Finite element modeling of dry frictional contact.

Yuan. Yao
University of Windsor

Follow this and additional works at: https://scholar.uwindsor.ca/etd

Recommended Citation
https://scholar.uwindsor.ca/etd/2109

This online database contains the full-text of PhD dissertations and Masters' theses of University of Windsor students from 1954 forward. These documents are made available for personal study and research purposes only, in accordance with the Canadian Copyright Act and the Creative Commons license—CC BY-NC-ND (Attribution, Non-Commercial, No Derivative Works). Under this license, works must always be attributed to the copyright holder (original author), cannot be used for any commercial purposes, and may not be altered. Any other use would require the permission of the copyright holder. Students may inquire about withdrawing their dissertation and/or thesis from this database. For additional inquiries, please contact the repository administrator via email (scholarship@uwindsor.ca) or by telephone at 519-253-3000ext. 3208.
INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600
FINITE ELEMENT MODELING OF DRY FRICTIONAL CONTACT

by

Yuan Yao

A Dissertation
Submitted to the Faculty of Graduate Studies and Research
through the Department of Physics
in Partial Fulfillment of the Requirements for
the Degree of Doctor of Philosophy at the
University of Windsor

Windsor, Ontario, Canada
January 2002
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author’s permission.

L’auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L’auteur conserve la propriété du droit d’auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-67678-1
Abstract

The Contact Problem is of great practical significance. The nature of contact and friction, however, is so complex that the science of contact interaction is not sufficiently advanced for assisting in real engineering problems. One of the difficulties with contact interaction is the nature of surface friction, another is the nonlinear character arising from the free boundary condition. The computation of the exact stress and deformation around the contact boundary region is vital for engineering analysis and design, but obtaining exact solutions based on the theory of elasticity has been a challenge for physicists and mathematicians since the end of nineteenth century. We have developed computing models to investigate static and quasi-static smooth and frictional contact between solid bodies with various two dimensional geometries. A variational inequality approach with penalty and multiplier optimization methods is used to study the contact problem. Friction is modelled according to the classical Coulomb friction law. To overcome the problem of the relative shift between particles on one surface and particles on the other surface, we have designed a scheme to solve each contact solid one by one iteratively. The solutions for the two elastic solid bodies in frictional contact are connected through the surface traction and surface deformation. This scheme has a good convergence rate. One new aspect of our approach to the frictional interaction problem is the application of cubic splines in approximating the contact surface. We also address the difference between Cauchy and Piola–Kirchoff stress, and show when it is significant. A numerical investigation of the stress dependence on the loading distribution and solid geometry is conducted. The stress distribution deviates from predictions of Hertz theory and subsequent research, and it is sensitive to the loading distribution and the geometrical shape of the contact solids. We therefore argue that an accurate analysis of the dry contact problem requires a more refined knowledge about loading conditions and the geometry of both solids.
Acknowledgments

The Author expresses his appreciation to Dr. M. Schlesinger for his direct guidance and sponsoring throughout this project. It is his scientific insight that brought my attention and curiosity toward this complex problem. I also thank Dr R. Kent, Dr W. E. Baylis, and Dr G. W. Drake for their helpful suggestions regarding computation experience and software. I also would like to thank our lab fellow Dr. L. F. McAven and E. Ibrahim for their valuable suggestions and discussions. Special thanks is given to Dr. L. F. McAven for helping me with correcting this thesis and creating Latex documents. Many of graduate students in our department gave me valuable help in the course of this project. among whom are A. Abdellatif, M. Catalin, M. Cassar, D. Keselica, H. Terek, H. Tareq, B. O'Neill, S. Hadi and G Trayling. Without their help and encouragement, I would have had much more difficulty finishing this project.
Contents

Abstract iv

Acknowledgments v

List of Figures viii

List of Tables x

Glossary of Symbols xi

1 Historic Review 1

1.1 Introductory Remarks and Research Motivation 1

1.2 Hertz Contact Problem 6

1.3 The Signorini Contact Problem 13

1.4 The Variational Approach to Numerical Modeling 17

2 Theoretical Introduction 21

2.1 An Introduction to Finite Element Methods and Variational Inequalities 21

2.2 Cauchy Stress and Piola–Kirchoff Stress 24

2.3 Theoretical Formulation of the Contact Problem 26

2.3.1 Frictionless Contact between Elastic Solid and Rigid Surface 27

2.3.2 Contact between a Rigid Body and an Elastic Foundation without Friction 36

2.3.3 Frictionless Contact between Two Elastic Solid Bodies 37

2.3.4 Contact Between an Elastic Solid and a Rigid Surface with Friction 39

2.3.5 Contact between Elastic Solids with Friction 51
List of Figures

1.1 Surface of contact between two elastic bodies ........................................ 8
1.2 Enlargement of Figure (1.1) in the contact region .................................. 9
1.3 Contact between elastic body and rigid body. ......................................... 14
1.4 Contact surface .................................................................................... 18

2.1 Shape function. ................................................................................... 23
2.2 The base function. .............................................................................. 31
2.3 The contact boundary condition in finite elements ................................. 35
2.4 Self-healing crack and lattice grid. ...................................................... 42
2.5 Transformation from global coordinates to natural coordinates. .......... 50

3.1 Flowchart of the algorithm for elastic frictional contact. ...................... 64

4.1 Geometry and loading of the first example. ......................................... 67
4.2 First example: Normal pressure. comparison with Hertz theory. .......... 68
4.3 First example: Stress under the surface. comparison with Hertz theory. . 69
4.4 Geometry and loading of the second example. ..................................... 70
4.5 Stress under the surface: First example versus second example. ............ 71
4.6 Displacement $v_z$: First example versus second example. .................. 71
4.7 Contact pressure: Loading condition I versus loading condition II. ......... 73
4.8 Displacement $v_z$: Loading condition I versus loading condition II. ...... 73
4.9 Geometry and loading of the third example. ........................................ 74
4.10 The mesh of the third example. ........................................................... 75
4.11 Surface traction of frictional contact for different loading histories. ......... 75
4.12 The final stress underneath the contact surface for the third example. ..... 76
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.13 Geometry and loading of the fourth example</td>
<td>77</td>
</tr>
<tr>
<td>4.14 Contact stress for the fourth example</td>
<td>79</td>
</tr>
<tr>
<td>4.15 Surface shape: Third example versus fourth example</td>
<td>79</td>
</tr>
<tr>
<td>4.16 Fourth example: Stress under the surface, comparison with Hertz theory</td>
<td>80</td>
</tr>
<tr>
<td>4.17 Displacement $v_x$: Frictionless versus frictional for equal loading</td>
<td>81</td>
</tr>
<tr>
<td>4.18 Displacement $v_x$: Frictionless versus frictional for equal loading</td>
<td>82</td>
</tr>
<tr>
<td>4.19 Displacement $v_x$: Different loadings</td>
<td>82</td>
</tr>
<tr>
<td>4.20 Schematic diagram of Brémond and Durelli’s experiment</td>
<td>84</td>
</tr>
<tr>
<td>4.21 The deformation of the finite element grid</td>
<td>85</td>
</tr>
<tr>
<td>4.22 Displacement of the contact surface</td>
<td>86</td>
</tr>
<tr>
<td>4.23 Numerically calculated contact stress</td>
<td>87</td>
</tr>
<tr>
<td>4.24 Stress under the surface</td>
<td>89</td>
</tr>
<tr>
<td>A.1 Points used in the Gauss quadrature for triangular elements</td>
<td>96</td>
</tr>
</tbody>
</table>
List of Tables

4.1  The estimated computation time for different number of nodes per body. . . 88

A.1  The Gauss quadrature for triangular element. . . . . . . . . . . . . . . . . . . . 95
Glossary of Symbols

\[ \beta_k = \frac{[\nabla f(x_k)]^{T}[\nabla f(x_k)]}{[\nabla f(x_{k-1})]^{T}[\nabla f(x_{k-1})]} \]

\[ \Gamma_{di} \quad \text{The boundary where the displacement is prescribed in } \Omega_i, i = 1, 2. \]

\[ \Gamma_{Fi} \quad \text{The boundary where the traction force is prescribed } \Omega_i, i = 1, 2. \]

\[ \Gamma_{ci} \quad \text{The boundary where contact can happen by certain loading } \Omega_i, i = 1, 2. \]

\[ \varepsilon_{ij} \quad \text{Strain tensor.} \]

\[ \varepsilon_x \quad \varepsilon_{xx}. \]

\[ \varepsilon_y \quad \varepsilon_{yy}. \]

\[ \varepsilon_{xy} \quad \text{Shear strain.} \]

\[ \lambda \quad \text{Lagrangian Multiplier.} \]

\[ \lambda^* \quad \text{Lagrangian Multiplier associated with the strict local minimum } x^*. \]

\[ \lambda_k \quad \text{Multiplier associated with augmented Lagrangian function at iteration } k. \]

\[ \tilde{\lambda} \quad \text{the optimizer of quadratic function through } (a, g(a)), (b, g(b)), (c, g(c)). \]

\[ \lambda_L, \mu_L \quad \text{Lamé coefficients } \lambda_L = \frac{E\nu}{(1-\nu)(1-2\nu)}; \quad \mu_L = \frac{E}{2(1-\nu)}. \]

\[ \theta_{(ac)} \quad \frac{x_{m1(ac)} - x_{ac}}{x_{m2(ac)} - x_{m1(ac)}}. \]

\[ \rho_0 \quad \text{Density before deformation.} \]

\[ \rho \quad \text{Density after deformation.} \]

\[ \sigma \quad \text{Cauchy stress tensor.} \]

\[ \tilde{\sigma} \quad \text{the second Piola–Kirchoff stress tensor.} \]

\[ \sigma_{ij} \quad \text{component form of Cauchy stress tensor.} \]

\[ \tilde{\sigma}_{ij} \quad \text{component form of the second Piola–Kirchoff stress tensor, some authors use } \tilde{T}_{ij}. \]

\[ \sigma_x \quad \text{Abbreviation for } \sigma_{xx}, \text{ the element } \sigma_{11} \text{ of } \sigma_{ij}. \]

\[ \sigma_y \quad \text{Abbreviation for } \sigma_{yy}, \text{ the element } \sigma_{22} \text{ of } \sigma_{ij}. \]

\[ \tau_1 \quad \text{Principal shear stress.} \]

\[ \sigma_n \quad \text{The magnitude of normal traction of the contact surface.} \]
\( \sigma_t \) The magnitude of tangential traction of the contact surface.

\( \sigma_t \) Tangential traction on the surface.

\( \sigma_n \) Normal traction on the surface.

\( \phi(x, y) \) Represents the contact surface of solid I before loading.

\( \phi'(x, y) \) Represents the contact surface of solid I after loading.

\( \phi_{\alpha}^{(e)} \) The base function \( \phi_{\alpha} \) evaluated at element \( e \).

\( \phi_{\alpha,l}^{(e)} \) The partial derivative \( \frac{\partial \phi_{\alpha}^{(e)}}{\partial x_l} \).

\( \phi_{\alpha}(x) \) Base function defined on \( \Omega_h \) satisfying \( \phi_{\alpha}(x_3) = \delta_{\alpha,3} \).

\( \phi(u) \) Regularization function of \( |u| \).

\( \psi(x, y) \) Represents the contact surface of solid II before deformation.

\( \psi'(x, y) \) Represents the contact surface of solid II deformation.

