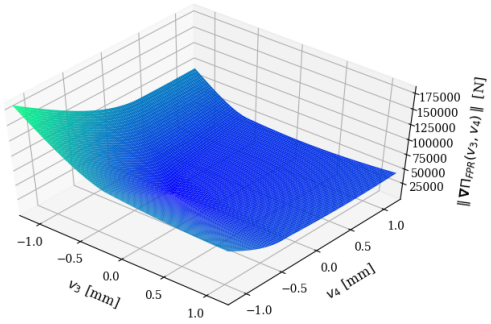


Figure 74: Euclidean norm of energy gradient — $\|\nabla\Pi_{FPR}(v_3, v_4)\|$ [N] — FEniCS

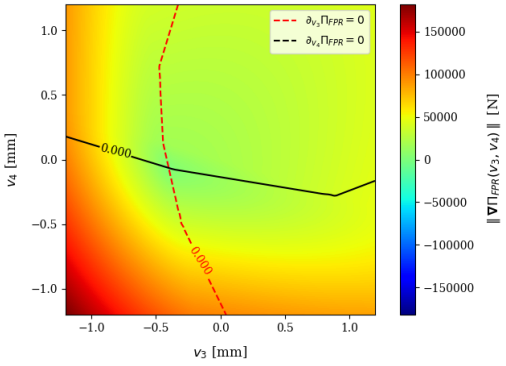


Figure 75: Isolines of zero derivatives, Euclidean norm of energy gradient in the background — FEniCS

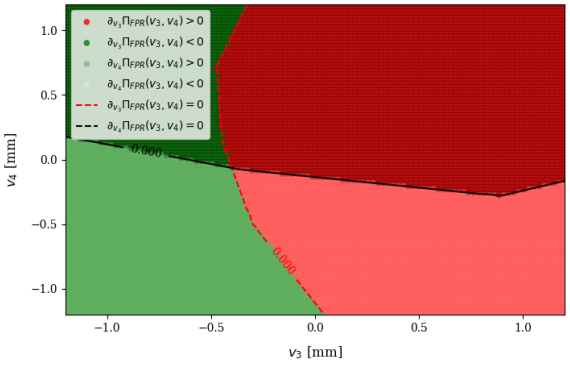


Figure 76: Isolines of zero derivatives, positive and negative regions of both derivatives (distinguished red-green for ∂_{v_3}, darker-lighter ∂_{v_4}) — FEniCS

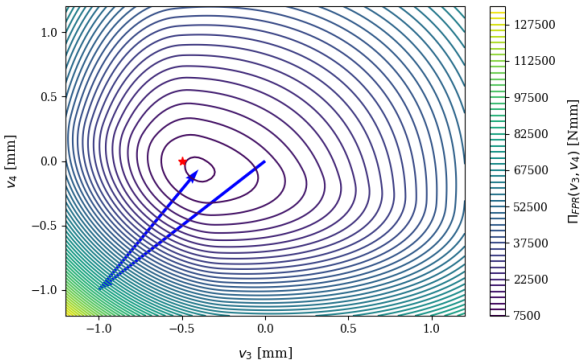


Figure 77: Iterations of Newton's method with contour plot of $\Pi_{FPR}(v_3, v_4)$ [N] — FEniCS, the star denotes the theoretical solution $(v_3, v_4) = (-0.5, 0)$

In this case, Newton's method converged in three iterations, which is also visible from Figure 77. The solution is $\mathbf{u} = (v_3, v_4) = (-0.40022, -0.06503)$. We would like to point out that while Nitsche-Wriggers is not able to converge in 100 iterations (as we described in the previous sections), Nitsche-FPR converges perfectly for the same penalty value. Overall, as will be apparent from the results of the other numerical simulations, with Nitsche-FPR method we nearly did not run into non-convergence problems which Nitsche-Wriggers suffers from. The reason for this is that (contrarily to the Nitsche-Wriggers method) it could be shown that the function obtained as the left-hand side of the discretised weak form of the Nitsche-FPR method is continuous with respect to the unknown DOFs. For the symmetric variant ($\theta = -1$) this means that the discretised energy functional is a C^1-differentiable function. This can be seen even intuitively, because the weak form of Nitsche-FPR does not include the term with the Heaviside function (as is the case of Nitsche-Wriggers), which is responsible for the discontinuity of the discretised weak form of the Nitsche-Wriggers method.

4.2.7 FEniCS solution for $\gamma = 40$ — Nitsche-FPR

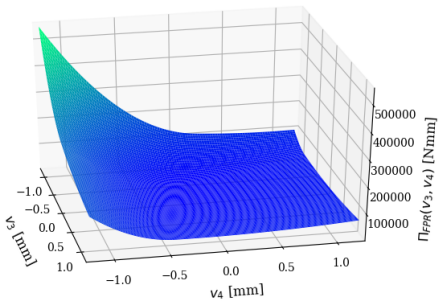


Figure 78: Energy functional $\Pi_{FPR}(v_3, v_4)$ [Nmm] — FEniCS

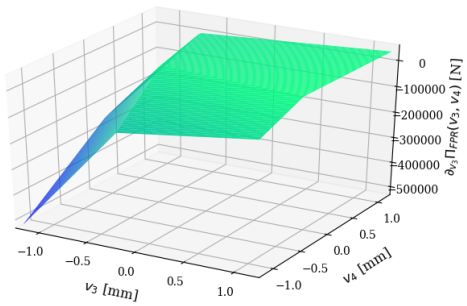


Figure 79: Derivative $\partial_{v_3}\Pi_{FPR}(v_3, v_4)$ [N] — FEniCS

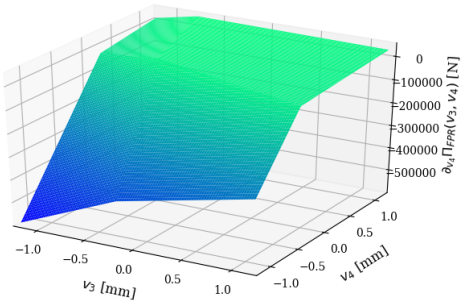


Figure 80: Derivative $\partial_{v_4}\Pi_{FPR}(v_3, v_4)$ [N] — FEniCS

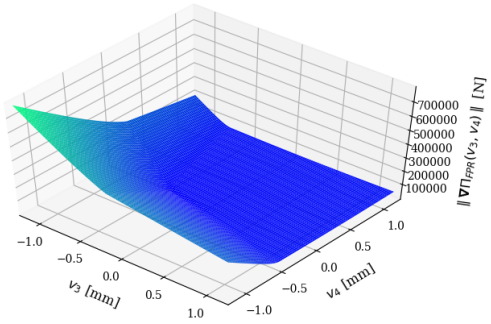


Figure 81: Euclidean norm of energy gradient — $\|\nabla\Pi_{FPR}(v_3, v_4)\|$ [N] — FEniCS

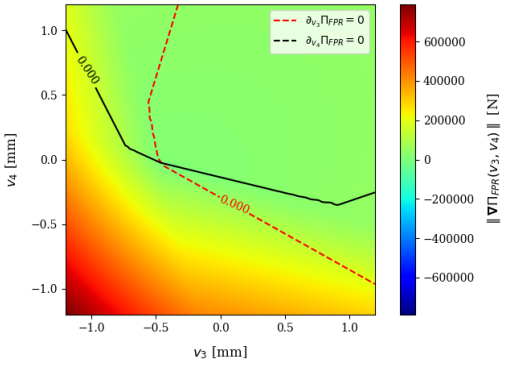


Figure 82: Isolines of zero derivatives, Euclidean norm of energy gradient in the background — FEniCS

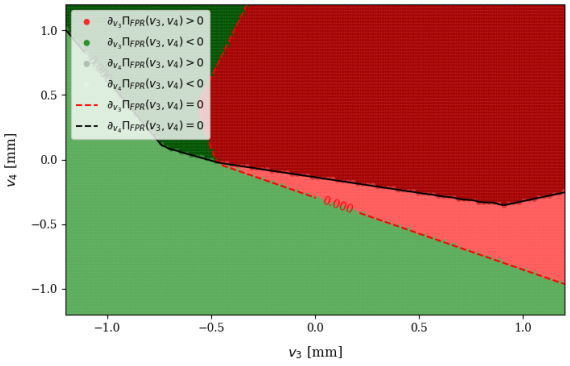


Figure 83: Isolines of zero derivatives, positive and negative regions of both derivatives (distinguished red-green for ∂_{v_3}, darker-lighter ∂_{v_4}) — FEniCS

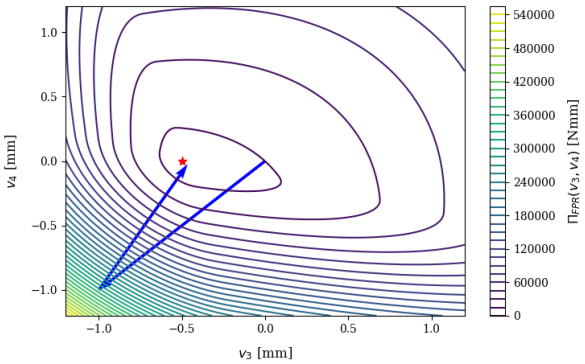


Figure 84: Iterations of Newton's method with contour plot of $\Pi_{FPR}(v_3, v_4)$ [N] — FEniCS, the star denotes the theoretical solution $(v_3, v_4) = (-0.5, 0)$

As we see in Figure 84, for penalty parameter $\gamma = 40$, Newton's method again converges in three iterations, same as for $\gamma = 10$. The obtained solution is $\mathbf{u} = (v_3, v_4) = (-0.46516, -0.02395)$. Because the penalty value is higher, the error of the method is also reduced and the theoretical solution is closer to the numerical one — compare Figures 77 and 84 and the listed values of converged displacements.

4.3 Two-dimensional examples

In this section, we will present several two-dimensional numerical examples. As our objective is to evaluate the accuracy and efficiency of all investigated methods for enforcing the contact boundary conditions, we will measure the relative error in displacements $\|\mathbf{u}\|_{rel}$, the relative error in predicted contact tractions $\|\sigma_n(\mathbf{u})\|_{rel}$ representing the accuracy of the method and the runtime of the solver representing the computational efficiency. These indicators are

- the relative error in displacements $\|\mathbf{u}\|_{rel}$ on the active contact zone $\overline{\Gamma}^{C,S}$ of the SNES solution in L^2 norm

$$\|\mathbf{u}\|_{rel} = \frac{\|\mathbf{u}_{SNES} - \mathbf{u}\|_{L^2}}{\|\mathbf{u}_{SNES}\|_{L^2}} \tag{216}$$

  where

$$\|\mathbf{u}_{SNES} - \mathbf{u}\|_{L^2} = \sqrt{\int_{\overline{\Gamma}^{C,S}} (\mathbf{u}_{SNES} - \mathbf{u})^2 \, \mathrm{d}A} \tag{217}$$

$$\|\mathbf{u}_{SNES}\|_{L^2} = \sqrt{\int_{\overline{\Gamma}^{C,S}} \mathbf{u}_{SNES}^2 \, \mathrm{d}A} \tag{218}$$

- the relative error in predicted normal contact tractions σ_n on the active contact zone $\overline{\Gamma}^{C,S}$ of the SNES solution in L^2 norm

$$\|\sigma_n(\mathbf{u})\|_{rel} = \frac{\|\sigma_n(\mathbf{u}_{SNES}) - \sigma_n(\mathbf{u})\|_{L^2}}{\|\sigma_n(\mathbf{u}_{SNES})\|_{L^2}} \tag{219}$$

  where

$$\|\sigma_n(\mathbf{u}_{SNES}) - \sigma_n(\mathbf{u})\|_{L^2} = \sqrt{\int_{\overline{\Gamma}^{C,S}} (\sigma_n(\mathbf{u}_{SNES}) - \sigma_n(\mathbf{u}))^2 \, \mathrm{d}A} \tag{220}$$

$$\|\sigma_n(\mathbf{u}_{SNES})\|_{L^2} = \sqrt{\int_{\overline{\Gamma}^{C,S}} (\sigma_n(\mathbf{u}_{SNES}))^2 \, \mathrm{d}A} \tag{221}$$

- the solver runtime in $[s]$

In the above error definitions, index $SNES$ denotes the solution obtained using the non-linear solver SNESVINEWTONRSLS from PETSc library. As we can find in PETSc manual [50], this is the reduced space active set solver for variational inequalities based on semi-smooth Newton's method. The details on the theory of this method can be found in [9]. This method provides a very precise solution as the non-penetration geometrical constraints are fulfilled nearly exactly, even on the discrete level. However, this method is quite costly for large-scale problems. For its precision, it was chosen as a reference method in the aforementioned comparisons. We would like to point out that while other considered methods require some value of penalty parameter, this method does not require any.

The absolute and relative tolerances of Newton's solver are set as $tol_{abs} = 10^{-9}$ and $tol_{rel} = 10^{-9}$. These tolerances limit the value of residuum (which could be interpreted as an unbalanced force) in every Newton iteration. The maximum number of Newton iterations is set to 100.

We will consider triangular elements and linear approximation functions for both displacements and Lagrange multipliers. It is also necessary to set the characteristic element size h needed for penalty value $E\gamma/h$. As we previously mentioned, h represents a 'characteristic size' of the mesh elements and is supposed to be constant over the mesh. For our 2D problems we define it as an arithmetic average of the largest and the smallest diameter of the circle circumscribed to the elements.

Two examples will be investigated — a skewed quadrilateral block and a 'cross-country ski', which will be described in more detail in the relevant sections. We would like to remind the reader that the quadrilateral example is similar to the one discussed in the previous chapter on the comparison of the analytical and the numerical solution. For both examples, the displacement $u_y = -\overline{u}$ is prescribed on a certain surface sub-domain. In the case of the skewed quadrilateral block, the prescribed displacement is chosen as $\overline{u} = 2$ mm and in the case of 'cross-country ski' as $\overline{u} = 20$ mm.

4.3.1 Square block meshed with 2 triangular elements

In this example, we will consider the square with sides of length $a = 10$ mm meshed with two triangular elements. There is an initial gap $g_0 = 1$ mm and prescribed downward displacement $\overline{u} = 2$ mm at the top side of the square. We consider Young's modulus $E = 15000$ MPa and Poisson ratio $\nu = 0.0$. The choice of vanishing Poisson ratio is motivated by the fact that no lateral displacements (in x direction) will occur and so we prescribe them directly as zeros to reduce the number of degrees of freedom. This example is evaluated for 20 values of penalty parameter defined by the geometric progression

$$\gamma = 10^4 \left(\frac{1}{2}\right)^{i-1}, \quad i = 1, 2, ..., 20 \tag{222}$$

This kind of series was selected so there would be enough points reflecting small penalty values and supposed log-log scale character of the plot. The largest value of γ is 10000 and the smallest 0.019. In this case, all methods converged for all the penalty parameter values, even Nitsche-Wriggers did.

We see that except for the penalty method, the error in displacements and tractions of other methods is similar, generally very small and probably results from the rounding error. The error produced by the penalty method is significantly larger and decreases with the increasing penalty value. The reason why we obtained so precise results with all the investigated methods (except the penalty method) is that the displacement is linear (with respect to y and independent of x) in this example. because we use linear approximation functions, the linear displacement is represented precisely even with two triangular elements. We see that the solver runtime is very short as there are only two DOFs (in terms of displacement). It is similar for the penalty method and all the Nitsche methods as all these methods work with the same number of DOFs. The solver runtime of the augmented Lagrangian method is significantly longer (though it is still very short), as there are additional DOFs (Lagrange multipliers) representing the normal contact traction values. The runtime of SNES is longer than for the Nitsche methods in this case. It should be pointed out that the points representing the SNES method are not present in the log-log graphs of errors, as SNES is considered to be the reference solution and its error is in this simple problem equal to zero.

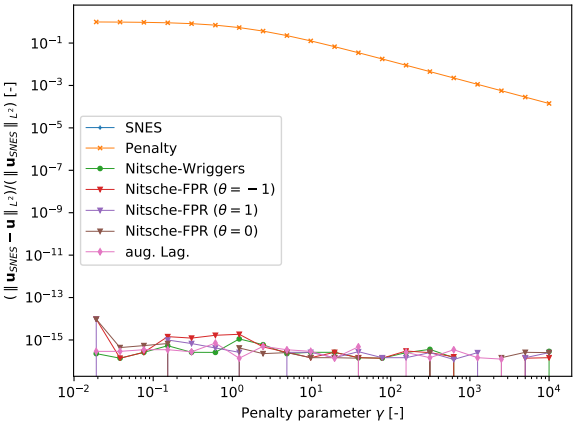


Figure 85: Relative error in displacements on $\overline{\Gamma}^C$ in L^2 norm for various penalty parameter values γ (log-log graph)

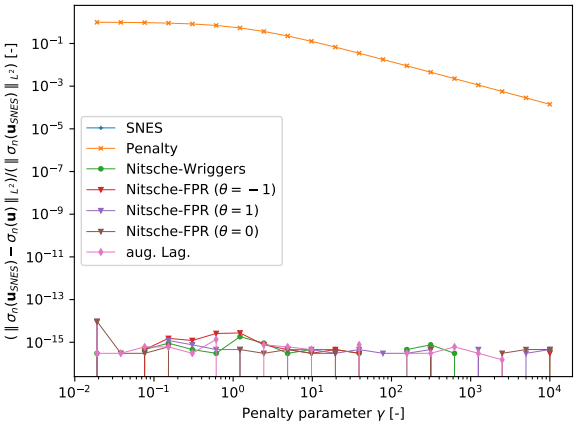


Figure 86: Relative error in the normal contact tractions on the active contact zone $\overline{\Gamma}^C$ in L^2 norm for various penalty parameter values γ (log-log graph)

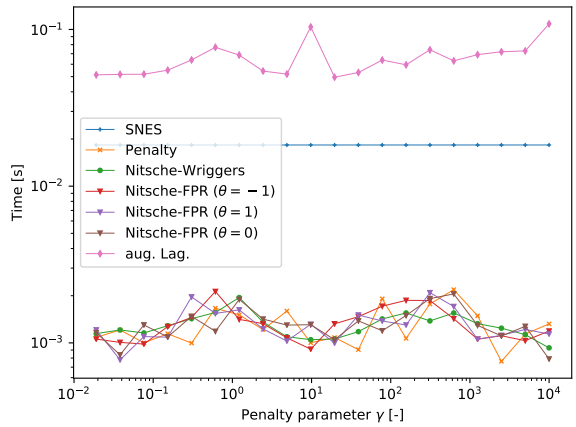


Figure 87: Runtime of solver for various penalty parameter values γ (log-log graph)

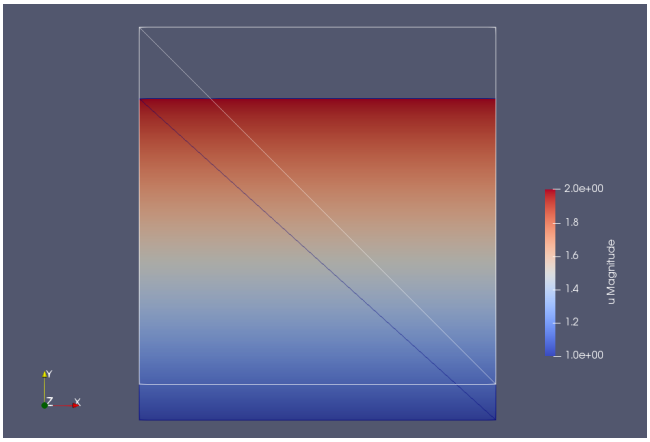


Figure 88: Undeformed mesh (white) and reference SNES solution (coloured)

4.3.2 Square block meshed with 384 triangular elements

This example is very similar to the previous one, except we consider a much finer mesh consisting of 384 triangular elements compared to 2 triangular elements in the previous example. We remind the parameters of the previous example. The length of the sides of the square is $a = 10$ mm and the initial gap $g_0 = 1$ mm. The downward displacement $\overline{u} = 2$ mm is prescribed at the top side of the square. Young's modulus is $E = 15000$ MPa and Poisson ratio is $\nu = 0.0$. Lateral displacements (in x direction) are prescribed as zero. Again, 20 values of penalty parameter are considered, and those are chosen to be the members of the geometric progression defined as (222). In this case, all methods (even Nitsche-Wriggers) converged for all the penalty parameter values.

The pattern of behaviour of all the methods is very similar to the same example meshed with two triangular elements (described in the previous section). Only in this case, the runtime is generally longer as there are much more DOFs and the runtime of SNES is closer to the runtime of the Nitsche methods and the penalty method than in the previous example. Again, the points representing the SNES method are not present in the log-log graphs of errors, as SNES is considered to be the reference solution and its error is again zero.

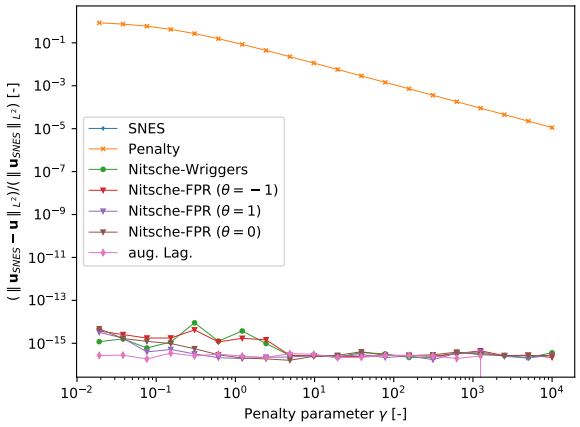


Figure 89: Relative error in displacements on $\overline{\Gamma}^C$ in L^2 norm for various penalty parameter values γ (log-log graph)

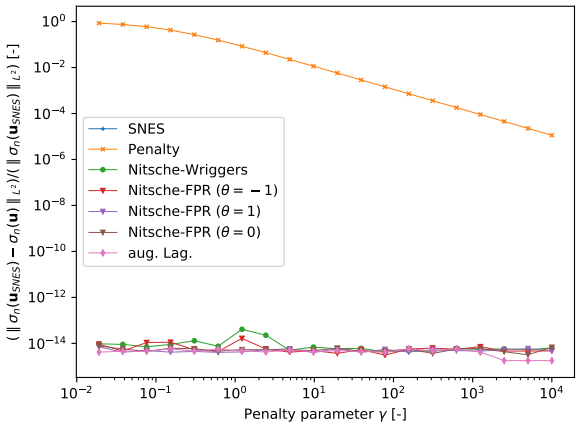


Figure 90: Relative error in the normal contact tractions on the active contact zone $\overline{\Gamma}^C$ in L^2 norm for various penalty parameter values γ (log-log graph)

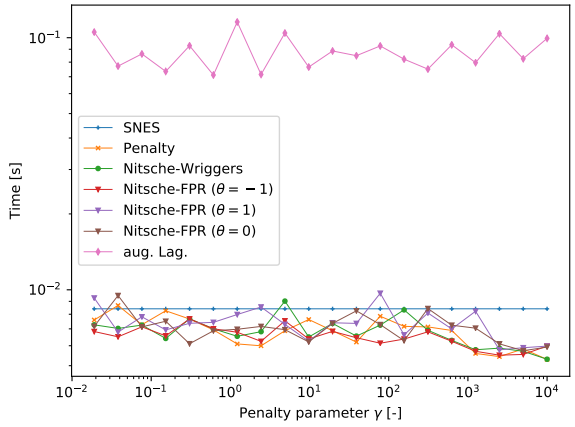


Figure 91: Runtime of solver for various penalty parameter values γ (log-log graph)

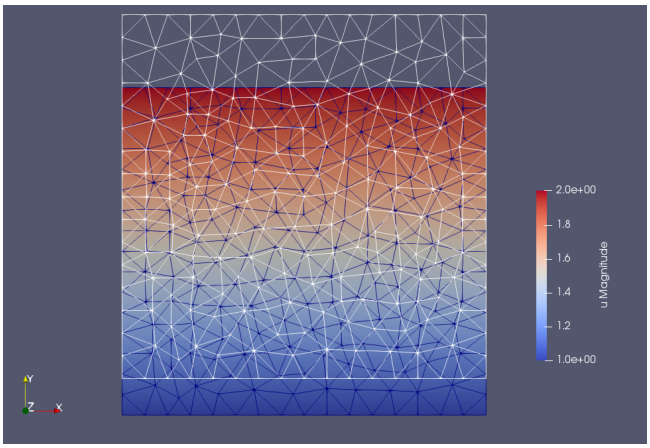


Figure 92: Undeformed mesh (white) and reference SNES solution (coloured)

4.3.3 Skewed guadrilateral block meshed with 2 triangular elements

In this example, we will investigate an example with the same geometry as in chapter 4.1 on the comparison of the numerical and the analytical solution — see Figure 3. So $a = 10$ mm and $\zeta \cdot a = 0.5$ mm. The downward displacement $\overline{u} = 2$ mm is prescribed at the top side of the guadrilateral. Contrarily to the example in chapter 4.1, the initial gap in the left node is $g_0 = 1$ mm. The quadrilateral is meshed with 2 triangular elements. We presume Young's modulus $E = 15000$ MPa and Poisson ratio $\nu = 0.0$. The choice of the Poisson ratio is motivated by the fact that no lateral displacements (in x direction) will occur and so we prescribe them directly as zeros to reduce the numerical error. Same as in the previous examples, the errors in displacement and tractions and the solver runtime are evaluated for 20 values of penalty parameter defined by the geometric progression (222). In this case, all methods except Nitsche-Wriggers converged for all penalty values. Nitsche-Wriggers did not converge for any penalty value.

In the figures below we see that for all the methods, the error decreases with the increasing penalty value. The penalty method produces the largest error both in displacements and tractions. The error of the Nitsche methods is similar with some small differences for small penalty values. Surprisingly, the Nitsche-Wriggers method, which did not converge for any penalty value, still produces a relatively precise solution. The precision is even similar to the Nitsche-FPR method. This is caused by the fact that the convergence of Nitsche-Wriggers was lost in the close vicinity of the solution where the solution oscillated between two values which were still very close to the theoretical solution. The augmented Lagrangian method produced the most precise solution, but on the other hand its solver runtime is the longest of all the methods (for all the investigated penalty values). We can observe that the solver runtime of Nitsche-Wriggers is higher than of the other Nitsche methods and the penalty method. This time corresponds to 100 Newton iterations which were set as the maximal number of Newton iterations. The SNES method is very effective in this case and its runtime is mostly shorter than of the other methods. We would like to remark here that compared to the previous examples with the square block, the displacement is now not linear with respect to x and y and therefore cannot be exactly represented by the linear approximation functions we consider. For this reason, the errors of the Nitsche methods and the augmented Lagrangian method are now larger. Again, the points representing the SNES method are not present in the log-log graphs of errors, as SNES is considered to be the reference solution and its error to be zero (in the sense that the gap at nodes 3 and 4 computed by this method is exactly zero).