\( \Omega \) The subset of \( \mathbb{R}^3 \) representing the contact solid.

\( \Omega_1 \) The subset of \( \mathbb{R}^3 \) representing the contact solid I.

\( \Omega_2 \) The subset of \( \mathbb{R}^3 \) representing the contact solid II.

\( \Omega_h \) The finite element computational domain corresponding to \( \Omega \).

\( a, b \) Semi-major and semi-minor of the contact ellipse.

\( a(u, v) \int_{\Omega} \sigma_{ij}(u)\varepsilon_{ij}(u)dx. \)

\( A, B \) Quadratic term coefficients in \( z_1 + z_2 = Ax^2 + By^2 \).

\( d_k \) Conjugate searching direction.

\( d_y \) Indentation displacement in the \( y \) direction.

\( E \) Young’s Modulus of the material in solid I.

\( E' \) Young’s Modulus of the material in solid II.

\( E_{ijkl} \) Elements of the stiffness tensor.

\( f_i \) Body force density. \( i=1,2,3 \).

\( f(v) \int_{\Omega} f_i u_i + \int_{\Gamma_p} t_i u_i ds. \)

\( f_{\alpha} \) Elements of the force matrix \( \int_{\Omega} f_i \phi_{\alpha} + \int_{\Gamma_p} t_i \phi_{\alpha} ds. \)

\( F^{-1} \) \( x\vec{\nabla} \vec{x}' \).

\( F \) Total pressure force.

\( F_y \) Total pressure force for the rigid pouch problem.
The document contains a page of mathematical expressions and definitions. Here are the key terms and concepts:

- $g_{ac}$: \max\{0, g_{\alpha}(v)\}
- $g_n(x)$: \frac{\psi(x,y) - \phi(x,y)}{\sqrt{1 + \left(\frac{\partial \psi}{\partial z} - \frac{\partial \phi}{\partial z}\right)^2}}$
- $x^\alpha$: the $x$ component of the coordinate of node $\alpha$.
- $g_n(x)$: The gap at $x$ before loading.
- $g_0(x)$: $g_n(x_{\alpha^i})$.
- $g_j$: A general inequality constraint $g_j \leq 0$.
- $g_j^+(x, \mu, c)$: \max\{g_j(x), -\frac{\mu}{\varepsilon}\}.
- $h_i$: $x_{i+1} - x_i$.
- $h$: The approach of two contact solid bodies toward each other.
- $k_{ab}$, $k_{ab}^\prime$: Quadratic term coefficients in the functions $z_1$ and $z_2$.
- $K$: \{v \in V|v_n(x) - g_n(x) \leq 0 \text{ on } \Gamma_c\}.
- $J(v, u)$: $J(u, v) = -\int_{\Gamma_c} \mu(\sigma_n \cdot n) |v_t| \, ds_b$.
- $J$: Jacobian matrix. Transforms natural to global coordinates.
- $x_{m_1(\alpha)}$: $\tilde{x}$ coordinates of two neighboring contact nodes of solid II.
- $J_s$: $\frac{\partial}{\partial \rho}$.
- $L_c$: Augmented Lagrangian function.
- $M_i$: $s''(x_i)$.
- $\tilde{N}$: The normal unit vector of the surface before deformation.
- $n$: $n_1, n_2, n_3$. Normal unit vector of the surface $\psi(x, y)$.
- $N_x$: Shape function corresponding to node $x$ of an element.
- $N_e(\alpha)$: The shape function corresponding to node $\alpha$ of element $e$.
- $N_{\alpha^i}^1, N_{\alpha^i}^2$: Components of unit vector normal to the contact.
- $P(x, y)$: The surface pressure as a function of surface position.
- $(x', y')$: The Cartesian coordinates of a particle after deformation.
- $P_{\text{max}}$: Maximum Hertz pressure.
$P(v)$  Penalty function.

$R, R'$  Radii of the two contact spheres.

$r$  $r(x, y, z)$ Position vector of a field point.

$R_1, R_2$  Radii of the cross section area of the two contact cylinders.

$s$  $-\mu(\sigma_n \cdot n)$.

$s(a)$  One of the node numbers $i, j, k, l, m, n$.

$s(x)$  Spline function.

$t_i$  Prescribed traction force on the boundary $\Gamma_F$.

$T_{\alpha}^{i}$  Tensor elements: $\int_{\Gamma_F^i} t_i \phi_\alpha ds$.

$u_j$  $z_j^2$.

$u_{k,l}$  The partial derivative $\frac{\partial u_k}{\partial x_l}$.

$u, v, w$  Components of the displacement vector $u$.

$U_i$  Prescribed displacement constraint on the boundary $\Gamma_d$.

$\dot{u}_j$  $-\left[\left(\frac{\partial u_j}{\partial x} + g_j(x)\right]\right.$.

$u_j^*$  $\max\{0, -\left[\left(\frac{\partial u_j}{\partial x} + g_j(x)\right]\right.$.

$v, \tilde{v}$  Displacement field vector of solid I and solid II.

$v$  Displacement vector defined on $\Omega$.

$(v_h)_i$  Displacement vector defined on $\hat{\Omega_h}$.

$u_t$  Tangential component of displacement vector.

$v_{zp}, v_{yp}$  Components of displacement vector at $p$ on the contact boundary side.

$V$  Normed linear space of real functions on $\Omega$.

$x, y, z$  Cartesian coordinates of an initial particle position $x(x, y, z)$.

$x_\alpha$  The location vector of node $\alpha$.

$x_{\alpha c}$  $x$ coordinate of the contact boundary node $\alpha_c$ of solid I.

$z_1$  Function to represent the contact surface of solid I $z_1(x, y)$.

$z_2$  Function to represent the contact surface of solid II $z_2(x, y)$.

$z_j$  A variable introduced to make $g_j(x) + z_j^2 = 0$. 

xv
Chapter 1

Historic Review

1.1 Introductory Remarks and Research Motivation

Contact and friction exist everywhere: almost every mechanical component is exposed to contact condition. It is the deformable contact surface that transfers the mechanical force, momentum and energy between components. Where there are mechanical contacts, there must be surface traction forces. Surface traction forces are usually classified into two parts: normal and tangential traction. Normal traction is usually called pressure, and tangential traction is usually called friction. These two fundamental force concepts were introduced into physics during the early development of classical mechanics. But the nature of these two forces and their relationship are still not fully understood.

Since most mechanical loadings result from such tractions, frictional contact problems are very important in engineering analysis. The behavior of elastic bodies in contact is critical for many areas of mechanics, such as balls and roller bearings, gears and connecting elements.

The widely used Coulomb friction law asserts that the relative sliding of two solid bodies in contact will occur when the net tangential force reaches a critical value proportional to the normal force pressing the two bodies together. This simple idea was presented by Amontons in 1699 [1] and extended by Coulomb in 1781 [2]. It was determined empirically from observations of macroscopic mechanical phenomena with the assumption that solid bodies are rigid. However whether or not this frictional law works for the pointwise situation or in atomic scale is still an open problem, especially when the contact surface
traction is considered in the language of the theory of elasticity. Normal traction, tangential traction and the surface deformation are all important elements affecting the behavior of mechanical contacts. Strictly speaking they cannot be considered independent of each other. The need for knowledge about the stress around the contact arises from engineering applications. It is known that the stress is much higher around the contact surface region than in the rest of the solid. The character of this small contact surface plays an important role in the motion of the solid body and has a direct impact on properties of the material. Contact stress in the contact region forms one of the main reasons for wear and surface fatigue of the material. However, the nature of contact pressure and friction is far more complex than what classical mechanics describes. It cannot just be simply modeled as a vector with a magnitude and a direction, as used in the early development of mechanics. Although it did approximate the gross motion of the solid as a rigid body very well, it is not sufficient to depict the exact situation in the material. It is known that surface fatigue and the initialization and the development of cracking are important phenomena in engineering practice, and they are closely dependent on the stress in the contact region. In order to guarantee the safety and quality of a mechanical structure, it becomes necessary to study stress conditions in the contact area.

But contact effects are rarely taken into account in normal structural analysis because the analysis of contact phenomena poses a theoretical difficulty in its inherent nonlinear nature. This nonlinearity arises from the free boundary condition. This problem was brought to the attention of physicists and mathematicians in the late 19th century by Hertz [3]. He solved the nonlinear frictionless contact problem analytically using linear elastic theory and the contemporary continuum mechanics. Hertz solved the pressure distribution of two smooth curved surfaces by Newtonian potential theory. Considerable theoretical research to investigate the stress within the solid based on the Hertz pressure distribution has been conducted since then. However, a lengthy integration process is necessary before any explicit analytical results can be obtained. The two-dimensional contact problem was solved analytically by McEwen [4] who found an expression for the stress at a general field point under study. Later, Poritsky and Schnectady [5] computed the stress and deflection for the contact of cylindrical bodies. Smith and Liu [6] calculated the 2D stress of tangential and normal loading acting together with a Hertz distribution. The
stress value for a three-dimensional system along a symmetry axis, under Hertz pressure, was calculated by Thomas Hoersch [7] and later by Lundberg and Sjöval [8]. Fessler and Ollerton [9] later measured the principal shear stress in a symmetry plane using the frozen stress method of photoelasticity. The measured contact size was somewhat greater than what the theory predicted for the smallest value $a/R$, where $a$ is the semi major axis of the contact area ellipse and $R$ is the principal radius of surface curvature of the contact surface. However there was good agreement with theoretical predictions for both contact area and internal stress at high load up to $a/R = 0.3$. It was not until 1980 that Sackfield and Hills [10] calculated the stress of a general field point. Hanson and Johnson [11] calculated the stress and displacement for a spherical Hertzian contact. Recently an analytical solution of a dynamic frictional contact problem was calculated by Zhariii [12], but the normal traction was also directly determined from the results of Hertz Theory.

In general, Hertz contact theory is the basis of all those calculations of contact stresses, and has formed the main direction of analytical research into contact problems during the last century. These results are still widely used by today's engineers to estimate the contact stress in appropriate cases. However, all these calculations assume that the tangential traction is proportional to the normal pressure, as prescribed by the Hertzian distribution. The reality is that the normal traction, the tangential traction and the surface deformation are dependent on each other and that the contact surface is a freely moving boundary. This makes the real normal and tangential traction deviate from Hertz results, eventually affecting the stress underneath the contact surface. All the above mentioned analytical calculations fail to address this problem. This shortcoming is the main reason people seek help from computational modeling.

Computational studies of the contact problem have become more common in the last few decades due to the rapid progress of computing technology. Generally speaking computational methods for solving contact problems can be classified by means of numerical implementation into three categories: finite difference methods, finite element methods and boundary element methods. Finite difference methods and boundary element methods are not discussed in this thesis.

The implementation of finite element methods forms a major part of the research literature because the unstructured grids used by finite element methods have more flexibility
than finite difference method in solving a situation with irregular geometries.

There are two main formulations of finite elements methods to solve frictional contact problems: The first formulation is named as variational equality formulation (VE). It makes use of virtual work equation. The variational equality approach has drawn considerable attention, especially from engineering researchers, as it leads to a convenient framework that can be incorporated into existing commercial finite element programs (ANSYS, ABAQUS, ADINA, LS-DYNA3D). Many strategies have been developed in this area that have enabled the solution of a wide range of practical frictional contact problems [14, 15]. Although the variational equality approach is successful in solving large-scale practical frictional contact problems, it is not so mathematically rigorous. Furthermore certain numerical difficulties are encountered, for example load step size and time step size for the numerical equation solver become critical to ensure convergence. Also the solution depends on several artificial numerical parameters [16, 17, 18]. Although this approach has been incorporated into general-purpose finite element programs, careful selection of numerical parameters by experience into the given problem is required to obtain a reliable solution.