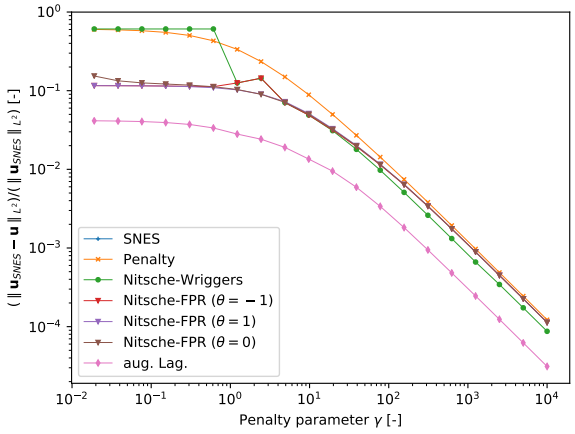


Figure 93: Relative error in displacements on $\overline{\Gamma}^C$ in L^2 norm for various penalty parameter values γ (log-log graph)

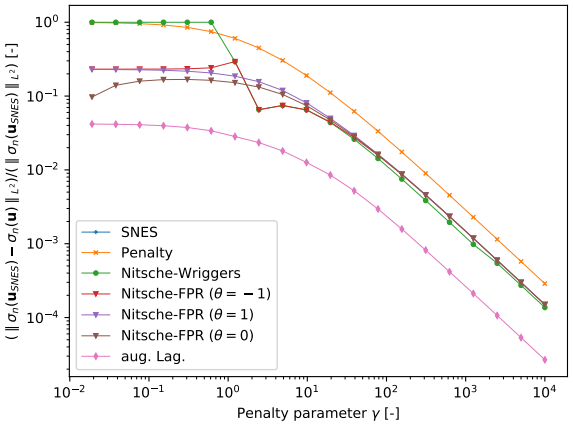


Figure 94: Relative error in the normal contact tractions on the active contact zone $\overline{\Gamma}^C$ in L^2 norm for various penalty parameter values γ (log-log graph)

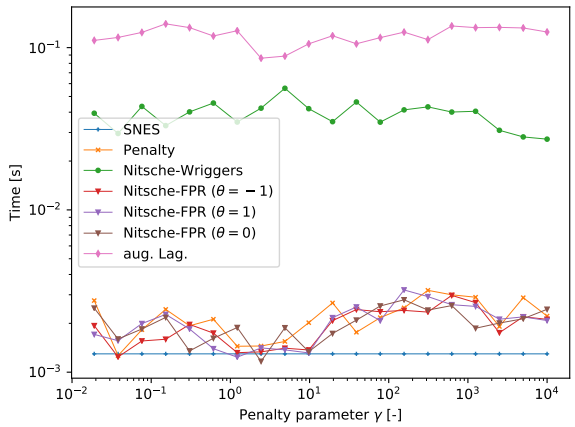


Figure 95: Runtime of solver for various penalty parameter values γ (log-log graph)

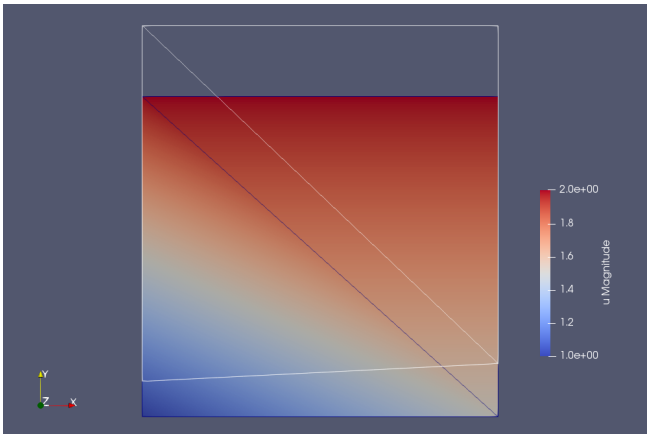


Figure 96: Undeformed mesh (white) and reference SNES solution (coloured)

4.3.4 Skewed quadrilateral block meshed with 364 triangular elements

This example is nearly the same as the one described in the previous section 4.3.3 except the quadrilateral is meshed with 364 elements. The geometry of the quadrilateral is depicted in Figure 3. The geometrical parameters and the considered penalty values are the same as in the previous section 4.3.3. Nitsche-Wriggers did not converge for any penalty value and Nitsche-FPR method did not converge for $\gamma = 2.441$. All the other methods converged for all the considered penalty values. In this case, the penalty method still produces the largest error. However, we can spot some significant differences in the error of different variants of the Nitsche-FPR method for small penalty values. As the Nitsche-Wriggers method did not converge, its precision is very poor in this case. One can also see that the SNES method is still very efficient in this case. We remind that the points representing the SNES method are not present in the log-log graphs of errors, as SNES is considered to be the reference solution and its error to be zero.

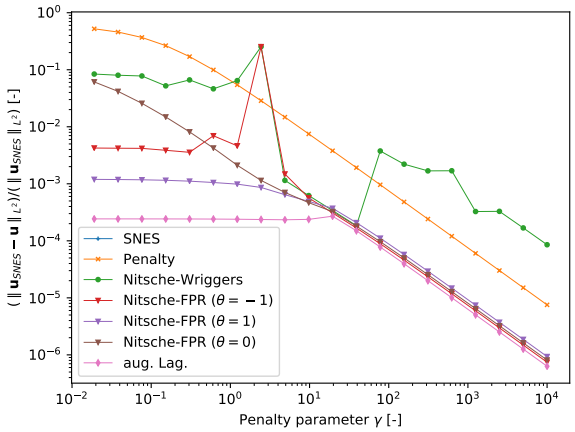


Figure 97: Relative error in displacements on $\overline{\Gamma}^C$ in L^2 norm for various penalty parameter values γ (log-log graph)

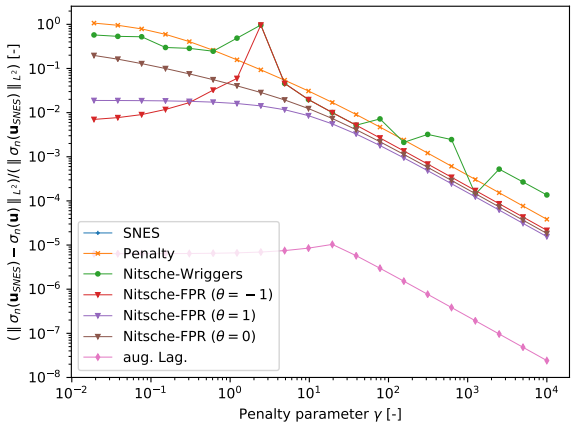


Figure 98: Relative error in the normal contact tractions on the active contact zone $\overline{\Gamma}^C$ in L^2 norm for various penalty parameter values γ (log-log graph)

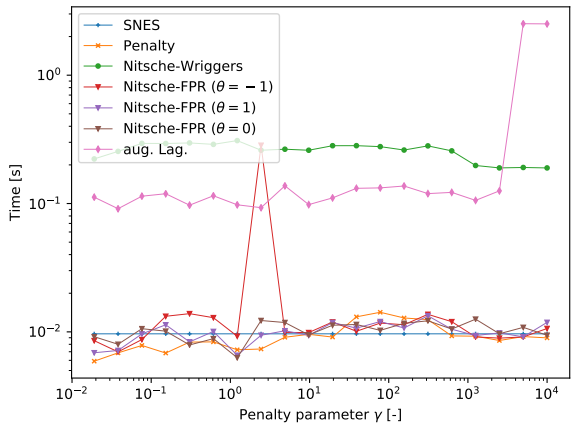


Figure 99: Runtime of solver for various penalty parameter values γ (log-log graph)

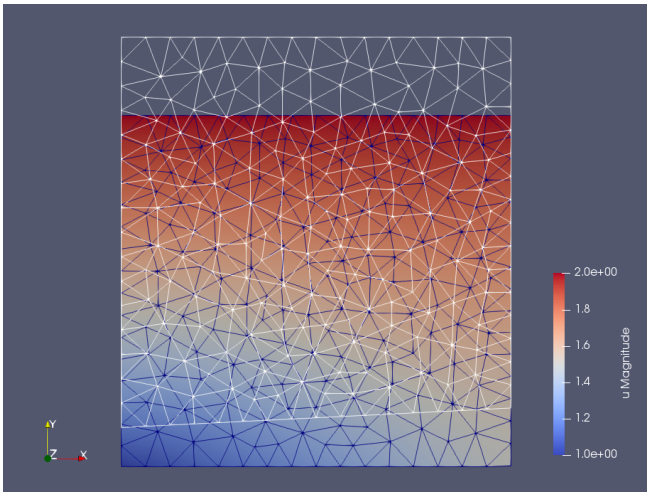


Figure 100: Undeformed mesh (white) and reference SNES solution (coloured)

4.3.5 'Cross-country ski'

The body we will investigate in this example inspired by the shape of cross-country ski. This was chosen because the potential contact zone is not a line segment in this case. For small prescribed displacement, there are two distant contact zones that enlarges with increasing prescribed displacement until they merge into one contact zone. We would like to point out that our objective was not to investigate the particular problem of ski but rather to investigate the errors produced in the case of more complicated shape of the potential contact zone. The length of the 'ski' is $l_{ski} = 3000$ mm and the height is $h_{ski} = 200$ mm. The 'bottom' boundary curve of the 'ski' is prescribed as

$$y = 20\left(1 - \sin\left(\frac{3}{l_{ski}}\pi x\right)\right) \tag{223}$$

Young's modulus is $E = 15000$ MPa, Poisson ratio $\nu = 0.20$. Plane stress is considered. The downward displacement $\overline{u} = 20$ mm is prescribed on a 300 mm long line segment on the top surface of the body. This is symmetric around the vertical axis (in y direction) of symmetry of the body. Lateral displacements $u_x = 0$ on this sub-domain. The body is meshed with 235,318 triangular elements. The penalty values are chosen in the same way as in the previous examples.

Figure 101: Surface of the 'cross-country ski'

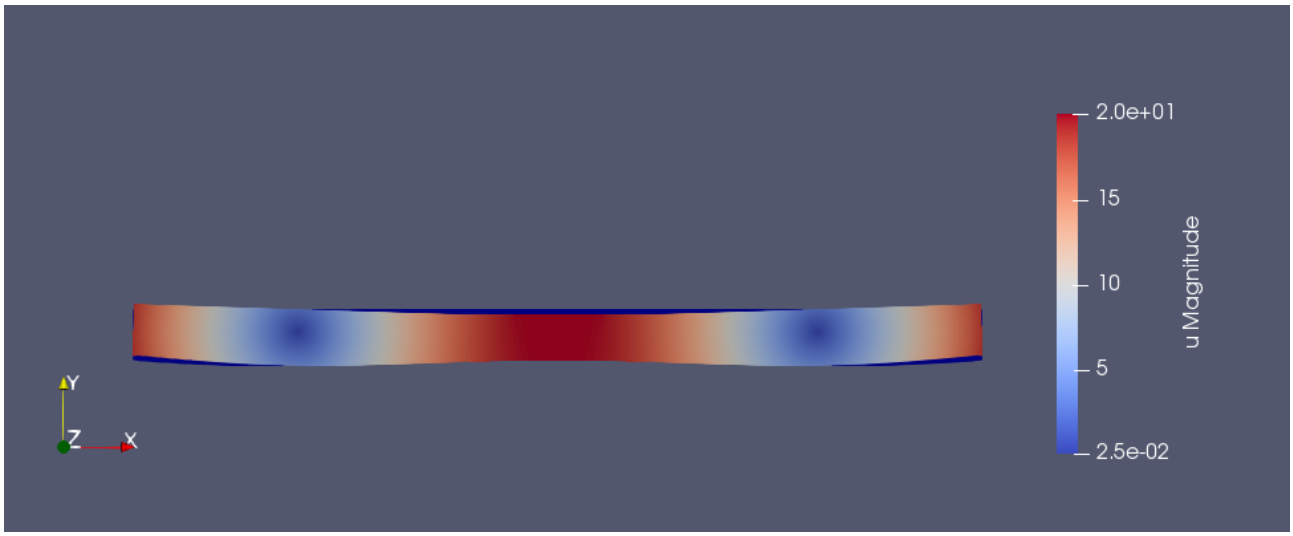


Figure 102: Undeformed mesh (dark blue) and reference SNES solution (coloured)

We would like to remark that in this thesis we use the SNES method as a reference solution. However, its runtime could become very long for a large number of DOFs. As we see in Figure 105, the Nitsche-FPR method and the penalty method is much faster for many penalty values. This suggests that the SNES method cannot be straightforwardly applied to large-scale problems as it would be too inefficient. Also, in Figure 105 we see that the effects of ill-conditioning become more and more apparent as the penalty value increases.

In this example, Nitsche-Wriggers again did not converge for any penalty value and Nitsche-FPR ($\theta = -1$) did not converge for $\gamma \leq 2.441$. Other methods converged for all the penalty values. We remind that the points representing the SNES method are not present in the log-log graphs of errors, as SNES is considered to be the reference solution and its error to be zero.

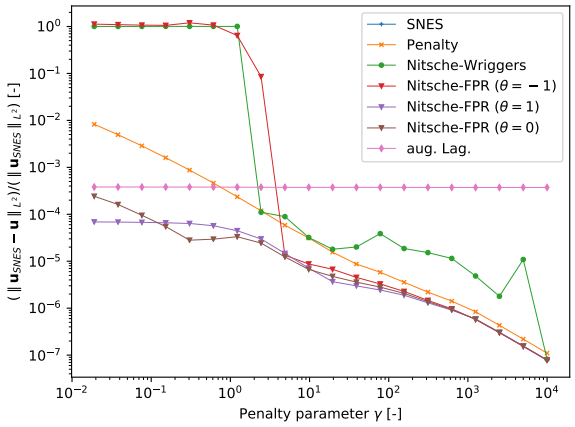


Figure 103: Relative error in displacements on $\overline{\Gamma}^C$ in L^2 norm for various penalty parameter values γ (log-log graph)

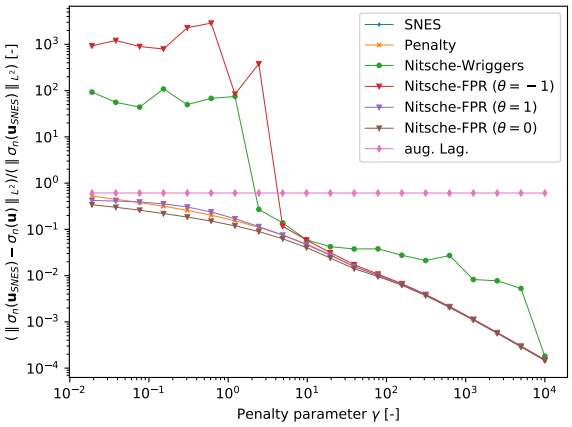


Figure 104: Relative error in the normal contact tractions on the active contact zone $\overline{\Gamma}^C$ in L^2 norm for various penalty parameter values γ (log-log graph)

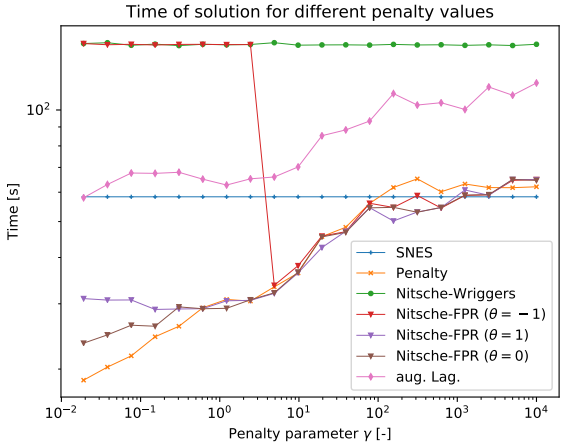


Figure 105: Runtime of solver for various penalty parameter values γ (log-log graph)

4.4 Three-dimensional examples

In this section, an example of an elastic hemisphere pressed against a rigid plane will be investigated. In this case, there exist an analytical solution by Hertz. The analytical solution of the contact force will be compared to the numerical predictions obtained from FEniCS for various methods for enforcing the contact boundary conditions. As in the previous chapter on two-dimensional examples, accuracy of all the investigated methods will be evaluated based on the relative error in displacements $\|\mathbf{u}\|_{rel}$ and the relative error in predicted contact tractions $\|\sigma_n(\mathbf{u})\|_{rel}$. Also, the computational efficiency represented by the runtime of the solver will be investigated for each method. The definitions of the mentioned indicators read

- Relative error in displacements $\|\mathbf{u}\|_{rel}$ on the active contact zone $\overline{\Gamma}^{C,S}$ of the SNES solution in L^2 norm

$$\|\mathbf{u}\|_{rel} = \frac{\|\mathbf{u}_{SNES} - \mathbf{u}\|_{L^2}}{\|\mathbf{u}_{SNES}\|_{L^2}} \tag{224}$$

  where

$$\|\mathbf{u}_{SNES} - \mathbf{u}\|_{L^2} = \sqrt{\int_{\overline{\Gamma}^{C,S}} (\mathbf{u}_{SNES} - \mathbf{u})^2 \, \mathrm{d}A} \tag{225}$$

$$\|\mathbf{u}_{SNES}\|_{L^2} = \sqrt{\int_{\overline{\Gamma}^{C,S}} \mathbf{u}_{SNES}^2 \, \mathrm{d}A} \tag{226}$$

- Relative error in predicted contact tractions σ_n on the active contact zone $\overline{\Gamma}^{C,S}$ of the SNES solution in L^2 norm

$$\|\sigma_n(\mathbf{u})\|_{rel} = \frac{\|\sigma_n(\mathbf{u}_{SNES}) - \sigma_n(\mathbf{u})\|_{L^2}}{\|\sigma_n(\mathbf{u}_{SNES})\|_{L^2}} \tag{227}$$

  where

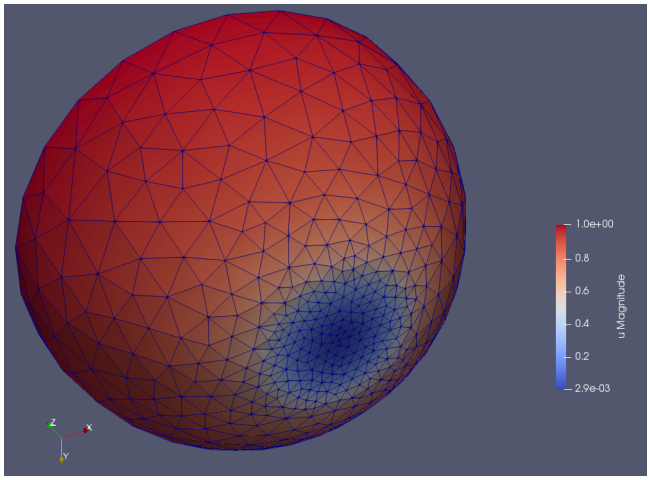
$$\|\sigma_n(\mathbf{u}_{SNES}) - \sigma_n(\mathbf{u})\|_{L^2} = \sqrt{\int_{\overline{\Gamma}^{C,S}} (\sigma_n(\mathbf{u}_{SNES}) - \sigma_n(\mathbf{u}))^2 \, \mathrm{d}A} \tag{228}$$

$$\|\sigma_n(\mathbf{u}_{SNES})\|_{L^2} = \sqrt{\int_{\overline{\Gamma}^{C,S}} (\sigma_n(\mathbf{u}_{SNES}))^2 \, \mathrm{d}A} \tag{229}$$

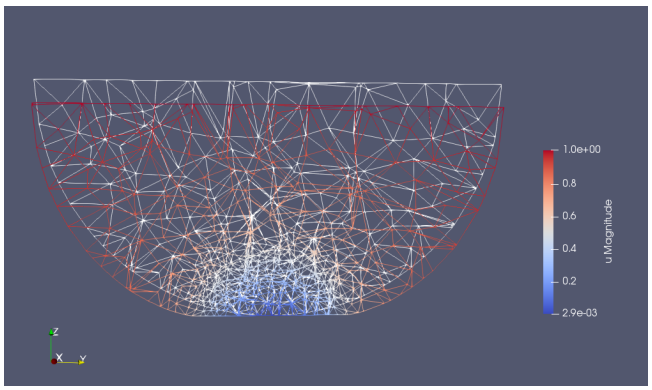
- Runtime of solver in $[s]$

We remind again that $SNES$ denotes the solution obtained using the non-linear solver SNESVINEWTONRSLS from PETSc library [50], which employs the reduced space active set solver for variational inequalities based on semi-smooth Newton's method [9]. For its very high precision, this method is used as a reference solution. One example of the SNES solution is showed in Figure 106. Again, we point out that this method does not require any value of penalty parameter.

For all the investigated problems, the radius of the hemisphere is $R = 10$ mm and the prescribed downward displacement $\overline{u} = 1$ mm. The prescribed displacement is set as $u_z = -1$ in the nodes on the flat part of the surface of the considered hemisphere. Zero displacements in the lateral directions are also prescribed in the nodes on the small circular region on the flat surface of the hemisphere. Mechanically, it would be of course possible to prescribe the lateral displacements just in two nodes but this is not natural to FEniCS where the user is supposed to provide the geometric parameters of domains where the boundary conditions are prescribed. The centre of the circle is located at the centre of the hemisphere. Young's elastic modulus $E = 15000$ MPa and Poisson ratio $\nu = 0.25$. The hemisphere is meshed with quadrilateral elements, and linear approximation functions for displacement and Lagrange multipliers are considered. An example of such a mesh is depicted in Figure 107. From this, we can also observe that the mesh is not uniform and is refined in the area around the potential contact zone. Also, we need to set a parameter h for the penalty value defined as $E\gamma/h$. We

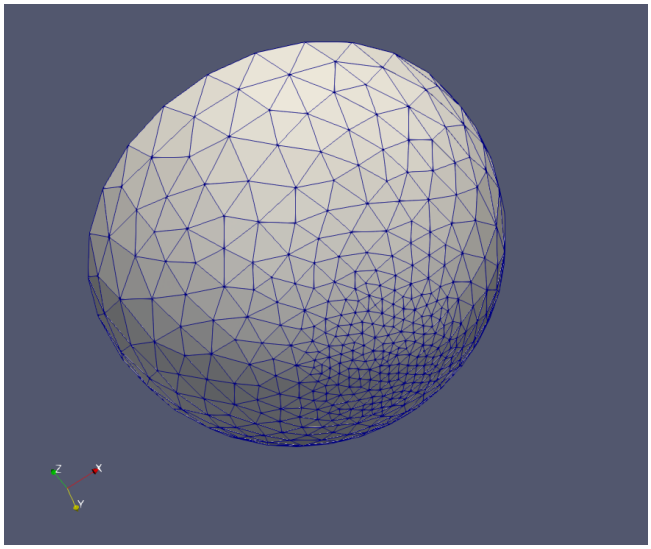


(a) Deformed mesh

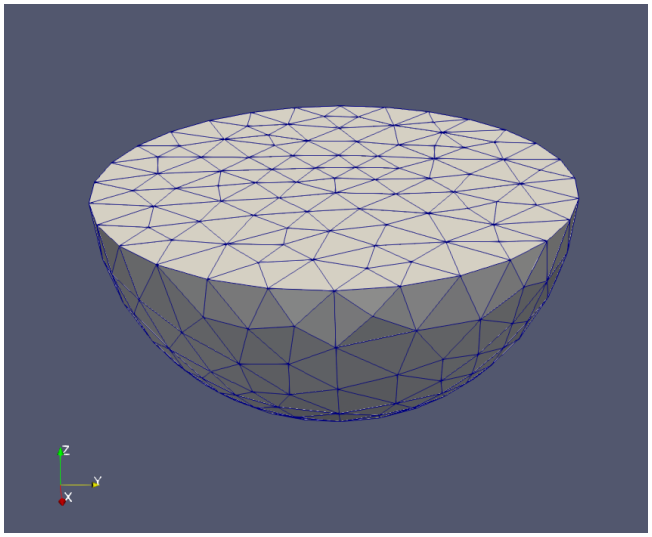


(b) Slice in $x = 0$ — deformed mesh (coloured) and undeformed mesh (white)

Figure 106: Reference SNES solution



(a) Mesh view 1



(b) Mesh view 2

Figure 107: Views of the mesh

remind that h represents the 'characteristic size' of the mesh elements and is supposed to be constant over the mesh. In this case, we define it as the arithmetic average of the largest and the smallest diameter of the sphere circumscribed to the elements.

The absolute and relative tolerance of Newton's solver are set as $tol_{abs} = 10^{-9}$ and $tol_{rel} = 10^{-9}$. By these tolerances, the value of residuum (which could be interpreted as an unbalanced force) is limited in every iteration of Newton's method. The maximum number of Newton's method iterations for convergence is 50.

4.4.1 Results for variable penalty parameter

In this case, the problem is evaluated for 20 values of the penalty parameter. These values are chosen to be the members of the geometric progression

$$\gamma = 10^4 \left(\frac{1}{2}\right)^{i-1}, \quad i = 1, 2, ..., 20 \tag{230}$$

This kind of series was selected so there would be enough points reflecting small penalty values and supposed log-log scale character of the plot. The largest value of γ is 10000 and the smallest 0.019. We used the mesh with 4588 quadrilateral elements.