The second formulation is variational inequalities formulation (VI). The variational inequality approach forms a relatively new direction in frictional contact research. It is mathematically rigorous and allows better modeling of the physics. The free boundary problem is addressed by this approach. Although some research has been done in this direction, it should be mentioned that a solution to the variational inequality for general frictional contact problems is still open. The usage is still restricted to the academic community, and solving large-scale practical frictional contact problems has not yet been attempted. This is the approach we use in this thesis.

The problem of an elastic solid in contact with a frictionless rigid body was first formulated by Signorini [19, 20]. Later Fichera [21] formulated the problem using a variational inequality and provided the condition for the existence and uniqueness of the solution to the Signorini problem. Duvaut and Lions [22] formulated the frictional contact problem in terms of variational inequality. Later computational studies of contact problems by variational inequality included those by Hlaváček and Lovíšek [23], Kalker [24], Oden and Carey [25]. Kikuchi and Oden [26] gave a systematic explanation of the formulation and
solution of the contact problem. Böhm [27] made a comparison of different contact algorithms in the literature. Refaat and Meguid [28, 29, 30] solved some practical static problems using variational inequalities. These kinds of numerical modelings have the advantage of being able to handle solids with more general geometries, as well as taking into account surface friction and the coupling of the normal traction and tangential traction. Recently, a review article by Mijar and Arora [31] gave a relatively complete introduction about the development and comparison of various approaches to the elastostatic frictional contact problem.

The physical mechanism for dynamic friction is much more complex than static friction. When one solid body slides on another, an apparent decrease of fractional resistance at the time of initial sliding has been observed for a long time, but some concepts for static friction are not applicable in dynamic situations. Moreover, the results of many experimental reports have little agreement with each other. Numerical solutions of dynamic frictional problems have been explored in finite element methods by Hughes, Taylor, and Sackman [32]. These dynamical problems have also been investigated from the variational inequality approach by Kikuchi [26], Wriggers and Scherf [33] and A. Czekanski. S. A. Meguid. N. El-Abbasi and M. H. Refaat [34, 35]. Generally, however, numerical approaches to dynamic frictional contact problem are still at an early stage of development. These kinds of problems do not fall within the purview of this thesis, but the issues addressed here are also relevant for such problems.

All of these finite element calculations are at the macroscopic level as they treat contact surfaces as continuous smooth surfaces. In other words, these calculations are descriptions in the framework of continuum mechanics and say nothing about the microscopic structures of contact surfaces. Therefore the results from finite element calculations can not be regarded as an ultimate understanding of the frictional contact problem. but it is reasonable to look upon the solution at a macroscopic point from finite element methods as an average value over the microscopic area around that macroscopic point. These calculations tell us the influence of macroscopic geometries on the average stress and strain at a macroscopic point, offer valuable information and pave the way for further studies.

Surface roughness was taken into account through finite difference methods in the computational study of contact problems by Ai [36] and Tian [37]. They paid attention
to the roughness of surfaces, but ignored the friction and macroscopic geometries of the solids. Computational methods at atomic scale were attempted by Gerde and Marder [38]. The contact solid was modelled as a system of discrete classical particles, which will be illustrated in Chapter 2. But both roughness calculations and calculations at the atomic scale are not treated in this thesis.

The purpose of this thesis is to research 2D frictionless and frictional contact on the basis of the variational inequality approach. An alternative method which iteratively couples the deformation of the two elastic solids is presented. Previous research has ignored the effect of elements such as loading distribution, loading step, total geometry and dimension of the solids. Our investigation considers the influence of the above elements. Our results show that the stress distribution deviates from Hertz model under frictional and elastic condition. The stress and deformation are also found to be sensitive to those additional elements, such as loading distribution. Relative transverse shift of the particles between the two contact surfaces does affect the stress. Before we can illustrate these different results, it is necessary to introduce the Hertz method and its foundation.

1.2 Hertz Contact Problem

Hertz's research on mechanics of contact is less known than his contribution to electromagnetic field theory, but the method he used in solving the contact problem had a great impact on later research. He developed a formula from the theory of elasticity to describe the contact pressure of a non-conforming contact surface with a regular smooth geometry. Hertz's original work is not easy to follow due to the evolution of scientific language and notation. Landau [39] reformulated the problem in modern notation following principles similar to Hertz's original work. Therefore, the convention of Landau's notation and approach is followed here.

The Hertz theory starts with two basic assumptions:

i) Each contact body is regarded as an elastic half-spaced material; this means it is flat at the surface of contact and extends infinitely towards the other side of the material. This assumption is based on the fact that the size of contact area is small compared to the dimension of each body and the radius $R$ of surface curve. This is acceptable in many
real situations.

ii) The surface is assumed to be frictionless; this means there is no tangential friction force around the contact area, which is also reasonable if the friction force is small compared to the normal pressure. The two surfaces are approximated as quadratic functions defined by

\[
z_1 = \frac{u_1^2}{2R_1} + \frac{v_1^2}{2R_1} = k_{\alpha\beta}x_\alpha x_\beta
\]

(1.1)

\[
z_2 = \frac{u_2^2}{2R_2} + \frac{v_2^2}{2R_2} = k'_{\alpha\beta}x_\alpha x_\beta
\]

(1.2)

see Figure (1.2). The approach \( h \) is defined as the distance between the starting contact points on both bodies \( o \) and \( o' \), if the solid bodies can penetrate into each other as pressure is applied. This is illustrated in Figure (1.1) and Figure (1.2). Figure (1.2) is an enlargement of Figure (1.1) in the region of the contact area \( \Omega_c \). The boundary condition within the contact region can be written as

\[(z_1 + u_1) + (z_2 + u_2) = \delta_1 + \delta_2 = h\]  

(1.3)

where \( u_1 \) and \( u_2' \) represent the displacements of particles on the two contact surfaces with same coordinates \( x \) and \( y \). The \( \delta_1 \) and \( \delta_2 \) are the displacement of the particles on the center of contact region \( \Omega_c \) for body 1 and body 2 respectively. From the diagram we have

\[(z_1 + u_1) + (z_2 + u_2) > h\]  

(1.4)

outside the contact area.

It is a reasonable approximation to apply the following formula (1.5) to the contact surfaces from the condition of the first Hertz Theory assumption. Suppose the flat surface of the semi-infinite solid is on the \( xy \) plane, then (1.5) describes the normal displacement \( u_z \) of any surface point in the semi-infinite solid media caused by a point loaded normal force \( P(x', y') \) applied at the surface point \( (x', y') \) of the solid body. The semi-infinite solid media has a modulus of elasticity \( E \) and Poisson ratio \( \nu \)

\[u_z (x, y) = \frac{1 - \nu^2}{\pi E} \frac{P(x', y')}{\Delta r}\]  

(1.5)

\( P \) is the normal force applied at the source point \( (x', y', 0) \) on the surface of the half spaced elastic material. Here \( \Delta r \) is the distance between the field point \( r (x, y, 0) \) and source point
Figure 1.1. Surface of contact between two elastic bodies, $h = \delta_1 + \delta_2$, $a$ is the radius of contact area. $\Omega_c$ is the contact region and $\Omega_1$ and $\Omega_2$ represent body 1 and body 2. The solid line represents the surface boundary after deformation. The dashed line represents the surface boundary without deformation if the two bodies can penetrate each other.
Figure 1.2. Enlargement of Figure (1.1), $\Omega_1$ and $\Omega_2$ represent solid 1 and solid 2, $\delta_1$ and $\delta_2$ are the displacement of the surface particle at the center of contact regions. $a$ is the radius of contact area, $u_{1z}, u_{2z}$ are the displacement of the particle on two surface. $\Omega_c$ is the area of contact.

The solid line represents the surface boundary after deformation. The dashed line represents the surface boundary without deformation if the two bodies can penetrate each other.
\( r' (x', y', 0) \) of the applied force.

\[
\Delta r = |r(x, y, 0) - r'(x', y', 0)| = \sqrt{(x-x')^2 + (y-y')^2}
\] (1.6)

Boussinesq [40] derived (1.5) from potential theory. For a detailed modern derivation see [39]. By linear superposition the total surface normal deflection at \((x, y)\) is

\[
u_{11}(x, y) = \frac{1-v^2}{\pi E} \int \int_{\Omega} \frac{P(x', y')}{\Delta r} \, dx' \, dy'
\] (1.7)

Similarly for the second solid body,

\[
u_{22}(x, y) = \frac{1-v^2}{\pi E'} \int \int_{\Omega} \frac{P(x', y')}{\Delta r} \, dx' \, dy'
\] (1.8)

where \(E, E'\) and \(\sigma, \sigma'\) are the moduli of elasticity and Poisson’s ratios of the two materials.

Substituting into the contact condition equation (1.3), within the contact region, we obtain

\[
\left( \frac{1-\sigma^2}{\pi E} + \frac{1-\sigma'^2}{\pi E'} \right) \int \int_{\Omega} \frac{P(x', y')}{r} \, dx' \, dy' = h - Ax^2 - By^2
\] (1.9)

where \(A\) and \(B\) depend on the radii of the two surfaces within the contact region. Since \(\Omega\) is unknown, it would be difficult to solve \(P(x', y')\) from the integral equation (1.9) directly, but it can be solved through analogy with electrostatic theory.

It is known [13] that the potential for a uniform charged ellipsoid with boundary surface

\[
\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1
\] (1.10)

is

\[
\phi(x, y, z) = \pi abc \int_0^\infty \frac{1 - \frac{x^2}{a^2 + \xi} - \frac{y^2}{b^2 + \xi} - \frac{z^2}{c^2 + \xi}}{\sqrt{(a^2 + \xi)(b^2 + \xi)(c^2 + \xi)}} \, d\xi
\] (1.11)

When \(c \to 0\), it becomes a flat ellipsoid in the \(z\) direction represented by

\[
\phi(x, y, z) = \pi ab \int_0^\infty \frac{1 - \frac{x^2}{a^2 + \xi} - \frac{y^2}{b^2 + \xi}}{\sqrt{(a^2 + \xi)(b^2 + \xi)}} \, d\xi
\] (1.12)

It is also known that an electrostatic potential can be written as

\[
\phi(x, y, z) = \int \int \frac{\rho(x', y', z')}{\sqrt{(x-x')^2 + (y-y')^2 + (z-z')^2}} \, dx' \, dy' \, dz'
\] (1.13)

after integration with respect to \(z\), the potential can also be written as

\[
\phi(x, y, z) = 2 \pi c \int \int_{\text{ellipse}} \frac{dz' \, dy'}{\sqrt{(x-x')^2 + (y-y')^2}} \sqrt{1 - \frac{x'^2}{a^2} - \frac{y'^2}{b^2}}
\] (1.14)
The integration domain is the region of the ellipse inside \( \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \). From (1.11) and (1.14) we have

\[
\int \int_{\text{ellipse}(a,b)} \frac{dx'dy'}{r} \sqrt{1 - \frac{x'^2}{a^2} - \frac{y'^2}{b^2}} = \frac{\pi ab}{2} \int_0^\infty \left\{ 1 - \frac{x^2}{a^2} \xi - \frac{y^2}{b^2} \xi \right\} d\xi
\]  