The Nitsche-Wriggers method did not converge for any value of the penalty parameter. This is evident also in Figure 110 where the time of every run of Nitsche-Wriggers takes nearly the same time corresponding to 100 iterations. In Figure 110 also we see that the convergence of Nitsche-FPR ($\theta = -1$) and Nitsche-FPR ($\theta = 0$) is lost if the penalty parameter is smaller than a certain value. The penalty values for which the Nitsche-FPR computation did not converge can be graphically identified from the solver runtime, because their runtime corresponds to the maximal allowed number of Newton iterations (100) — the runtime is then nearly the same as for Nitsche-Wriggers, which did not converge for any penalty value. In terms of the penalty parameters we tested, the largest one for which Nitsche-FPR ($\theta = -1$) did not converge was $\gamma = 19.531$ and for Nitsche-FPR ($\theta = 0$) it was $\gamma = 0.305$. In each case, the computation did not converge for any penalty value smaller than these values. On the other hand, Nitsche-FPR ($\theta = 1$) converged for all the tested penalty parameters. This reflects the fact that (as we mentioned in the section on the mathematical properties of the Nitsche-FPR method) Nitsche-FPR ($\theta = -1$) and Nitsche-FPR ($\theta = 0$) requires a certain minimal value of the penalty parameter so that the problem is well-posed. On the other hand, Nitsche-FPR ($\theta = 1$) is well posed for every positive penalty value.

It is quite noteworthy that for small penalty values, Nitsche-FPR ($\theta = 1$) produces the smallest error in displacements (see Figure 108) which is even smaller than the error of the augmented Lagrangian method (which in other examples is usually the most precise method). On the other hand, for these penalty values, the computational time is longer than for the penalty method, but still shorter than for the augmented Lagrangian method. Also, in Figure 109 we see that except for a few very small penalty values, the error in normal contact tractions of Nitsche-FPR ($\theta = 1$) is larger than of the other methods except Nitsche-FPR ($\theta = -1$).

One can also observe that the SNES method is remarkably efficient in terms of the runtime. However, it is presumed that the runtime for large-scale problems (which were not tested in this thesis) would be much longer compared to other methods.

Also, effects of high penalty values (causing ill-conditioning of the stiffness matrix) on increasing runtime of solver could be observed in Figure 110. The points representing the SNES method are not present in the log-log graphs of errors, as SNES is considered to be the reference solution and its error to be zero.

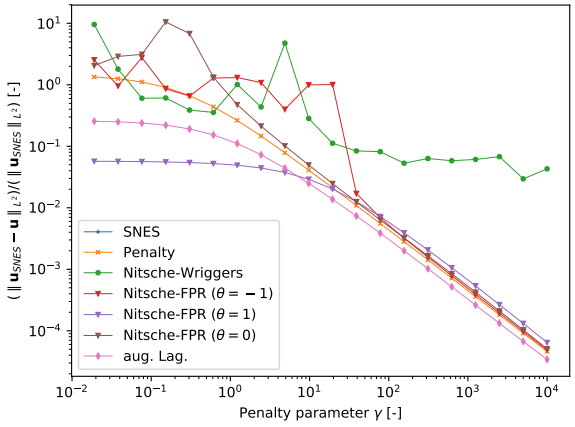


Figure 108: Relative error in displacements on $\overline{\Gamma}^C$ in L^2 norm for various penalty parameter values γ (log-log graph)

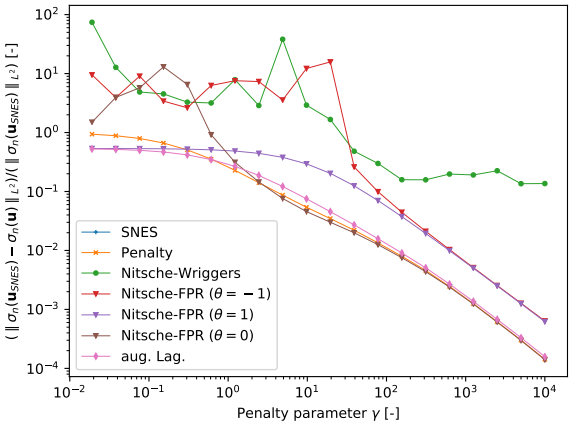


Figure 109: Relative error in the normal contact tractions on the active contact zone $\overline{\Gamma}^C$ in L^2 norm for various penalty parameter values γ (log-log graph)

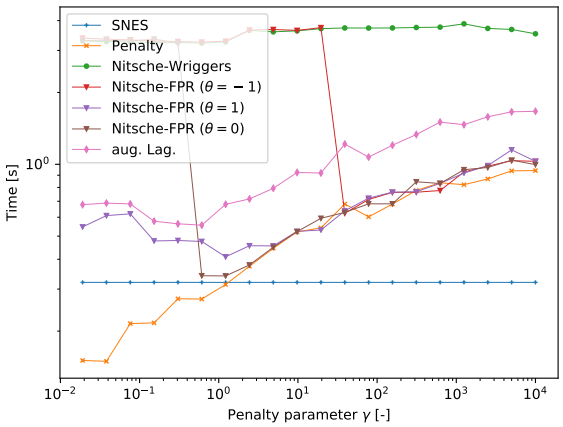


Figure 110: Runtime of solver for various penalty parameter values γ (log-log graph)

4.4.2 Results for variable number of DOFs

Because in the rest of the investigated examples we consider the same mesh, and change some parameters as the penalty or the prescribed downward displacement, it is of interest whether using a different (finer) mesh could provide significantly different results. For this reason, we measured the investigated accuracy and efficiency indicators for 5 meshes with different numbers of DOFs. The maximal considered number of DOFs was 38283 and the minimal 14556. This number includes only the number of DOFs connected with displacements. In the augmented Lagrangian method, there are additional DOFs representing Lagrange multipliers, but these are not included in the previous numbers. The penalty parameter for all methods (except SNES which does not require it) was set to 50. All methods converged except Nitsche-Wriggers method, which did not converge for any value of the penalty parameter. In Figure 113 we see that the used direct solver is relatively efficient, as it is effectively utilizing the sparse structure of the stiffness matrix. The points representing the SNES method are not present in the log-log graphs of errors, as SNES is considered to be the reference solution and its error to be zero.

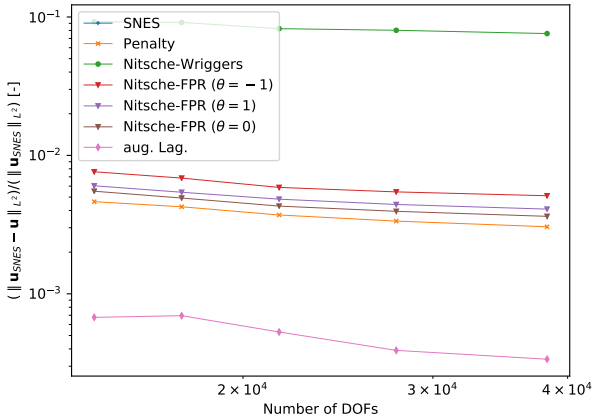


Figure 111: Relative error in displacements on $\overline{\Gamma}^C$ in L^2 norm for various number of DOFs (log-log graph)

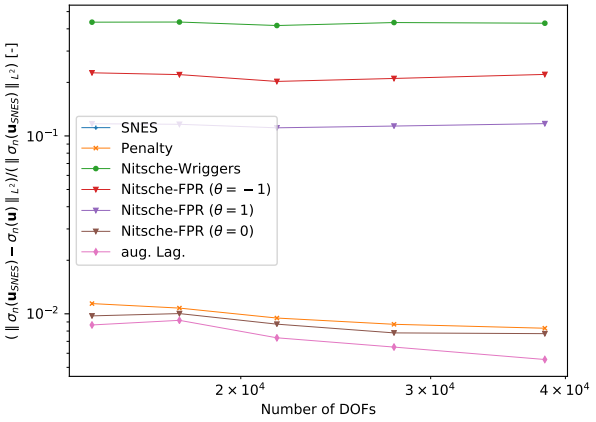


Figure 112: Relative error in the normal contact tractions on the active contact zone $\overline{\Gamma}^C$ in L^2 norm for various number of DOFs (log-log graph)

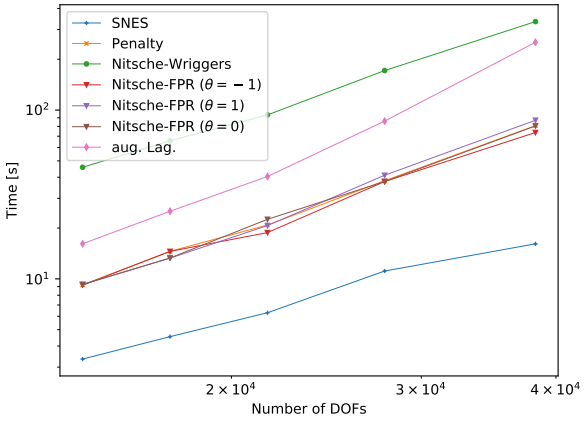


Figure 113: Runtime of solver for various number of DOFs (log-log graph)

4.4.3 Results for variable prescribed displacement

For the case of contact of a linear elastic hemisphere with a rigid plane (considering small deformation) there exists an analytical solution by Hertz (see the classic book [33, Chapter 4.2, p.92-93]). So, we can analytically calculate the contact force F_{HERTZ} as

$$F_{HERTZ} = \frac{4a^3 E^*}{3R} \tag{231}$$

where

$$E^* = \frac{E}{1-\nu^2} \tag{232}$$

$$a = \sqrt{R\overline{u}} \tag{233}$$

In the previous examples we considered the constant value of the prescribed downward displacement $\overline{u} = 1$. This stood for $u_z = -1$ in the nodes on the flat part of the surface of the considered hemisphere. Now, we will change the prescribed displacement and compare the numerical results to the analytical solution. The value of penalty parameter for all the methods is chosen 50, and for all the calculations the same mesh with 7738 quadrilateral elements is used. The prescribed displacement values are chosen to be the members of the geometric progression

$$\gamma = 2\left(\frac{1}{2}\right)^{i-1}, \quad i = 1, 2, ..., 8 \tag{234}$$

so the largest prescribed displacement is 2 mm and the smallest 0.0156 mm. Except for Nitsche-Wriggers method, all computations converged.

From the graphs below, we can observe that with decreasing prescribed displacement, up to a certain value, the numerical predictions are closer to the analytical one. However, if the relative prescribed displacement $\overline{u}/R$ is too small, the difference between numerical and analytical solution increases. This is caused by the fact that too few elements are present in the active contact zone, so the numerical predictions start to be very inaccurate. Also, we can observe that for our mesh, even the SNES method, which we use as a reference solution, significantly differs from the Hertz solution. For a finer mesh, better agreement between numerical predictions and Hertz solution can be expected. It is important to note the analytical Hertz solution is also an approximation based on many assumptions. However, for small prescribed displacement values it is a very precise approximation. For larger prescribed displacement, the numerical solution (for example our reference SNES method) could provide even more precise estimate of the contact force than the Hertz solution.

To measure the error in the predicted contact force, we calculated the contact force on the potential contact zone $\Gamma^{C,S}$ and also the reaction force on the flat surface of the hemisphere where the displacement is prescribed. The forces are evaluated as the volume integrals over the elements which have facets on the given boundary region. For the reaction force, this region is the flat surface of the hemisphere. For the contact force, it is the potential contact zone. In the ideal case, when there would be zero normal contact tractions on the non-active part of the potential contact zone, the forces would be in balance. However, in Figures 115 and 117 we see that the forces differ. One could incorrectly interpret this as a violation of equilibrium. However, the body is in equilibrium (in the weak sense) and this discrepancy is caused by the way we measure the forces. The normal tractions on the non-active part of the potential contact zone are not completely zero as there is some numerical error. We would like to note that there are other ways how to evaluate the contact force. One could for example integrate the normal contact tractions (projected to the z direction) over the active contact zone. However, it could be shown that this would lead to an incorrect result, unless the gap function is exactly zero on the active contact zone (which is not always true because there is usually some small numerical error).

From the gained results it seems that the contact and the reaction force predicted by the augmented Lagrangian method are nearly the same as the forces predicted by the reference SNES method. Interestingly, even the penalty method seems to provide relatively precise estimate of the contact and the reaction force.

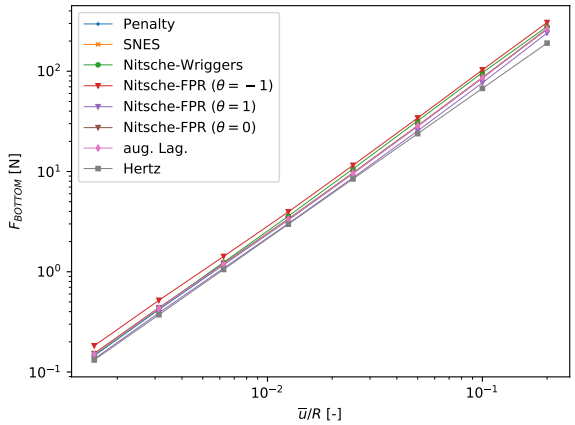


Figure 114: Absolute contact force calculated on the active contact zone $\overline{\Gamma}^C$ for the various relative prescribed displacement values (log-log graph)

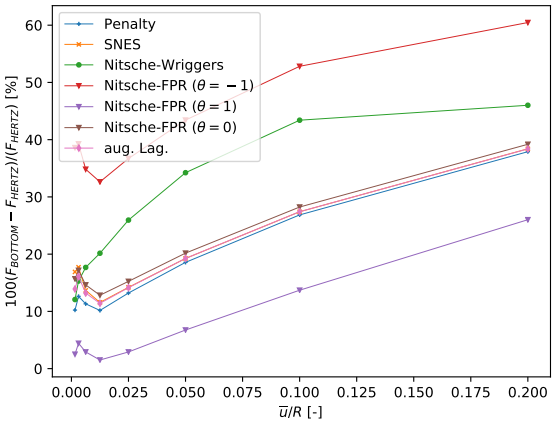


Figure 115: Relative error in contact force calculated on the active contact zone $\overline{\Gamma}^C$ for the various relative prescribed displacement values

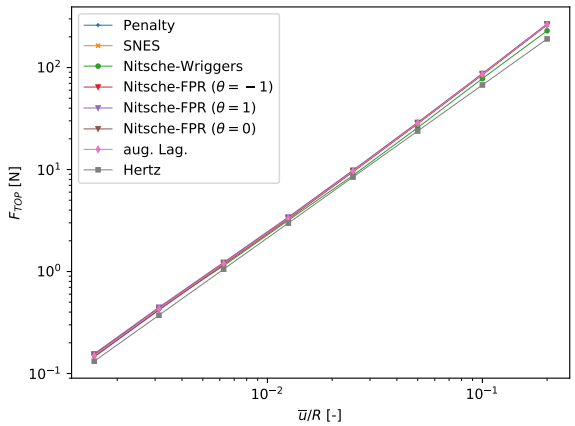


Figure 116: Absolute contact force calculated on the flat surface of hemisphere for the various relative prescribed displacement values (log-log graph)

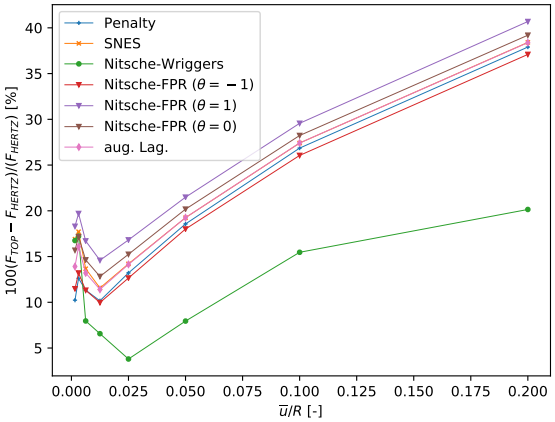


Figure 117: Relative error in the reaction force calculated on the flat surface of hemisphere for the various relative prescribed displacement values

5 Conclusion

The purpose of this thesis was to analyse a wide range of methods for enforcing the contact boundary conditions. Within theoretical formulations for a continuum, we considered two linear elastic bodies in frictionless contact. We adopted the master-slave approach, in which the whole domain Ω is separated into a slave body Ω^S and a master body Ω^M such that $\Omega = \Omega^S \cup \Omega^M$. Boundaries of the slave and master body are separated into three non-overlapping domains — $\Gamma^{C,S}$ and $\Gamma^{C,M}$, where contact conditions are prescribed, $\Gamma^{\sigma,S}$ and $\Gamma^{\sigma,M}$, where tractions $\overline{\mathbf{t}}$ are prescribed and $\Gamma^{D,S}$ and $\Gamma^{D,M}$ where displacements $\overline{\mathbf{u}}$ are prescribed. It is important to note that as in contact problems we generally do not know the actual contact region in advance, $\Gamma^{C,S}$, $\Gamma^{C,M}$ are the so-called potential contact zones. It is supposed that active contact zones $\overline{\Gamma}^{C,S}$ and $\overline{\Gamma}^{C,M}$ are the sub-domains of potential contact zones, i.e., $\overline{\Gamma}^{C,S} \subset \Gamma^{C,S}$ and $\overline{\Gamma}^{C,M} \subset \Gamma^{C,M}$. In terms of the master-slave approach, contact boundary conditions are prescribed only on $\Gamma^{C,S}$. In chapter 2, the strong form of the elastic frictionless contact problem was presented in (16) along with the contact boundary conditions (also called the Hertz–Signorini–Moreau conditions) on $\Gamma^{C,S}$ (17). As we consider a small elastic deformation, the Cauchy stress tensor $\boldsymbol{\sigma}(\mathbf{u})$ and the tensor of small deformation $\boldsymbol{\varepsilon}(\mathbf{u})$ are utilized and so the strong form of contact problem (16), (17) reads

$$\begin{aligned} -\boldsymbol{\nabla}\cdot\boldsymbol{\sigma}\left(\mathbf{u}\right) &= \rho\overline{\mathbf{b}} && \text{in} && \Omega \\ \boldsymbol{\sigma}\left(\mathbf{u}\right) &= \mathbf{D}\boldsymbol{\varepsilon}\left(\mathbf{u}\right) && \text{in} && \Omega \\ \boldsymbol{\varepsilon}(\mathbf{u}) &= \frac{1}{2}\left(\boldsymbol{\nabla}\mathbf{u} + (\boldsymbol{\nabla}\mathbf{u})^T\right) && \text{in} && \Omega \\ \mathbf{u} &= \overline{\mathbf{u}} && \text{on} && \Gamma^D \\ \mathbf{n}\cdot\boldsymbol{\sigma} &= \overline{\mathbf{t}} && \text{on} && \Gamma^\sigma \end{aligned} \tag{235}$$

$$\begin{aligned} g_n &\geq 0 && \text{on} && \Gamma^{C,S} \\ \sigma_n &\leq 0 && \text{on} && \Gamma^{C,S} \\ \sigma_n g_n &= 0 && \text{on} && \Gamma^{C,S} \end{aligned} \tag{236}$$

Here, g_n is the gap function, which is the metric of the distance (which, if negative, becomes penetration) of a master surface from a slave surface, defined as

$$g_n(\mathbf{u}) = \mathbf{X}^S\cdot\mathbf{n}^M + \mathbf{u}^S\cdot\mathbf{n}^M - \mathbf{X}^M\cdot\mathbf{n}^M - \mathbf{u}^M\cdot\mathbf{n}^M \tag{237}$$

where $\mathbf{X}$ denotes the position vector in the original (undeformed) configuration of Ω and $\mathbf{n}^M$ is a unit normal vector to the master surface. By σ_n we mean the normal contact traction. Contact tractions on Γ^C are defined as

$$\mathbf{t}^S = \mathbf{n}^S\cdot\boldsymbol{\sigma}^S = -\sigma_n^S\mathbf{n}^M \tag{238}$$

where

$$\sigma_n^S = -\mathbf{n}^S\cdot\boldsymbol{\sigma}^S\cdot\mathbf{n}^M \tag{239}$$

and

$$\mathbf{t}^M = \mathbf{n}^M\cdot\boldsymbol{\sigma}^M = \sigma_n^M\mathbf{n}^M \tag{240}$$

where

$$\sigma_n^M = \mathbf{n}^M\cdot\boldsymbol{\sigma}^M\cdot\mathbf{n}^M \tag{241}$$

For the purposes of application of the finite element method, a weak form of (16), (17) must be found. In chapter 2, it is shown that this could take the form of a variational inequality (30)

$$\int_\Omega \boldsymbol{\sigma}\left(\mathbf{u}\right) : \boldsymbol{\nabla}\mathbf{v}\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\mathbf{v}\,\mathrm{d}V - \int_{\Gamma^\sigma}\overline{\mathbf{t}}\cdot\mathbf{v}\,\mathrm{d}A \geq 0 \qquad \forall\mathbf{v}\in\mathbf{K} \tag{242}$$

$$\mathbf{V} = \left\{ \mathbf{v} \in H^1(\Omega)^d ; \mathbf{v} = \mathbf{0} \text{ on } \Gamma^D \right\} \tag{243}$$

$$\mathbf{K} = \left\{ \mathbf{v} \in \mathbf{V}; \delta g_n = \mathbf{v}^S \cdot \mathbf{n}^M - \mathbf{v}^M \cdot \mathbf{n}^M \geq 0 \text{ on } \overline{\Gamma}^{C,S} \right\} \tag{244}$$

Even though this formulation is possible, Yastrebov mentions in [57, Chapter 4.4, p.144] that: '...the variational inequality is hard to apply for contact with finite sliding and/or rotations. That is why, nowadays, most of the practical studies in contact mechanics are based on the so-called variational equalities, which are easy to introduce in a finite element framework and does not require totally new minimization techniques.' For this reason, the focus of this thesis is on the methods that allow us to formulate the elastic frictionless contact problem as a variational equality. In this thesis we considered only so-called unconstrained methods. In terms of our thesis this means that the contact problem is formulated by a certain equality and no additional conditions have to be fulfilled.

From the optimization theory, three classes of approaches to this problem could be distinguished, namely the penalty methods, mixed methods and Nitsche methods. The first two approaches are nowadays standard and represent the backbone of contact computations in the currently leading finite element software packages such as ANSYS, ABAQUS and COMSOL [57, Chapter 1.1.1, p.7]. Although the idea of Nitsche method comes from the 1970's paper by Joachim A. Nitsche, who proposed a method for treating interface conditions between meshes, the application of Nitsche methods to contact problems is a relatively novel topic of research. Some major advances applying this method to contact mechanics were made by Hansbo, A., Hansbo, P. and co-authors [25], [26] and by Wriggers and Zavarise [56] at the beginning of the new millennium and recently by Chouly, Hild and Renard [29], [30] in 2013 and 2015 respectively.

In this thesis, we focus on the Nitsche method, as it seems that it has the potential to overcome the most important drawbacks of the penalty and mixed methods. Penalty methods are known to be easy to implement, but they are not consistent,[9] and the contact conditions are not fulfilled exactly. A small amount of penetration is allowed, depending on the value of the penalty parameter. The higher it is, the smaller penetration is obtained. Theoretically, we should choose as high value as possible, but in reality, the maximum value is limited, as choosing the too high value leads to the ill-conditioning of a discrete problem. On the contrary, mixed methods are consistent, and contact conditions are fulfilled exactly, but they introduce a new unknown variable (except displacements $\mathbf{u}$) called the Lagrange multiplier. This has the physical meaning of the normal contact traction. However, the resulting problem is a saddle point of the corresponding Lagrangian, and so fulfilling Babuška-Brezzi condition at the discrete level is required. This means that approximation functions must be chosen appropriately for displacements and Lagrange multiplier, as naive some choice could lead to an unstable discrete problem. Also, the number of DOFs of the discrete problem is significantly enlarged as Lagrange multiplier values are represented by additional DOFs. For large problems, this poses a serious problem. As suggested, Nitsche methods have the potential to overcome weak points of both penalty and mixed methods. They are consistent, and at the continuum level, contact conditions are fulfilled exactly irrespectively of the penalty value. Even though some minimal value of the penalty parameter is still necessary, it could be chosen significantly smaller than for the penalty method, and serious ill-conditioning is therefore avoided. At the same time, Nitsche method does not introduce any additional unknown, so the number of DOFs and computational cost of a discrete problem are not increased. Also, we do not have to concern ourselves with fulfilling Babuška-Brezzi condition.

For the suggested reasons, we analysed two versions of the Nitsche method in chapter 3. Moreover, in the

[9] The meaning of consistency was specified in the chapter on the mathematical properties of Nitsche-FPR. By consistency we mean that if there is a solution $\mathbf{u}$ to the strongly formulated contact problem (16), (17) and $\mathbf{u}$ is sufficiently regular then $\mathbf{u}$ is also the solution to the weak problem formulated with our method.

same chapter, we also presented the augmented Lagrangian method and one version of the penalty method, which are closely connected to the Nitsche methods we analysed. One of the presented Nitsche methods is by Wriggers and Zavarise (denoted as Nitsche-Wriggers) [55, Chapter 5, p.106], [56] and the other, more recent, is by Fabré, Pousin and Renard from 2016 [15] (denoted as Nitsche-FPR). The Nitsche-Wriggers method is based on the paper by Wriggers and Zavarise [56]. However, we have not found any derivation of this method for the case of frictionless contact in the scientific literature, so in this thesis, we presented a derivation of this method inspired by the derivation of the Discontinuous Galerkin method. The Nitsche-FPR method is adopted from [15] where the derivation is provided, and we presented it as well. We note that the authors of Nitsche-FPR based their method on the ideas of Chouly, Hild and Renard [29], [30] published in 2013 and 2015. It comes quite surprising that both methods are presented as Nitsche methods by their authors, but the resulting weak forms are not the same. To our knowledge, we do not know of any analysis of their differences in the scientific literature, so we provided one in this thesis. It turned out that on the active contact zone, the weak forms of both methods coincide, but they are different on the parts of the potential contact zone, where contact is not active. Also, while Nitsche-Wriggers uses the gap as the contact indicator, Nitsche-FPR uses the sign of $g_n + \frac{h}{\gamma E}\sigma_n$, i.e., of the gap plus some multiple of the traction normal to the contact surface. On the theoretical continuum level, these differences do not play any role, but at a discrete level, they certainly do, as was documented by numerical simulations. Also, during the numerical tests it was observed that the Nitsche-Wriggers method has significant troubles to converge in FEniCS and very careful implementation of some terms of its weak form is necessary. On the contrary, Nitsche-FPR proved itself to be quite stable for sufficiently high penalty values during the simulations we carried out.