(1.15)

The right side of (1.15) is quadratic with respect to \( x \) and \( y \). By comparing (1.9) and (1.15) and guessing that the contact region \( \Omega \) is in the form of a ellipse \( \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \), we can derive the \( P_z(x',y') \) as

\[
P_z(x',y') = \frac{3F}{2\pi ab} \sqrt{1 - \frac{x'^2}{a^2} - \frac{y'^2}{b^2}}
\]  

(1.16)

where \( F \) is the total normal load. The approach \( h \), semi-major axis \( a \) and semi-minor axis \( b \) can be derived from

\[
h = \frac{FD}{\pi} \int_0^\infty \frac{d\xi}{\sqrt{(a^2 + \xi)(b^2 + \xi)}}
\]  

(1.17)

\[
A = \frac{FD}{\pi} \int_0^\infty \frac{d\xi}{(a^2 + \xi)\sqrt{(a^2 + \xi)(b^2 + \xi)}}
\]  

(1.18)

\[
B = \frac{FD}{\pi} \int_0^\infty \frac{d\xi}{(b^2 + \xi)\sqrt{(a^2 + \xi)(b^2 + \xi)}}
\]  

(1.19)

In the general case, \( a, b \) can only be solved numerically from \( A, B \), but if the contact surfaces are spherical with radii \( R \) and \( R' \), then

\[
A = B = \left( \frac{1}{2R} + \frac{1}{2R'} \right)
\]  

(1.20)

so the contact ellipse becomes a circle with

\[
a = b = \sqrt{\frac{FDRR'}{R + R'}}
\]  

(1.21)

and the approach becomes

\[
h = F^3 \sqrt{\frac{D^2 (R + R')}{RR'}}
\]  

(1.22)

These results are for the three dimensional problem. If it is a two dimensional problem, which represents two cylinders in contact with their axes parallel, it is simplified into a plane strain problem. Hertz theory gives the half contact semi-width as the limit of semi-major axis at \( \frac{b}{a} \to \infty \) as

\[
a = \sqrt{\frac{8DF}{3\pi A}} = 2 \frac{F \left( \frac{1 - \sigma^2}{E} + \frac{1 - \sigma^2}{E} \right)}{\pi \left( \frac{1}{K_1} + \frac{1}{K_2} \right)}
\]  

(1.23)
and the contact pressure distribution as

$$P(x', y') = \frac{2E}{\pi a} \sqrt{1 - \frac{x'^2}{a^2}}$$  \hspace{1cm} (1.24)

One of the problems with the theory is that the approach is of unbound since the integration diverges

$$h = \frac{1}{\pi} \left( \frac{1 - \sigma^2}{E} + \frac{1 - \sigma'^2}{E'} \right) F \int_0^\infty \frac{d\xi}{\sqrt{(a^2 + \xi)\xi}}$$  \hspace{1cm} (1.25)

Although Hertz's theory has stood for a century, its theoretical justification is incomplete in several aspects. One example involves Boussinesq's formula given in equation (1.7), which is derived for infinite half-spaced material with an ideal flat boundary surface under a point loaded force. Any force has a finite area on which it acts. If the area shrinks to a mathematical point, that means the pressure will reach infinity, which no real material can stand. If the size of the solid body is finite and small, for example a steel ball between the race of a roller, Boussinesq's formula would not be suitable. Furthermore, Hertz did not account for how sensitive the stress of the contact region would be to the applied load distribution. Independence from the loading distribution is built into Hertz's theory. This may be true for solid bodies of large size, but for small size solid bodies, there may be significant differences generated by different distributions of normal loading, different geometries and different size of the two solid bodies. This needs to be tested and may be accomplished through numerical methods.

The second important aspect is that Hertz's theory does not specify the boundary condition of other parts of the surfaces of the two contact bodies. It is known from elasticity theory that a solution for the stress exists and is unique only when all the boundary conditions are explicitly defined. The total solid body $\Gamma$ is made up of the displacement boundary $\Gamma_d$, where $u_{x \in \Gamma_d} = U(x)$ and the force boundary $\Gamma_F$ where $\tau_{x \in \Gamma_F} = F(x)$. $U(x)$ is the known displacement distribution on the boundary $\Gamma_d$ and $F(x)$ is the known traction force distribution on the boundary $\Gamma_F$. Both $u$ and $\tau$ are undefined on the boundary of the contact region for the contact problem. Hertz did not address this problem mathematically.

Another shortcoming is that Hertz's theory is designed to tackle a very idealized model with frictionless contact. Friction at the contact surface and elastic deformation of the second solid body always exists, The entanglement of friction and elastic deformation, plus
the microslip of the part of the contact surface make the real frictional contact problem very difficult to model analytically. The results of Hertz theory are used to investigate the stress distribution within the material and underneath the contact region later by Sackfield & Hills [10] and Johnson [41]. The displacement of a field point can be calculated by substituting the Hertz pressure (1.24) into (1.7). Since we are concerned only about the results for the two dimensional plane strain problem, the principle stresses along the symmetry axis starting from the middle of contact area can be calculated as

$$\sigma_x = -2\nu P_{\text{max}} \left( \sqrt{1 + \frac{y^2}{b^2}} - \frac{y}{b} \right)$$  \hspace{1cm} (1.26)

$$\sigma_y = -\frac{2P_{\text{max}}}{\sqrt{1 + \frac{y^2}{b^2}}} \hspace{1cm} (1.27)$$

$$\sigma_z = \nu (\sigma_x + \sigma_y) \hspace{1cm} (1.28)$$

The principle shear stress is computed as

$$\tau_1 = \frac{1}{2} (\sigma_y - \sigma_x) \hspace{1cm} (1.29)$$

It was found by above calculation that the maximum shear stress along the vertical axis direction is around 30\% of the $P_{\text{max}}$, and it is located at position $y = 0.78a$. where $a$ is the semi-width. We will compare results of our numerical modeling with this result in Chapter 4 using the same parameter.

1.3 The Signorini Contact Problem

Hertz theory avoids facing the changing boundary condition in the contact region by treating the contact patch as an ellipse. While true for many regular smooth surfaces which can be approximated with quadratic functions at the contact region, it is not true for general geometries. One of major drawbacks of Hertz theory is not considering the changing boundary of the contact area, which is dependent on the solution of the stress of the field and also affects the solution. Free boundary contact adds a nonlinear nature to the contact problem even with a linear elastic material. Signorini [19] first formulated the frictionless contact in terms of a free boundary condition, which is known as Signorini contact problem. The following paragraph is a short summary of the Signorini Problem.
Figure 1.3. Contact between elastic body and rigid body. $\Omega_1$ is an elastic solid, and $\Omega_2$ is a rigid solid. $\Gamma_{d_1}$ is the displacement boundary of $\Omega_1$, $\Gamma_{c_2}$ is the displacement boundary of $\Omega_2$. $\Gamma_{F_1}$ is the traction boundary of $\Omega_1$, and $\Gamma_{F_2}$ is the traction boundary of $\Omega_2$. $\tau_1$ is the surface traction on $\Omega_1$, $\tau_2$ is the surface traction on $\Omega_2$. 

a) before contact  b) after contact
with a schematic diagram as Figure (1.3). The effect of a solid body in contact with a smooth frictionless rigid surface is described by a vector displacement field \( u(x) \) satisfying the equilibrium equation

\[
\frac{\partial \sigma_{ij}(u)}{\partial x_j} + f_i = 0
\]  

for \( \Omega \) where \( f_i \) is the body force density in the direction of \( i \) for \( (i = 1, 2, 3) \), which represent the \( x, y, z \) directions, and \( \Omega \) is the field domain of the solid. The stress \( \sigma_{ij}(u) \) is related to displacement as

\[
\sigma_{ij}(u) = E_{ijkl} u_{k,l}
\]  

where \( E_{ijkl} \) is the tensor of modulus of elasticity for the linear elastic material and \( u_{k,l} = \frac{\partial u_k}{\partial x_l} \). The boundary surface of \( \Omega \) is denoted by \( \Gamma \). With the boundary conditions

\[
u_i = U_i
\]  

on \( \Gamma_d \)

\[
\sigma_{ij}(u) n_j = t_i
\]  

is the condition on \( \Gamma_F \), where \( \Gamma_d \) is the part of the boundary surface in which the displacement constraint is defined. \( \Gamma_F \) is the part of the boundary surface in which the force condition is defined. \( \Gamma_c \) is the part of the boundary surface which includes the real contact area and the part of the surface which is traction free. \( \Gamma_c \) is chosen to be large enough to cover the real contact area. \( \Gamma = \Gamma_d \cup \Gamma_F \cup \Gamma_c \). The division of the boundary surface \( \Gamma \) into \( \Gamma_d, \Gamma_F, \) and \( \Gamma_c \) is well defined before the solution, but the real contact area and the exact boundary between the real contact area and other parts of the boundary surface is the problem to be solved. The boundary conditions on the contact region \( \Gamma_c \) are given by

\[
\sigma_t = 0
\]  

since the surface is smooth, and \( \sigma_t \) is the magnitude of the tangential traction force.

\[
\sigma_n = 0 \quad \text{when } u_n - g_n(x) \leq 0
\]

\[
\sigma_n \leq 0 \quad \text{when } u_n - g_n(x) = 0
\]  

where \( g_n(x) \) is the normal gap at the point \( x \). It is the projection of the gap between two surface \( \psi(x, y) \) and \( \phi(x, y) \) on the normal direction of \( \psi(x, y) \). Putting the two conditions together

\[
\sigma_n (u_n - g_n(x)) = 0
\]  

15
where \( \sigma_n \) is the normal traction.

\[
\sigma_n = \sigma_{ij}n_jn_i
\]  \hspace{1cm} (1.37)

\[
u_n = u_in_i
\]  \hspace{1cm} (1.38)

The normal direction unit vector at the contact surface is \( \mathbf{n} = (n_1, n_2, n_3) \). The above inequality boundary condition is called the unilateral boundary condition. Our derivation of the gap function \( g_n(\mathbf{x}) \) follows Kikuchi [26] and Refaat [29]. Let \( \phi(x, y) \) be the function representing the contact boundary surface of the elastic body before deformation and let \( \psi(x, y) \) be the function representing the rigid solid surface in contact with the elastic body. The displacement of every particle in the domain can be represented as a vector field with three components.

\[
u = (u_x(x), u_y(x), u_z(x))
\]  \hspace{1cm} (1.39)

The contacting surface \( z = \phi(x, y) \) becomes the new surface \( z' = \phi'(x', y') \) after the deformation, where

\[
x' = x + u_x(x), \quad y' = y + u_y(x), \quad z' = z + u_z(x)
\]  \hspace{1cm} (1.40)

The unilateral contact condition is a strict restriction that no overlapping of solid material occurs. It may be formulated as

\[
\phi(x, y) + u_z(x) = \phi'(x + u_x(x), y + u_y(x))
\]  \hspace{1cm} (1.41)

\[
\phi'(x + u_x(x), y + u_y(x)) \geq \psi(x + u_x(x), y + u_y(x))
\]  \hspace{1cm} (1.42)

Since the displacement field is small, applying a Taylor expansion and ignoring the 2nd order and above terms, the equation (1.42) becomes

\[
u_z(x) - \frac{\partial \psi}{\partial x}|_{x,y} u_x(x) - \frac{\partial \psi}{\partial y}|_{x,y} u_y(x) \geq \psi(x) - \phi(x)
\]  \hspace{1cm} (1.43)

This can also be rewritten as

\[
\hat{n}_\psi(x,y) \cdot \mathbf{u} \geq g_n(x)
\]  \hspace{1cm} (1.44)

where

\[
\hat{n}_\psi(x,y) = \left( n_1\psi(x,y), n_2\psi(x,y), n_3\psi(x,y) \right)
\]  \hspace{1cm} (1.45)
and
\[
\begin{align*}
n_1 \psi(x,y) &= \frac{-\frac{\partial \psi}{\partial x}|_{x,y}}{\sqrt{1 + \left(-\frac{\partial \psi}{\partial x}|_{x,y}\right)^2 + \left(-\frac{\partial \psi}{\partial y}|_{x,y}\right)^2}} \\
n_2 \psi(x,y) &= \frac{-\frac{\partial \psi}{\partial y}|_{x,y}}{\sqrt{1 + \left(-\frac{\partial \psi}{\partial x}|_{x,y}\right)^2 + \left(-\frac{\partial \psi}{\partial y}|_{x,y}\right)^2}} \\
n_3 \psi(x,y) &= \frac{1}{\sqrt{1 + \left(-\frac{\partial \psi}{\partial x}|_{x,y}\right)^2 + \left(-\frac{\partial \psi}{\partial y}|_{x,y}\right)^2}}
\end{align*}
\]  
(1.46)

are the three components of the unit normal vector of \(\psi\) at the point \((x, y)\). The gap is defined as
\[
g_n(x) = \frac{\psi(x,y) - \phi(x,y)}{\sqrt{1 + \left(-\frac{\partial \psi}{\partial x}|_{x,y}\right)^2 + \left(-\frac{\partial \psi}{\partial y}|_{x,y}\right)^2}}
\]  
(1.49)

see Figure (1.4). Kikuchi and Oden [26], Refaat and Meguid [29] use an opposite contact normal direction \(\hat{\mathbf{n}} = -\hat{\mathbf{n}}_\psi(x,y)\), which points toward the inside of surface of \(\Omega_2\), so the contact condition becomes
\[
\hat{\mathbf{n}} \cdot \mathbf{u} \leq \frac{-[\psi(x) - \phi(x)]}{\sqrt{1 + \left(-\frac{\partial \psi}{\partial x}|_{x,y}\right)^2 + \left(-\frac{\partial \psi}{\partial y}|_{x,y}\right)^2}}
\]  
(1.50)

1.4 The Variational Approach to Numerical Modeling

Although the formulation of the Signorini problem is well defined, it is still a great challenge to solve it analytically by the traditional elasticity theory of solids, because the indefinite boundary condition does not fit into any category of boundary condition problem such as the Dirichlet and Neumann boundary conditions we have ever met before. The displacement and traction are both unknown on \(\Gamma_c\) and dependent on the solution of the problem. One approach with great potential for numerical solution is the variational inequality method. The solution can be found through the weak solution of a variational inequality. The weak solution is a solution \(\mathbf{u}(x)\) for any \(\mathbf{v} \in \mathbf{K}\) satisfying
\[
\int_{\Omega} E_{ijkl} u_{k,l} (v_{i,j} - u_{i,j}) \, dx \geq \int_{\Omega} f_i (v_i - u_i) + \int_{\Gamma_t} t_i (v_i - u_i) \, ds
\]  
(1.51)
where
\[ K = \{ v | v_i = U_i \text{ on } \Gamma_d, \ u_{n_o} - g_{n_o} (x) \geq 0 \text{ on } \Gamma_c \} \]  
(1.52)

The above weak solution is proven to be equivalent to the solution of the Signorini problem above [26]. The problem of the weak solution written in concise notation is
\[ a(u, v - u) \geq f(v - u) \]  
(1.53)

where
\[ a(u, v - u) = \int_\Omega E_{ij,kl} u_{k,l} (v_{i,j} - u_{i,j}) \, dx \]  
(1.54)
\[ f(v - u) = \int_\Omega f_i (v_i - u_i) + \int_{\Gamma_f} t_i (v_i - u_i) \, ds \]  
(1.55)

since
\[ \epsilon_{ij} (v - u) = \frac{1}{2} [(v_i - u_i)_{,j} + (v_j - u_j)_{,i}] \]  
(1.56)

and the stress and strain tensor is symmetric, we have
\[ \int_\Omega \sigma_{ij} (u) \epsilon_{ij} (v - u) \, dx = \int_\Omega \sigma_{ij} (u) (v_i - u_i)_{,j} \, dx \]  
(1.57)

We can solve the classical Signorini problem through seeking the weak solution \( u \) which satisfies the above inequality. The advantage of the variational inequality formulation

18
of the free boundary problem is that it can make use of the method of optimization of
a functional to find the close answer to the real one. Functional optimization is a well
developed computational method, which will be discussed in Chapter 3.

A variational approach to solve the Signorini problem provides a new angle of research
into contact and friction. But the Signorini contact problem is still an approximated ide-
alistic model for the real contact problem. In reality there are many factors which affect
the solution, for example the interaction of two elastic bodies, which is more complicated
than just one elastic body in contact with the smooth rigid surface. In addition to this,
friction or tangential traction on the contact region, and the coupling between the tan-
gential and normal traction make the real contact problem more complicated. Thermal
and dynamic effects play an important role too, but are outside the scope of this work.
All these considerations make the accurate modelling of frictional contact very difficult.
A simple model at the beginning is always helpful.

The classical Coulomb frictional law is used widely [26, 22] to model the friction at a
contact point. Let \( \sigma_n \) denote the vector of the contact surface traction, so that \( \sigma_n = \sigma_n \cdot n \)
is the magnitude of normal traction force at the contact point

\[
\sigma_t = \sigma_n - (\sigma_n \cdot n)n
\]  

(1.58)

is the vector representing tangential traction force on the contact surface. The Coulomb
frictional law assumes that the contact surface can be characterized by a positive coefficient
\( \mu \). When

\[
|\sigma_t| < -\mu (\sigma_n \cdot n)
\]  

(1.59)

there is no slippage between the two surfaces. Pending slippage is triggered at the juncture
if

\[
|\sigma_t| = -\mu (\sigma_n \cdot n) .
\]  

(1.60)

and there exists a \( \lambda > 0 \), such that \( \dot{u}_t = -\lambda \sigma_t \). The variational inequality applied to the
frictional contact problem is represented as

\[
a(u, v - u) + J(v, u) - J(u, u) \geq f(v - u)
\]  

(1.61)

for all \( v \) subject to \( v = 0 \) on \( \Gamma_d \) with \( n \cdot u - g \leq 0 \) on \( \Gamma_c \). The terms in equation (1.61) are
with \( s = -\mu (\sigma_n \cdot n) \)

\[
a(u, v) = \int_{\Omega} \sigma_{ij} (u) \varepsilon_{ij} (v) \, dV \tag{1.62}
\]

\[
J(u, v) = \int_{\Gamma_c} -\mu (\sigma_n \cdot n) |v| \, ds_b = \int_{\Gamma_e} s \, |v| \, ds_b \tag{1.63}
\]

where \( ds_b \) is the area of differential element on the boundary surface \( \Gamma_c \)

\[
f(v) = \int_{\Omega} f_i u_i \, dV + \int_{\Gamma_i} t_i u_i \, ds \tag{1.64}
\]

Equation (1.61) demands that the virtual work done by external force is not sufficient to overcome the potential energy increase due to both elastic and frictional contact on the boundary \( \Gamma_c \). This approach treats frictional force as the gradient of a frictional potential energy, but in fact this is a kind of approximation, assumes there is no plastic deformation, which means contact deformation has to be within the elastic limit. It is regarded as a static case at each equilibrium point, the thermal dissipation through frictional force at each incremental step is taken to be negligible. Equation (1.61) has been studied by Duvant [22], but the question whether there exists a solution or a unique solution is still open. Refaat [29] and Kikuchi [26] started from (1.61) to conduct numerical modeling and did find a convergent solution. This increased their confidence that the frictional contact problem has a unique stable solution. Our research work extends the search for the solution to (1.61) by using finite element methods to implement the variational inequality in studying the two dimensional frictional contact problem. The computational algorithm can deal with contact between two elastic solid bodies, and it can solve static and quasi-static cases. The example used to conduct numerical experiments is chosen for comparison with the theoretical results [41] and previous numerical practice. The solution exists but it is sensitive to many other conditions such as loading, boundary conditions, overall geometry and size, all of which are ignored by the Hertz theory. Another thing worth mentioning here is that all the previous theoretical analyses and many numerical calculations use the concept of Cauchy stress to describe the stress. The difference between Piola–Kirchhoff stress and Cauchy stress is addressed in the next chapter, and this distinction is shown to be important for discussing the contact problem.
Chapter 2

Theoretical Introduction

2.1 An Introduction to Finite Element Methods and Variational Inequalities

Finite element methods are employed in our modeling. We present here a short summary of the finite element approximation and how the solution of a variational inequality is implemented numerically by using it. The original idea of the finite element approximation starts by dividing the computational domain into a number of small elements called the mesh. Every element is made up of a certain number of node points. Any physical variable \( \mathbf{v}_h \) of a field point \( \mathbf{x} \) is approximated by taking linear combinations of the base function corresponding to each node in the computational domain. For example

\[
(\mathbf{v}_h)_i = \sum_{\alpha=1}^{P} v_i^{\alpha} \phi_{\alpha}(\mathbf{x})
\]

(2.1)

where \( \mathbf{x} \in \bar{\Omega}_h \), and \( \bar{\Omega}_h \) is the meshed computational domain corresponding to original domain \( \Omega \) representing elastic solid. Any notation with subindex \( h \) means being associated with the meshed computational domain. The coefficient \( v_i^{\alpha} \) corresponds to the value of the \( i \) th \((i = 1, 2, 3)\) component of the vector \( \mathbf{v} \) at the node position \( \alpha \), \( \phi_{\alpha}(\mathbf{x}) \) is the base function of node \( \alpha \) satisfying \( \phi_{\alpha}(\mathbf{x}) = \delta_{\alpha}^\beta \) at any node \( \mathbf{x}^\beta \in \bar{\Omega}_h \). The functional

\[
a_h(\mathbf{u}, \mathbf{v}) = \int_{\Omega_h} E_{ijkl} u_{k,l} v_{i,j} d\mathbf{x}
\]

(2.2)

can be evaluated as

\[
a_h(\mathbf{u}, \mathbf{v}) = E_{\alpha\beta} u_{k} v_{i}^{\beta}\phi_{\alpha}(\mathbf{x})
\]

(2.3)
where $E_{a\beta}^{ik} = \int_{\Omega_a} E_{ijkl} \phi_{a,j} \phi_{\beta,l} \, dx$. Equation (1.64) can also be written as

$$f_h (v) = f^i_a v^a_i$$  \hspace{1cm} (2.4)

where $f^i_a = \int_{\Omega_a} f_i \phi_a \, dx + \int_{\Gamma^i} t_i \phi_a \, ds$. The variational inequality (1.53) written in finite element form is therefore

$$E_{a\beta}^{ik} u^\beta_k (v^a_i - u^a_i) \geq f^i_a (v^a_i - u^a_i)$$  \hspace{1cm} (2.5)