Besides the Nitsche-Wriggers method and the Nitsche-FPR method, we also analysed the augmented Lagrangian method and one version of the penalty method. The augmented Lagrangian method is taken from Yastrebov in [57, Chapter 4.7.2, p.197-180], who in turn based his formulation on Rockafellar ([46] and [47]). The penalty method is also based on the formulation by Yastrebov [57, Chapter 4.5.1, p.145-146]. The presented penalty method could be seen as an inconsistent version of the Nitsche-Wriggers method (as some of the Nitsche-Wriggers terms are missing). On the other hand, Nitsche-FPR has a strong connection to the augmented Lagrangian method and the stabilised augmented Lagrangian method by Barbosa and Hughes [6], which allow us to circumvent the discrete Babuška-Brezzi condition.

To sum up, the following four interconnected methods are presented — the Nitsche-Wriggers method, the Nitsche-FPR method, the augmented Lagrangian method and the penalty method. Their weak forms are summarized below. In these, E is Young's modulus and h is the characteristic size of the mesh elements (considered to be constant over the mesh). These parameters are introduced so that the penalty parameter γ is dimensionless.

1) Nitsche-Wriggers method (113)

$$\begin{aligned}
&\int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{v})\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\mathbf{v}\,\mathrm{d}V - \int_{\Gamma^\sigma} \overline{\mathbf{t}}\cdot\mathbf{v}\,\mathrm{d}A \\
&+ \int_{\Gamma^{C,S}} \langle\sigma_n(\mathbf{u})\rangle H(-g_n(\mathbf{u}))\delta g_n(\mathbf{v})\,\mathrm{d}A - \int_{\Gamma^{C,S}} \langle\sigma_n(\mathbf{v})\rangle\{-g_n(\mathbf{u})\}\,\mathrm{d}A \\
&- \frac{\gamma E}{h}\int_{\Gamma^{C,S}} \{-g_n(\mathbf{u})\}\delta g_n(\mathbf{v})\,\mathrm{d}A = 0 \quad \forall\mathbf{v}\in\mathbf{V}
\end{aligned} \tag{245}$$

2) Penalty method (157)

$$\begin{aligned}
&\int_\Omega \boldsymbol{\sigma}\,(\mathbf{u}) : \boldsymbol{\nabla}\mathbf{v}\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\mathbf{v}\,\mathrm{d}V - \int_{\Gamma^\sigma} \overline{\mathbf{t}}\cdot\mathbf{v}\,\mathrm{d}A \\
&- \frac{\gamma E}{h}\int_{\Gamma^{C,S}} \{-g_n(\mathbf{u})\}\delta g_n(\mathbf{v})\,\mathrm{d}A = 0 \quad \forall\mathbf{v}\in\mathbf{V}
\end{aligned} \tag{246}$$

3) Nitsche-FPR method(134)

$$\begin{aligned}&\int_\Omega \boldsymbol{\sigma}\left(\mathbf{u}\right) : \boldsymbol{\nabla}\mathbf{v}\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\mathbf{v}\,\mathrm{d}V - \int_{\Gamma^\sigma}\overline{\mathbf{t}}\cdot\mathbf{v}\,\mathrm{d}A\\&+\int_{\Gamma^{C,S}}\theta\frac{h}{\gamma E}\langle\sigma_n(\mathbf{u})\rangle_t\langle\sigma_n(\mathbf{v})\rangle_t\,\mathrm{d}A\\&-\int_{\Gamma^{C,S}}\frac{\gamma E}{h}\left\{-g_n(\mathbf{u})-\frac{h}{\gamma E}\langle\sigma_n(\mathbf{u})\rangle_t\right\}\left(\delta g_n(\mathbf{v})-\theta\frac{h}{\gamma E}\langle\sigma_n(\mathbf{v})\rangle_t\right)\mathrm{d}A=0\qquad\forall\mathbf{v}\in\mathbf{V}\end{aligned}\tag{247}$$

This formulation encompasses a family of methods depending on the value of parameter θ. According to [15], this parameter should be chosen from $[-1,1]$ and three choices are of special interest. The choice $\theta=-1$ leads to the symmetric version of the method. For choices $\theta=0$ and $\theta=1$, the non-symmetric versions are obtained. Thorough mathematical analysis of the variants of these methods for the unilateral contact problem is provided in [31].

4) Augmented Lagrangian method (177), (178)

$$\begin{aligned}&\int_\Omega \boldsymbol{\sigma}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{v})\,\mathrm{d}V - \int_\Omega \rho\overline{\mathbf{b}}\cdot\mathbf{v}\,\mathrm{d}V - \int_{\Gamma^\sigma}\overline{\mathbf{t}}\cdot\mathbf{v}\,\mathrm{d}A\\-&\int_{\Gamma^C}\frac{\gamma E}{h}\left\{-g_n(\mathbf{u})-\frac{h}{\gamma E}\lambda\right\}\delta g_n(\mathbf{v})\,\mathrm{d}A=0\qquad\forall\mathbf{v}\in\mathbf{V}\end{aligned}\tag{248}$$

$$-\int_{\Gamma^C}\frac{h}{\gamma E}\left(\lambda\delta\lambda+\frac{\gamma E}{h}\left\{-g_n(\mathbf{u})-\frac{h}{\gamma E}\lambda\right\}\delta\lambda\right)\mathrm{d}A=0\qquad\forall\delta\lambda\in L^2\left(\Gamma^C\right)\tag{249}$$

All of these methods were implemented in FEniCS, and they were tested on various problems concerning contact of a linear elastic body with a rigid plane (called the Signorini problem [57, Chapter 1.1, p.5]). We wish to point out that although the analysed methods for enforcing contact boundary conditions are formulated for the case of contact of two linear elastic bodies in this thesis, the numerical examples are restricted to the contact of one elastic body with a rigid plane. This is because the contact detection is much easier in this case and does not require any additional sophisticated algorithms which importantly influence the efficiency and overall performance of the contact implementation. For this reason, it is important for us to avoid them, as we intended to compare different methods and not the way they interact with contact detection algorithms.

Considering the numerical solution in FEniCS, standard Newton's method was employed, and at every iteration step, sparse LU decomposition (Gaussian elimination) [52] was used as a linear solver. In order to simplify the problem, integration is performed over the original instead of the current configuration of the bodies in contact.

We encountered some difficulties with the convergence of Nitsche-Wriggers. To cast light on this, a simple two-dimensional example of a skewed quadrilateral block was analysed in section 4.2.4. The quadrilateral block was modelled with four nodes and two triangular elements. Linear approximation functions were considered. This problem was solved both analytically with the help of Wolfram Mathematica software and numerically in FEniCS. The results suggest that the function obtained as the left-hand side of the discretised weak form of the Nitsche-Wriggers method is not continuous with respect to the unknown DOFs. In the numerical examples we investigated, the discontinuity was located near the theoretical solution. This led to the fact that the Nitsche-Wriggers method failed to converge in the vicinity of the theoretical solution. The reason is that in every iteration step, standard Newton's method was used for the linearisation of our non-linear problem. However, one of the requirements of guaranteed convergence of Newton's method are the continuous derivatives of the minimized function, which in this case is not satisfied. It could be shown that function obtained as the left-hand side of the discretised weak form of the Nitsche-FPR method is continuous with respect to the unknown DOFs. In the case of the symmetric variant (for $\theta=-1$) this means that the discretised energy functional is the

C^1-differentiable function of the unknown DOFs. For these reasons, the Nitsche-FPR method did not suffer from the convergence problems, if the sufficiently large value of the penalty parameter was provided.

Moreover, we observed that a very careful implementation of Nitsche-Wriggers is necessary, especially of the term containing the Heaviside function which reads $H(-g_n(\mathbf{u}))\delta g_n(\mathbf{v})$. It was found out that this terms causes the method to be very sensitive to a small numerical error in displacements and has to be artificially replaced with $H(\varepsilon_g - g_n(\mathbf{u}))\delta g_n(\mathbf{v})$. The reason for this behaviour is that, as suggested above, Nitsche-Wriggers uses the gap as the contact indicator.

For the Nitsche-Wriggers method, the Nitsche-FPR method, the augmented Lagrangian method and the penalty method implemented in FEniCS, we measured their performance in terms of their accuracy and efficiency. Three following indicators were calculated and compared: the relative error in displacements on an active contact zone of the SNES solution, the relative error in the predicted normal contact tractions on an active contact zone of the SNES solution and the solver runtime. By SNES we mean the non-linear solver SNESVINEWTONRSLS from PETSc library. This is the reduced space active set solver for variational inequalities based on semi-smooth Newton's method. For its high accuracy, it was chosen as a reference method in the aforementioned comparisons.

The numerical results, documented in chapter 4, suggest that the penalty method and the augmented Lagrangian method are very robust in FEniCS. In most of the tested examples (and for most of the considered penalty values) the augmented Lagrangian method provides us with the most precise solution in terms of the indicators we compared (the relative error in displacements and tractions on an active contact zone). Except the 'cross-country ski' example, the error decreased with the increasing penalty value. However, as expected, for the given penalty parameter, the solver runtime was considerably longer than for the other methods. This is caused by the presence of the additional DOFs representing the Lagrange multipliers which are enlarging the solved linear system. The solver runtime for the penalty method and the Nitsche methods was therefore shorter than for the augmented Lagrangian method as there are fewer degrees of freedom. The error produced by the penalty method was the largest in most of the two-dimensional examples. In terms of three-dimensional examples we tested, it turned out that for sufficiently large penalty values, the penalty method provides a relatively precise solution. For small penalty values, the precision of the penalty method was, as expected, always limited (both in the case of two-dimensional and three-dimensional examples). The effect of the ill-conditioning, causing the increase of the solver runtime, with the increasing penalty parameter was observed.

The reference method called SNES (the reduced space active set solver for variational inequalities based on semi-smooth Newton's method) turned out to be relatively effective in terms of the solver runtime. However, on the example of the 'cross-country ski' (which was considerably larger in terms of the number of DOFs than the other problems we tested) it was documented that except for some very large penalty values, the solver runtime of all the methods (except the augmented Lagrangian method) was shorter than the runtime of the SNES method. This suggests that for large-scale problems (which were not tested in this thesis), the SNES method would be too inefficient.

Concerning the Nitsche methods, the investigation of which was our main objective, it turned out that the Nitsche-Wriggers method suffers from severe problems with convergence. These are caused by the discontinuity of the discretised Nitsche-Wriggers weak form located near the theoretical solution of the problem. If converged, Nitsche-Wriggers provides us mostly with more precise solution than the penalty method and the solver runtime is shorter than the one of the augmented Lagrangian method. However, for the examples we tested, the Nitsche-Wriggers method converged only rarely.

Compared to Nitsche-Wriggers, Nitche-FPR proved itself to be a very robust and stable method in FEniCS, if sufficiently large penalty value was provided. In contrast to Nitsche-Wriggers, the Nitsche-FPR method uses the sign of $g_n + \frac{h}{\gamma E}\sigma_n$, i.e., of the gap plus some multiple of the traction normal to the contact surface, which turned out to be more stable in terms of the small numerical error than the indicator based exclusively on the gap, which is used by Nitsche-Wriggers. The negative aspect of Nitsche-FPR is that at a discrete level, it leads to a certain reduction of the stiffness in the inactive part of the potential contact zone. In the case of two-dimensional examples, the produced errors in displacements and contact tractions were mostly smaller than the error of the penalty method (for the same value of the penalty parameter). At the same time, the solver runtime was considerably shorter than for the augmented Lagrangian method. In the case of three-dimensional examples, for large penalty values, the accuracy of the Nitsche-FPR method was relatively similar to the penalty method. For small penalty values, Nitsche-FPR ($\theta = 1$) produces a smaller error in displacements than the penalty method.

We would like to point out that in the case of Nitsche-FPR ($\theta = -1$) and Nitsche-FPR ($\theta = 0$), it is necessary to choose certain minimal value of the penalty parameter so that the discrete problem could be well-posed. However, Nitsche-FPR ($\theta = 1$) is well-posed for every positive value of the penalty parameter. In terms of the numerical experiments we conducted, Nitsche-FPR ($\theta = 1$) always converged. An important drawback of Nitsche-FPR ($\theta = 1$) and Nitsche-FPR ($\theta = 0$) is that the discrete problem becomes non-symmetric even if the original problem (without considering contact) is symmetric. This does not allow us to use some efficient solvers. However, in the case of many more complicated problems, the problem could be non-symmetric even without considering contact and Nitsche-FPR ($\theta = 1$) could become very competitive in this case.

Overall, the results of this thesis suggest that despite being overlooked for many years, Nitsche methods (meaning the Nitsche-FPR method) provide a very interesting alternative to standard penalty and mixed methods. They circumvent classic drawbacks of both approaches, as they are consistent and they require a much smaller value of the penalty parameter to achieve the same accuracy as the penalty method. The problem of ill-conditioning of the discrete system, resulting from using high penalty values, thus could be avoided. At the same time, no additional unknown (a Lagrange multiplier) is introduced, so the discrete system (and the computational cost) is not enlarged, and it is not necessary to worry about fulfilling the Babuška-Brezzi condition. These properties make Nitsche methods potentially very attractive, especially for large-scale problems. Also, we would like to point out that non-symmetric variants of Nitsche-FPR should not be ignored, as they seem to be very robust and sometimes providing even better results than their symmetric counterparts. Despite the fact that we focused our attention on FEniCS in this thesis, the results are not limited to this platform and could be utilized even for implementations of the mentioned methods in different finite element based software packages.