A particular base function $\phi_a (x)$ has to be chosen in order to evaluate the above integral. Current finite elements methods choose $\phi_a (x)$ as a piecewise continuous function, which is non-zero only at the element that contains the node $a$. Also the condition $\phi_a (x^3) = \delta^a_\alpha$ is demanded for all basis functions. There are many way of choosing basis functions, sometimes called shape functions in finite elements text books [25, 42]. We generate the shape function $\phi_a (x)$ from polynomial interpolation functions, which takes the exact value $\delta^a_\alpha$ at all nodes of the elements. The higher order the polynomial approximation, the more smoother the change of the base function within the elements, but also the greater the node density. This will cause increases in computational time and memory load. Our modeling here uses 6 nodes per complex element, its shape function is a polynomial of order 2 and thus

$$T = \alpha_1 + \alpha_2 x + \alpha_3 y + \alpha_4 xy + \alpha_5 x^2 + \alpha_6 y^2$$  \hspace{1cm} (2.6)

where the coefficients $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6$ depend on the coordinates of the nodes in one element. In finite element analysis it is more convenient to use natural coordinates $(L_1, L_2, L_3)$ rather than global coordinates $(x, y)$. Note that $0 \leq L_1, L_2, L_3 \leq 1$ and $L_1 + L_2 + L_3 = 1$. The transformation from natural coordinate to global coordinate for a two dimensional triangular element takes the form

$$x = L_1 x_i + L_2 x_j + L_3 x_k$$  \hspace{1cm} (2.7)

$$y = L_1 y_i + L_2 y_j + L_3 y_k$$  \hspace{1cm} (2.8)

in the current work. We build the matrix array by choosing the 2nd order isoparametric triangular element mesh. Each element $e$ is made up of 6 nodes $(x_i, y_i), (x_j, y_j), (x_k, y_k), (x_l, y_l), (x_m, y_m)$ and $(x_n, y_n)$. The coordinates of the nodes in an element define the shape
Figure 2.1. Shape function $N_i, N_j, N_k, N_l, N_m, N_n$ corresponding to the 6 nodes $i, j, k, l, m, n$ of each complex element.

function as

$$\{N\} = \{N_i, N_j, N_k, N_l, N_m, N_n\}$$  \hspace{1cm} (2.9)

The shape functions within one element are (2.10), in terms of the natural coordinates $L_1, L_2, L_3$

$$
N_i = L_1(2L_1 - 1)
$$

$$
N_j = L_2(2L_2 - 1)
$$

$$
N_k = L_3(2L_3 - 1)
$$

$$
N_l = 4L_1L_2
$$

$$
N_m = 4L_2L_3
$$

$$
N_n = 4L_3L_1
$$

They are illustrated in Figure (2.1).

The global coordinates $(x, y)$ can be written in terms of natural coordinates within an isoparametric element,

$$
x = x_iN_i + x_jN_j + x_kN_k + x_lN_l + x_mN_m + x_nN_n
$$

$$
y = y_iN_i + y_jN_j + y_kN_k + y_lN_l + y_mN_m + y_nN_n
$$

This takes the same form as a variable function

23
\[ \phi(x, y) = \phi_t N_t + \phi_j N_j + \phi_k N_k + \phi_l N_l + \phi_m N_m + \phi_n N_n \] (2.12)

### 2.2 Cauchy Stress and Piola–Kirchoff Stress

Various descriptions of stress and strain have been used in the theory of elasticity, and it is necessary to distinguish between them here. Hertz theory and many analytical calculations actually use the Cauchy stress to describe the stress at the field point, which implies a spatial description is used. This means the independent variable is the present position \( x' \) occupied by a particle at time \( t \). Most numerical computation results are directly related to the 2nd Piola–Kirchoff stress tensor based on the reference description of motion, where the independent variable is chosen as the position \( x \) of the particle in a reference configuration. More precisely the reference configuration is chosen to be the initial configuration at \( t_0 = 0 \) before deformation. The symmetric Cauchy Stress tensor \( \sigma_{ij} \) is computed as a function of spatial coordinate, and Cauchy’s equation of motion is formulated as

\[ \frac{\partial \sigma_{ij}}{\partial x_j} + \rho b_i = \rho \frac{dv_i}{dt} \] (2.13)

where \( b_i \) denotes the body force density and \( v_i \) is the velocity of particle at the present position of the field point \( x \). This equation is derived from Newton’s second law. The left side of equation (2.13) represents all the force acting on an element of a unit volume and the right side of the equation is the acceleration term of this unit element. It is more convenient to use the reference coordinate to implement the finite element method due to a fixed reference coordinate. Piola–Kirchoff stress tensor is derived from the Lagrangian description [43]. There are two definitions of the Piola–Kirchoff Stress tensors. The Piola–Kirchoff stress tensor used in this thesis is the 2nd Piola–Kirchoff Stress tensor \( \tilde{\sigma} \) defined as

\[ \left( \tilde{\sigma} \cdot \tilde{N} \right) dS_0 = F^{-1} \cdot \sigma \cdot \tilde{n} dS \] (2.14)

where \( F^{-1} \) is a transformation matrix

\[ F^{-1} = x \tilde{V}_{x'} = \begin{pmatrix} \frac{\partial x}{\partial x'} & \frac{\partial y}{\partial x'} & \frac{\partial z}{\partial x'} \\ \frac{\partial x}{\partial y'} & \frac{\partial y}{\partial y'} & \frac{\partial z}{\partial y'} \\ \frac{\partial x}{\partial z'} & \frac{\partial y}{\partial z'} & \frac{\partial z}{\partial z'} \end{pmatrix} \] (2.15)
where \((x, y, z)\) is the reference point and \((x', y', z')\) is the current point. \((u, v, w)\) are displacements from the reference point to the current position.

\[
(x', y', z') = (x, y, z) + (u, v, w)
\] (2.16)

The second Piola–Kirchoff stress tensor \(\tilde{\sigma}\) can be calculated from displacement field of finite element methods. Conversion between \(\tilde{\sigma}\) and \(\sigma\) can be done with the help of

\[
\tilde{\sigma} = \frac{\rho_0}{\rho} (F^{-1})^T \cdot \tilde{\sigma} \cdot F^{-1} \cdot F^T
\] (2.18)

or

\[
\sigma = \frac{1}{J} F \cdot \tilde{\sigma} \cdot F^T
\] (2.19)

where \(J\) is the Jacobian

\[
J = \frac{\rho_0}{\rho} = \begin{vmatrix}
\frac{\partial x}{\partial x'} & \frac{\partial x}{\partial y'} & \frac{\partial x}{\partial z'} \\
\frac{\partial y}{\partial x'} & \frac{\partial y}{\partial y'} & \frac{\partial y}{\partial z'} \\
\frac{\partial z}{\partial x'} & \frac{\partial z}{\partial y'} & \frac{\partial z}{\partial z'}
\end{vmatrix}
\] (2.20)

and

\[
F = \begin{pmatrix}
1 + \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\
\frac{\partial v}{\partial x} & 1 + \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \\
\frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} & 1 + \frac{\partial w}{\partial z}
\end{pmatrix}
\] (2.21)

More specifically in our two-dimensional case \(w = 0\) with the plane strain condition

\[
J \begin{pmatrix}
\sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\
\sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\
\sigma_{zx} & \sigma_{zy} & \sigma_{zz}
\end{pmatrix} = \begin{pmatrix}
1 + \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & 0 \\
\frac{\partial v}{\partial x} & 1 + \frac{\partial v}{\partial y} & 0 \\
0 & 0 & 1
\end{pmatrix} \begin{pmatrix}
\tilde{\sigma}_{xx} & \tilde{\sigma}_{xy} & 0 \\
\tilde{\sigma}_{yx} & \tilde{\sigma}_{yy} & 0 \\
0 & 0 & \tilde{\sigma}_{zz}
\end{pmatrix}
\] (2.22)

\[
\begin{pmatrix}
1 + \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & 0 \\
\frac{\partial v}{\partial x} & 1 + \frac{\partial v}{\partial y} & 0 \\
0 & 0 & 1
\end{pmatrix}^T
\]

where \(J = \begin{vmatrix}
1 + \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\
\frac{\partial v}{\partial x} & 1 + \frac{\partial v}{\partial y}
\end{vmatrix} \). Expanding the matrix above gives

\[
\sigma_{xx} = \frac{1}{J} \left(1 + \frac{\partial u}{\partial x} \right) \left[ \tilde{\sigma}_{xx} \left(1 + \frac{\partial u}{\partial x} \right) + \tilde{\sigma}_{yx} \frac{\partial u}{\partial x} \right] \\
+ \frac{1}{J} \left(1 + \frac{\partial u}{\partial y} \right) \left[ \tilde{\sigma}_{xy} \left(1 + \frac{\partial u}{\partial y} \right) + \tilde{\sigma}_{yy} \frac{\partial u}{\partial y} \right] (2.23)
\]
\[ \sigma_{xy} = \frac{1}{J} \left( 1 + \frac{\partial u}{\partial x} \right) \left[ \sigma_{xx} \frac{\partial u}{\partial x} + \sigma_{xy} \left( 1 + \frac{\partial u}{\partial y} \right) \right] \] 
\[ + \frac{1}{J} \frac{\partial u}{\partial y} \left[ \sigma_{yx} \left( 1 + \frac{\partial u}{\partial x} \right) + \sigma_{yy} \frac{\partial u}{\partial y} \right] \] 
\[ (2.24) \]

\[ \sigma_{yz} = \frac{1}{J} \frac{\partial v}{\partial x} \left[ \sigma_{xx} \left( 1 + \frac{\partial u}{\partial x} \right) \right] \] 
\[ + \frac{1}{J} \left( 1 + \frac{\partial v}{\partial y} \right) \left[ \sigma_{yx} \left( 1 + \frac{\partial u}{\partial x} \right) + \sigma_{yy} \frac{\partial v}{\partial y} \right] \] 
\[ (2.25) \]

\[ \sigma_{yy} = \frac{1}{J} \left( \frac{\partial v}{\partial y} \right) \left[ \sigma_{xx} \frac{\partial v}{\partial x} + \sigma_{xy} \left( 1 + \frac{\partial v}{\partial y} \right) \right] \] 
\[ + \frac{1}{J} \left( 1 + \frac{\partial v}{\partial y} \right) \left[ \sigma_{yx} \left( 1 + \frac{\partial u}{\partial x} \right) + \sigma_{yy} \frac{\partial v}{\partial y} \right] \] 
\[ (2.26) \]

\[ \sigma_{zz} = \frac{1}{J} \sigma_{zz} \] 
\[ (2.27) \]

Thus the traction on the surface with the normal vector \( n \) can be written as \( \sigma_i = \sigma_i n_j \).

The normal pressure (normal traction) can be calculated using \( \sigma_n = \sigma_i n_i \), while the tangential traction on the surface can be calculated from

\[ \sigma_t = \begin{bmatrix} \sigma_1 & \sigma_2 \end{bmatrix} \begin{bmatrix} -n_2 \\ n_1 \end{bmatrix} \] 
\[ (2.28) \]

Normal pressure and tangential traction are calculated from the Cauchy stress instead of the Piola–Kirchoff stress. This is one place where current modeling differs from methods in the previous documents [29]. These two kinds of stress do not make much difference in small strain problem, but the difference becomes obvious in large strain. The Cauchy stress is more accurate in the calculation of the surface traction such as normal pressure and tangential traction.

2.3 Theoretical Formulation of the Contact Problem

A computational scheme to solve the two-dimensional contact problem based on the finite element approximation of a variational inequality is presented in this section. The Signorini problem is the simplest case in contact problem. It models an elastic solid body in contact with a frictionless rigid surface. But real contact involves both friction and elasticity in
two solid bodies. Frictional contact and the interaction between two elastic bodies can be solved on the basis of Signorini problem. We will discuss the numerical implementation of Signorini problem first.

### 2.3.1 Frictionless Contact between Elastic Solid and Rigid Surface

Our finite element formulation in constructing the variational inequality follows the same notation as Kikuchi [26, 29], but we choose penalty methods with multiplier to solve the variational inequality instead of the projected gradient method [29]. The variational inequality (1.53) is written in finite element form as (2.5) with the inequality constraint conditions expressed as

\[ n_\psi(x^\alpha) \cdot \mathbf{v}^\alpha - g_n|_{x^\alpha} = n_{\psi(x^\alpha)}^1 v_1^\alpha + n_{\psi(x^\alpha)}^2 v_2^\alpha - g_n|_{x^\alpha} \geq 0 \]  

(2.29)

on the contact boundary \( \Gamma_c \). We define \( g_n|_{x^\alpha} \) as the gap \( g_n \) defined in the last chapter evaluated at node point \( \alpha \). The \( n_\psi(x^\alpha) \) is the normal unit vector of the smooth rigid surface at location \( x = x^\alpha \) with components \( n_{\psi(x^\alpha)}^1 \) and \( n_{\psi(x^\alpha)}^2 \), which can be computed from the coordinates of node points on the surface of rigid body by the spline interpolation method.

This is another different method we have chosen to construct the modeling. Kikuchi [26] and Refaat [29] modelled the surface by connecting the neighboring nodes on the surface by straight lines, which could not produce a very smooth curve with continuous \( n_{\psi(x^\alpha)}^1 \) and \( n_{\psi(x^\alpha)}^2 \). The spline interpolation method proves to be a very successful way to generate a smooth curve with a continuous derivative. It will be discussed in the next chapter. N. El-Abbasi et al, showed a similar method in a very recent paper [44].

Kikuchi [26] and Duvaut & Lions [22] have proven that the solution of (2.5) is equivalent to the minimization problem of a quadratic functional \( I(\mathbf{v}) \), where

\[ I(\mathbf{v}) = \frac{1}{2} E_{ij} \epsilon_{ji} \mathbf{v}^\alpha \mathbf{v}_i^\alpha - f_i^\alpha \mathbf{v}_i^\alpha \]  

(2.30)

with inequality constraints associated with the contact boundary conditions

\[ \forall \mathbf{v}, \ n_\psi(x^\alpha) \cdot \mathbf{v}^\alpha (x) - g_n|_x \geq 0 \text{ on } \Gamma_c \]  

(2.31)

and equality constraints associated with displacement boundary conditions

\[ \mathbf{v}^\alpha_1 = U_{\alpha_1}^1 \text{ and } \mathbf{v}^\alpha_2 = U_{\alpha_2}^2 \text{ on } \Gamma_d \]  

(2.32)
Since each element of \( \{ E^{ik}_{i\alpha} \} \) is greater than zero and the highest order of the polynomial
\[
\frac{1}{2} E^{ik}_{i\alpha} v^\alpha_k v^\alpha_i - f^i_{i\alpha} v^\alpha_i
\]
is two, it is proven [22] that this quadratic functional has a unique minimum point. There are numerical schemes available for the optimization of quadratic functional. We will present such a scheme in chapter 3. To implement the scheme for the above problem it is necessary to build matrices \( \{ E^{ik}_{i\alpha} \}, \{ f^i_{i\alpha} \}, g_{n|x}, n^1_{\varphi(z_o)} \) and \( n^2_{\varphi(z_o)} \) from the coordinate information of the nodes and elements of the meshed computational domain, the surface traction information and the geometric shape of the two contact surfaces. Only the two dimensional plane strain problem is modeled here, which is an approximation to contact between two cylindrical bodies with their symmetry axes parallel to each other. Suppose the z axis is perpendicular to the cross section of the cylindrical body, then the z components of displacement are zero, i.e. \( u_z = 0 \). Furthermore, assume that the x and y components of displacement \( u_x, u_y \) are independent of \( z \)
\[
\frac{\partial u_x}{\partial z} = 0, \quad \frac{\partial u_y}{\partial z} = 0. \tag{2.33}
\]
Examining the definition of strain
\[
\epsilon_{ij} = \frac{1}{2} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \tag{2.34}
\]
we can deduce that any strain component containing \( z \) is zero. The independent function \( x_i \) means that we are using reference coordinate to define strain and stress. The stress is defined by
\[
\sigma_{ij} = \lambda_L \epsilon_{kk} \delta_{ij} + 2 \mu_L \epsilon_{ij} \tag{2.35}
\]
so the six component stress tensor are
\[
\sigma_{xx} = \lambda_L \epsilon_{yy} + (\lambda_L + 2 \mu_L) \epsilon_{xx} \tag{2.36}
\]
\[
\sigma_{yy} = \lambda_L \epsilon_{xx} + (\lambda_L + 2 \mu_L) \epsilon_{yy} \tag{2.37}
\]
\[
\sigma_{zz} = \lambda_L (\epsilon_{xx} + \epsilon_{yy}) = \frac{\lambda_L}{2(\lambda_L + \mu_L)} (\sigma_{xx} + \sigma_{yy}) \tag{2.38}
\]
\[
\sigma_{xy} = 2 \mu_L \epsilon_{xy} \tag{2.39}
\]
\[
\sigma_{xz} = 0 \tag{2.40}
\]
\[
\sigma_{yz} = 0 \tag{2.41}
\]
where $\lambda_L$ and $\mu_L$ are constant called Lamé coefficients related to the elastic property of a material

$$\lambda_L = \frac{E\nu}{(1 + \nu)(1 - 2\nu)} \quad \text{and} \quad \mu_L = \frac{E}{2(1 + \nu)}$$  \hspace{1cm} (2.42)

The stress element $\sigma_{zz}$ is not zero in the plane strain case, which means the two ends of the cylinder are not load free. We thus need to apply a load force to maintain the plain strain condition. If the dimension of the $z$ direction is much longer than the cross section, $u_z$ is so small compared with $u_x, u_y$ that plain strain condition is a reasonable approximation for solids with cylindrical symmetry. The dynamic equation will be reduced to

$$\frac{\partial}{\partial x} (\tilde{\sigma}_{xx}) + \frac{\partial}{\partial y} (\tilde{\sigma}_{xy}) + \rho b_z = \rho \frac{\partial^2 u}{\partial t^2} \hspace{1cm} (2.43)$$

$$\frac{\partial}{\partial x} (\tilde{\sigma}_{yz}) + \frac{\partial}{\partial y} (\tilde{\sigma}_{yy}) + \rho b_y = \rho \frac{\partial^2 v}{\partial t^2} \hspace{1cm} (2.44)$$

plus the compatibility equation

$$\frac{\partial^2 \varepsilon_x}{\partial x^2} + \frac{\partial^2 \varepsilon_y}{\partial y^2} = 2 \frac{\partial \varepsilon_{xy}}{\partial x \partial y} \hspace{1cm} (2.45)$$

applied to two dimensional case it is

$$\varepsilon_x = \frac{\partial u_x}{\partial x}, \quad \varepsilon_y = \frac{\partial u_y}{\partial y}, \quad \varepsilon_{xy} = \frac{1}{2} \left( \frac{\partial u_y}{\partial x} + \frac{\partial u_x}{\partial y} \right) \hspace{1cm} (2.46)$$

where $\varepsilon_x, \varepsilon_y$ are the short form notation for $\varepsilon_{xx}$ and $\varepsilon_{yy}$ respectively. From (2.35) we have

$$\tilde{\sigma}_{ij} = E_{ijkl} u_{k,l} \hspace{1cm} (2.47)$$

where $u_{k,l} = \frac{\partial u_k}{\partial x_l}$

$$E_{ijkl} = E_{jikl} = E_{ijkl} = E_{klij} \hspace{1cm} (2.48)$$

since

$$\begin{bmatrix} \tilde{\sigma}_x \\ \tilde{\sigma}_y \\ \tilde{\sigma}_{xy} \end{bmatrix} = \begin{bmatrix} \tilde{\sigma}_{11} \\ \tilde{\sigma}_{22} \\ \tilde{\sigma}_{12} \end{bmatrix} = \alpha \begin{bmatrix} 1 - \nu & \nu & 0 \\ \nu & 1 - \nu & 0 \\ 0 & 0 & 1 - 2\nu \end{bmatrix} \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{12} \end{bmatrix} \hspace{1cm} (2.49)$$

where

$$\begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{xy} \end{bmatrix} = \begin{bmatrix} \varepsilon_x \\ \varepsilon_y \\ \varepsilon_{xy} \end{bmatrix} = \begin{bmatrix} u_{1,1} \\ u_{2,2} \\ \frac{1}{2} (u_{1,2} + u_{2,1}) \end{bmatrix} \hspace{1cm} (2.50)$$
It can be known from (2.47)

\[
\begin{align*}
E_{1111} &= \alpha(1 - v) & E_{1112} &= 0 \\
E_{1121} &= 0 & E_{1122} &= \alpha v \\
E_{1211} &= 0 & E_{1212} &= \frac{1}{2} \alpha(1 - 2v) \\
E_{1221} &= \frac{1}{2} \alpha(1 - 2v) & E_{1222} &= 0 \\
E_{2111} &= 0 & E_{2112} &= \frac{1}{2} \alpha(1 - 2v) \\
E_{2121} &= \frac{1}{2} \alpha(1 - 2v) & E_{2212} &= 0 \\
E_{2211} &= \alpha v & E_{2212} &= 0 \\
E_{2221} &= 0 & E_{2222} &= \alpha(1 - v)
\end{align*}
\]

(2.51)

with \( \alpha = \frac{p}{(1+\nu)(1-2\nu)} \). The matrices \( \{E_{ik}^{\alpha}\} \) and \( \{f_i^\alpha\} \) are constructed by choosing the 2nd order isoparametric triangular element mesh with shape function

\[
\{N\} = \{N_i, N_j, N_k, N_l, N_m, N_n\}
\]

(2.52)

in terms of natural coordinates \( L_1, L_2, L_3 \). The shape function within the element is described in section 2.1. For isoparametric elements, the coordinate \((x, y)\) and variable function \(\phi(x, y)\) can be written in terms of natural coordinates as in (2.11) and (2.12).

Global base function is constructed from the piecewise continuous shape function by regarding \(\phi_\alpha(x) = 0\) in elements which do not contain the node \(\alpha\). Although 6 node triangular element is less accurate, it is easier to implement than higher node elements.