There are many directions in which the further research of the topics we covered could proceed. Firstly, from the technical point of view, we did not devote any attention to the implementation of the problem of two elastic bodies to FEniCS, as we concentrated only on unilateral contact considering the numerical examples. However, as we give the relevant weak forms, it is an interesting question how to efficiently implement these methods and more advanced contact detection algorithms in FEniCS. Also, as we considered only linear elastic behaviour of the material (and integrated over the undeformed configuration during numerical solution), it would be very interesting to extend the proposed methods to the case of large deformation or other material models. Last but not least, the master-slave approach can be omitted entirely and replaced with an unbiased approach, as suggested in [41], [32].

References

[1] Annavarapu, C., Hautefeuille, M., Dolbow, J.E., 2012. A robust Nitsche's formulation for interface problems. Computer Methods in Applied Mechanics and Engineering 225–228, 44–54. https://doi.org/10.1016/j.cma.2012.03.008

[2] Amann, H., Helfrich, H.-P., Scholz, R., 1997. Joachim A. Nitsche (1926-1996). Jahresbericht der Deutschen Mathematikervereinigung 99, 90-100.

[3] Apostolatos, A., Schmidt, R., Wüchner, R., Bletzinger, K.-U., 2014. A Nitsche-type formulation and comparison of the most common domain decomposition methods in isogeometric analysis. International Journal for Numerical Methods in Engineering 97, 473–504. https://doi.org/10.1002/nme.4568

[4] Arnold, D., Douglas, N., 1982. An Interior Penalty Finite Element Method with Discontinuous Elements. SIAM Journal on Numerical Analysis 19. https://doi.org/10.1137/0719052

[5] Babuška, I., 1973. The Finite Element Method with Penalty. Mathematics of Computation 27, 221–228. https://doi.org/10.2307/2005611

[6] Barbosa, H.J.C., Hughes, T.J.R., 1992. Circumventing the Babuška-Brezzi condition in mixed finite element approximations of elliptic variational inequalities. Computer Methods in Applied Mechanics and Engineering 97, 193–210. https://doi.org/10.1016/0045-7825(92)90163-E

[7] Becker R., Hansbo P., 1999. A finite element method for domain decomposition with non-matching grids. Technical Report No.3613, INRIA, Sophia Antipolis

[8] Becker, R., Hansbo, P., Stenberg, R., 2003. A finite element method for domain decomposition with non-matching grids. ESAIM: Mathematical Modelling and Numerical Analysis 37, 209–225. https://doi.org/10.1051/m2an:2003023

[9] Benson, S.J., Munson, T.S., 2006. Flexible complementarity solvers for large-scale applications. Optimization Methods and Software 21, 155–168. https://doi.org/10.1080/10556780500065382

[10] Boffi, D., Brezzi, F., Fortin, M., 2013. Mixed Finite Element Methods and Applications, Springer Series in Computational Mathematics. Springer-Verlag, Berlin Heidelberg.

[11] Bristol, U. of, n.d. September: ASiMoV Prosperity Partnership — News and features — University of Bristol [WWW Document]. URL https://www.bristol.ac.uk/news/2018/september/asimov-prosperity-partnership.html (accessed 4.14.20).

[12] Burman, E., Hansbo, P., Larson, M., 2016. Augmented Lagrangian finite element methods for contact problems. ESAIM: Mathematical Modelling and Numerical Analysis 53. https://doi.org/10.1051/m2an/2018047

[13] Burman, E., Hansbo, P., Larson, M.G., Stenberg, R., 2017. Galerkin least squares finite element method for the obstacle problem. Computer Methods in Applied Mechanics and Engineering 313, 362–374. https://doi.org/10.1016/j.cma.2016.09.025

[14] Duvaut G., Lions J.L., 1971. Elasticité avec frottement, Journal de Mécanique, vol. 10, pp. 409–420.

[15] Fabre, M., Pousin, J., Renard, Y., 2014. A fictitious domain method for frictionless contact problems in elasticity using Nitsche's method. SMAI Journal of Computational Mathematics 2. https://doi.org/10.5802/smai-jcm.8

[16] FEniCS Project [WWW Document], n.d. . FEniCS Project. URL https://fenicsproject.org/ (accessed 4.21.20).

[17] Fichera, G., 1995. La nascita della teoria delle disequazioni variazionali ricordata dopo trent'anni. Incontro scientifico italo-spagnolo. Roma, 21 ottobre 1993, Atti dei Convegni Lincei (in Italian), 114, Roma: Accademia Nazionale dei Lincei, pp. 47–53.

[18] Fletcher R., 1970. A Class of Methods for Nonlinear Programming with Termination and Convergence Properties. Integer and Nonlinear Programming, Amsterdam, 1970, pp. 157-173.

[19] Galántai, A., 2012. Properties and construction of NCP functions. Comput Optim Appl 52, 805–824. https://doi.org/10.1007/s10589-011-9428-9

[20] Georgoulis, E.H., 2011. Discontinuous Galerkin Methods for Linear Problems: An Introduction, in: Georgoulis, E.H., Iske, A., Levesley, J. (Eds.), Approximation Algorithms for Complex Systems, Springer Proceedings in Mathematics. Springer Berlin Heidelberg, pp. 91–126. https://doi.org/10.1007/978-3-642-16876-5_5

[21] Glowinski, R., Tallec, P.L., 1987. Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics. Society for Industrial and Applied Mathematics, Philadelphia.

[22] Hansbo, P., Larson, M.G., 2002. Discontinuous Galerkin methods for incompressible and nearly incompressible elasticity by Nitsche's method. Computer Methods in Applied Mechanics and Engineering 191, 1895–1908. https://doi.org/10.1016/S0045-7825(01)00358-9

[23] Hansbo, P., Rashid, A., Salomonsson, K., 2016. Least-squares stabilized augmented Lagrangian multiplier method for elastic contact. Finite Elements in Analysis and Design 116, 32–37. https://doi.org/10.1016/j.finel.2016.03.005

[24] Hansbo, P., 2005. Nitsche's method for interface problems in computational mechanics. GAMM-Mitteilungen 28. https://doi.org/10.1002/gamm.201490018

[25] Hansbo, A., Hansbo, P., 2004. A finite element method for the simulation of strong and weak discontinuities in solid mechanics. Computer Methods in Applied Mechanics and Engineering 193, 3523–3540. https://doi.org/10.1016/j.cma.2003.12.041

[26] Heintz, P., Hansbo, P., 2006. Stabilized Lagrange multiplier methods for bilateral elastic contact with friction. Computer Methods in Applied Mechanics and Engineering 195, 4323–4333. https://doi.org/10.1016/j.cma.2005.09.008

[27] Hertz, H., 1882. Ueber die Berührung fester elastischer Körper. Journal für die reine und angewandte Mathematik 1882, 156–171. https://doi.org/10.1515/crll.1882.92.156

[28] Hestenes, M.R., 1969. Multiplier and gradient methods. J Optim Theory Appl 4, 303–320. https://doi.org/10.1007/BF00927673

[29] Chouly, F., Hild, P., 2013. A Nitsche-Based Method for Unilateral Contact Problems: Numerical Analysis. SIAM J. Numer. Anal. 51, 1295–1307. https://doi.org/10.1137/12088344X

[30] Chouly, F., Hild, P., Renard, Y., 2014. Symmetric and non-symmetric variants of Nitsche's method for contact problems in elasticity: Theory and numerical experiments. Mathematics of Computation 84, 1089–1112. https://doi.org/10.1090/S0025-5718-2014-02913-X

[31] Chouly, F., Fabre, M., Hild, P., Mlika, R., Pousin, J., Renard, Y., 2017. An Overview of Recent Results on Nitsche's Method for Contact Problems, in: Bordas, S.P.A., Burman, E., Larson, M.G., Olshanskii, M.A. (Eds.), Geometrically Unfitted Finite Element Methods and Applications, Lecture Notes in Computational Science and Engineering. Springer International Publishing, pp. 93–141.

[32] Chouly, F., Mlika, R., Renard, Y., 2018. An unbiased Nitsche's approximation of the frictional contact between two elastic structures. Numer. Math. 139, 593–631. https://doi.org/10.1007/s00211-018-0950-x

[33] Johnson, K.L., 1985. Contact mechanics. Cambridge University Press, Cambridge.

[34] Kikuchi, N., Oden, J.T., 1987. Contact Problems in Elasticity: A Study of Variational Inequalities and Finite Element Methods. Society for Industrial and Applied Mathematics, Philadelphia, Pa.

[35] Konyukhov, A., Izi, R., 2015. Introduction to Computational Contact Mechanics: A Geometrical Approach, 1 edition. ed. Wiley, Chichester.

[36] Logg, A., Mardal, K.-A., Wells, G. (Eds.), 2012. Automated Solution of Differential Equations by the Finite Element Method: The FEniCS Book, Lecture Notes in Computational Science and Engineering. Springer-Verlag, Berlin Heidelberg.

[37] Lu, K., Augarde, C.E., Coombs, W.M., Hu, Z., 2019. Weak impositions of Dirichlet boundary conditions in solid mechanics: A critique of current approaches and extension to partially prescribed boundaries. Computer Methods in Applied Mechanics and Engineering 348, 632–659. https://doi.org/10.1016/j.cma.2019.01.035

[38] Luenberger, D.G., Ye, Y., 2015. Linear and Nonlinear Programming, 4th ed. 2016 edition. ed. Springer, Cham.

[39] Lumping a mass matrix — Numerical tours of continuum mechanics using FEniCS master documentation [WWW Document], n.d. URL https://comet-fenics.readthedocs.io/en/latest/demo/tips_and_tricks/mass_lumping.html (accessed 4.21.20).

[40] Middleton J., Pange G. (eds), 1985. Penalty and Augmented Lagrangian Formulation for Contact Problems, Elsevier

[41] Mlika, R., Renard, Y., Chouly, F., 2017. An unbiased Nitsche's formulation of large deformation frictional contact and self-contact. Computer Methods in Applied Mechanics and Engineering 325, 265–288. https://doi.org/10.1016/j.cma.2017.07.015

[42] multiphenics — mathLab innovating with mathematics [WWW Document], n.d. URL https://mathlab.sissa.it/multiphenics (accessed 4.22.20).

[43] New Prosperity Partnership to develop world first in high-fidelity engineering simulations — EPCC at The University of Edinburgh [WWW Document], n.d. URL https://www.epcc.ed.ac.uk/blog/2018/11/21/new-prosperity-partnership-develop-world-first-high-fidelity -engineering-simulations (accessed 4.14.20).

[44] Nitsche, J., 1971. Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind. Abh.Math.Semin.Univ.Hambg. 36, 9–15. https://doi.org/10.1007/BF02995904

[45] Powell M.J.D., 1969. A Method for Nonlinear Constraints in Minimization Problems. In: Fletcher, R., Ed., Optimization, Academic Press, New York, NY, 283-298.

[46] Rockafellar, R.T., 1973. A dual approach to solving nonlinear programming problems by unconstrained optimization. Mathematical Programming 5, 354–373. https://doi.org/10.1007/BF01580138

[47] Rockafellar, R.T., 1973. The multiplier method of Hestenes and Powell applied to convex programming. J Optim Theory Appl 12, 555–562. https://doi.org/10.1007/BF00934777

[48] Ruess, M., Schillinger, D., Özcan, A.I., Rank, E., 2014. Weak coupling for isogeometric analysis of non-matching and trimmed multi-patch geometries. Computer Methods in Applied Mechanics and Engineering 269, 46–71. https://doi.org/10.1016/j.cma.2013.10.009

[49] Signorini A., 1933. Sopra alcune questioni di elastostatica. Atti della Societa Italiana per Il Progresso delle Scienze.

[50] SNESVINEWTONRSLS [WWW Document], n.d. URL https://www.mcs.anl.gov/petsc/petsc-current/docs/manualpages/SNES/SNESVINEWTONRSLS.html (accessed 4.14.20).

[51] Signorini problem [WWW Document], n.d. URL https://en.wikipedia.org/wiki/Signorini_problem#CITEREFFichera1995 (accessed 4.14.20).

[52] Solving PDEs in Python - The FEniCS Tutorial Volume I [WWW Document], n.d. URL https://fenicsproject.org/pub/tutorial/html/._ftut1017.html (accessed 4.14.20).

[53] Stenberg, R., 1995. On some techniques for approximating boundary conditions in the finite element method. Journal of Computational and Applied Mathematics, Proceedings of the International Symposium on Mathematical Modelling and Computational Methods Modelling 94 63, 139–148. https://doi.org/10.1016/0377-0427(95)00057-7

[54] UFL — Unified Form Language (UFL) 2019.2.0.dev0 documentation [WWW Document], n.d. URL https://fenics.readthedocs.io/projects/ufl/en/latest/ (accessed 4.21.20).

[55] Wriggers, P., 2006. Computational Contact Mechanics, 2nd ed. Springer-Verlag, Berlin Heidelberg.

[56] Wriggers, P., Zavarise, G., 2008. A formulation for frictionless contact problems using a weak form introduced by Nitsche. CompM 41, 407–420. https://doi.org/10.1007/s00466-007-0196-4

[57] Yastrebov, V.A., 2013. Numerical Methods in Contact Mechanics, 1 edition. ed. Wiley-ISTE, London.