The principle is the same in using other elements to construct the base function. The final base function is formulated as

\[
\phi^{(e)}_{\alpha} = \begin{cases} 
N^{(e)}_{s(\alpha)} & \text{if } \alpha \in e(i, j, k, l, m, n) \\
0 & \text{if } \alpha \notin e(i, j, k, l, m, n)
\end{cases}
\]

(2.53)

where \(N^{(e)}_{s(\alpha)}\) means the shape function of the node at position \(s(\alpha)\) in element \(e\). The notation \(\alpha \in e(i, j, k, l, m, n)\) means node number \(\alpha\) matches the node number of one of the six nodes \{i, j, k, l, m, n\} of the element \(e\). Alternately when \(s(\alpha) = i, j, k, l, m, n\) of \(e\), then the base function is equal to the shape function at its corresponding element. In general the base function is a piecewise continuous function, and all the piece are made up shape function of relevant elements. This is illustrated in Figure (2.2). The derivative of
Figure 2.2. The base function $\phi_\alpha (x)$.

the base function is

$$
\phi^{(e)}_{\alpha,i} = \frac{\partial}{\partial x_i} \phi^{(e)}_\alpha = \begin{cases} 
\frac{\partial}{\partial x_i} N_{s(\alpha)}^{(e)} & \text{if } \alpha \in e \ (i, j, k, l, m, n) \\
0 & \text{if } \alpha \notin e \ (i, j, k, l, m, n)
\end{cases} \quad (2.54)
$$

where $s(\alpha) = i, j, k, l, m, n$ of $e$ and $l = 1, 2$. Thus $N_{s(\alpha)}^{(e)}$ is well defined by $\alpha$ and $e$. The integration of the stiffness matrix

$$
E^{ik}_{\alpha\beta} = \int_{\Omega_h} E_{ijkl} \phi_{\alpha,i} \phi_{\beta,l} dx
$$

(2.55)

is evaluated by expanding the integrand as

$$
E^{ik}_{\alpha\beta} = \int_{\Omega_h} E_{i1k1} \phi_{\alpha,1} \phi_{\beta,1} dx + \int_{\Omega_h} E_{i1k2} \phi_{\alpha,1} \phi_{\beta,2} dx
\quad (2.56)
$$

$$
+ \int_{\Omega_h} E_{i2k1} \phi_{\alpha,2} \phi_{\beta,1} dx + \int_{\Omega_h} E_{i2k2} \phi_{\alpha,2} \phi_{\beta,2} dx.
$$

We know $\phi_{\alpha,i} \phi_{\beta,l} \neq 0$ only when the node $\alpha$ and $\beta$ are situated in the same element, so the summation $\sum_e$ over all the elements takes effect only at the particular element that
contains both $\alpha$ and $\beta$.

$$E^{\epsilon k}_{\alpha \beta} = \sum_{\epsilon} \int_{\Omega_{\epsilon}} E_{\epsilon 1k1} \frac{\partial N^{(e)}_{s(a)}}{\partial x} \frac{\partial N^{(e)}_{s(b)}}{\partial x} dx$$

$$+ \sum_{\epsilon} \int_{\Omega_{\epsilon}} E_{\epsilon 1k2} \frac{\partial N^{(e)}_{s(a)}}{\partial x} \frac{\partial N^{(e)}_{s(b)}}{\partial y} dy$$

$$+ \sum_{\epsilon} \int_{\Omega_{\epsilon}} E_{\epsilon 2k1} \frac{\partial N^{(e)}_{s(a)}}{\partial y} \frac{\partial N^{(e)}_{s(b)}}{\partial x} dx$$

$$+ \sum_{\epsilon} \int_{\Omega_{\epsilon}} E_{\epsilon 2k2} \frac{\partial N^{(e)}_{s(a)}}{\partial y} \frac{\partial N^{(e)}_{s(b)}}{\partial y} dy$$

(2.57)

To evaluate the matrix element above, each term $\frac{\partial N_{x}}{\partial x}$, $\frac{\partial N_{y}}{\partial y}$ is needed. It can be calculated with the help of natural coordinates.

$$\begin{bmatrix} \frac{\partial N_{x}}{\partial L_{1}} \\ \frac{\partial N_{x}}{\partial L_{2}} \end{bmatrix} = J_{s} \begin{bmatrix} \frac{\partial N_{x}}{\partial x} \\ \frac{\partial N_{x}}{\partial y} \end{bmatrix}$$

(2.58)

$J_{s}$ is Jacobian of the transformation from natural coordinates to global coordinates

$$J_{s} = \begin{bmatrix} \frac{\partial x}{\partial L_{1}} & \frac{\partial x}{\partial L_{2}} \\ \frac{\partial y}{\partial L_{1}} & \frac{\partial y}{\partial L_{2}} \end{bmatrix}$$

(2.59)

so

$$\begin{bmatrix} \frac{\partial N_{x}}{\partial x} \\ \frac{\partial N_{y}}{\partial y} \end{bmatrix} = J_{s}^{-1} \begin{bmatrix} \frac{\partial N_{x}}{\partial L_{1}} \\ \frac{\partial N_{x}}{\partial L_{2}} \end{bmatrix}$$

(2.60)

The inverse of Jacobian is $J_{s}^{-1} = \frac{1}{|J_{s}|} \begin{bmatrix} \frac{\partial y}{\partial L_{2}} & -\frac{\partial y}{\partial L_{1}} \\ -\frac{\partial x}{\partial L_{2}} & \frac{\partial x}{\partial L_{1}} \end{bmatrix}$ where

$$|J_{s}| = \left( \frac{\partial x}{\partial L_{1}} \frac{\partial y}{\partial L_{2}} - \frac{\partial x}{\partial L_{2}} \frac{\partial y}{\partial L_{1}} \right)$$

(2.61)

Substituting this into (2.58) gives

$$\begin{bmatrix} \frac{\partial N_{x}}{\partial x} \\ \frac{\partial N_{y}}{\partial y} \end{bmatrix} = \frac{1}{|J_{s}|} \begin{bmatrix} \frac{\partial y}{\partial L_{2}} \frac{\partial N_{x}}{\partial L_{1}} - \frac{\partial y}{\partial L_{1}} \frac{\partial N_{x}}{\partial L_{2}} \\ -\frac{\partial x}{\partial L_{2}} \frac{\partial N_{x}}{\partial L_{1}} + \frac{\partial x}{\partial L_{1}} \frac{\partial N_{x}}{\partial L_{2}} \end{bmatrix}$$

(2.62)

Now we need to evaluate each term of (2.62) using (2.10) and (2.11).

$$\frac{\partial x}{\partial L_{1}} = (-x_{i} - 3x_{k} + 4x_{m} + (4x_{i} + 4x_{k} - 8x_{n}) L_{1}$$

$$+ (4x_{k} + 4x_{l} - 4x_{m} - 4x_{n}) L_{2}$$

(2.63)
\[
\frac{\partial x}{\partial L_2} = (-x_j - 3x_k + 4x_m) + (4x_k + 4x_l - 4x_m - 4x_n) L_1 + (4x_j + 4x_k - 8x_m) L_2.
\]

(2.64)

\[
\frac{\partial y}{\partial L_1} = (-y_i - 3y_k + 4y_m) + (4y_i + 4y_k - 8y_n) L_1 + (4y_k + 4y_l - 4y_m - 4y_n) L_2.
\]

(2.65)

\[
\frac{\partial y}{\partial L_2} = (-y_i - 3y_k + 4y_m) + (4y_k + 4y_l - 4y_m - 4y_n) L_1 + (4y_j + 4y_k - 8y_m) L_2.
\]

(2.66)

Substituting these terms into \[
\begin{bmatrix}
\frac{\partial N^{(e)}_i}{\partial x} \\
\frac{\partial N^{(e)}_i}{\partial y}
\end{bmatrix}
\]
leaves the problem of numerically evaluating the integrals.

\[
E^{ik}_{\alpha\beta} = \sum \int_{L_1=0}^{L_1=1} \int_{L_2=0}^{L_2=1} E_{ikij} \frac{\partial N^{(e)}_{s(\alpha)}}{\partial x} \frac{\partial N^{(e)}_{s(\beta)}}{\partial y} |J_s| dL_1 dL_2 \bigg|_{\Omega_{(e)}}
\]

(2.67)

\[
\sum \int_{L_1=0}^{L_1=1} \int_{L_2=0}^{L_2=1} E_{ik2j} \frac{\partial N^{(e)}_{s(\alpha)}}{\partial y} \frac{\partial N^{(e)}_{s(\beta)}}{\partial x} |J_s| dL_1 dL_2 \bigg|_{\Omega_{(e)}}
\]

\[
\sum \int_{L_1=0}^{L_1=1} \int_{L_2=0}^{L_2=1} E_{12k1} \frac{\partial N^{(e)}_{s(\alpha)}}{\partial x} \frac{\partial N^{(e)}_{s(\beta)}}{\partial y} |J_s| dL_1 dL_2 \bigg|_{\Omega_{(e)}}
\]

\[
\sum \int_{L_1=0}^{L_1=1} \int_{L_2=0}^{L_2=1} E_{12k2} \frac{\partial N^{(e)}_{s(\alpha)}}{\partial y} \frac{\partial N^{(e)}_{s(\beta)}}{\partial x} |J_s| dL_1 dL_2 \bigg|_{\Omega_{(e)}}
\]

The double integration can be evaluated using Gauss quadratures which give

\[
\int_{L_1=0}^{L_1=1} \int_{L_2=0}^{L_2=1} f (L_1, L_2) dL_1 dL_2 = \sum w_m [f (L_1, L_2)]_{L_1(m), L_2(m)}
\]

(2.68)

The detail of Gauss quadratures [25, 42] are given in appendix A. The next matrix necessary for computations is the force matrix

\[
f_\alpha = \int_{\Omega_{L_1}} f_i \phi_\alpha dx + \int_{\Gamma} t_i \phi_\alpha ds
\]

(2.69)

where \( f_1 (i = 1, 2) \) is the component of the body force density. The first term can be manipulated as

\[
\int_{\Omega_{L_1}} f_i (x) \phi_\alpha (x) dx = f_i \int_{\Omega_{L_1}} \phi_\alpha (x) dx = f_i \sum \int_{\Omega_{(e)}} \phi^{(e)}_\alpha (x) dx
\]

(2.70)
Since \( \phi^{(e)}_\alpha(x) \neq 0 \) only when \( \alpha \in \{i, j, k, l, m, n\} \). That is, the term in the summation is non-zero only for the element which contain \( \alpha \)

\[
\int_{\Omega^{(e)}} \phi^{(e)}_\alpha dx = \left\{ \begin{array}{ll}
\int_0^1 \int_0^1 N_{s(\alpha)}^{(e)} |J_s| dL_1 dL_2 & \text{where } \alpha \in \{i, j, k, l, m, n\} \\
0 & \text{where } \alpha \notin \{i, j, k, l, m, n\}
\end{array} \right.
\]

(2.71)

Integration by Gauss Quadrature gives

\[
\int_0^1 \int_0^1 N_{s(\alpha)}^{(e)} |J_s| dL_1 dL_2 = \sum_m w_m \left[ N_{s(\alpha)}^{(e)} |J_s| \right]_{L_{1m} L_{2m}}
\]

(2.72)

Since the degree of the polynomial is 4, 12 point Gauss Quadratures is chosen (see appendix A). The second term of the force matrix is denoted by

\[
\sigma^{ij}_\alpha = \int_{\Gamma} t_i \phi^{(e)}_\alpha ds
\]

(2.73)

and is non-zero only at boundary node points. This matrix can be evaluated by dividing the boundary into a series of small sections, and computing the length of each section and the shape function at the middle of the section, then summing over this boundary.

The term \( g_n|_{x=\alpha} \) represents the gap along the normal direction of \( \psi(x) \) at \( x = x^\alpha \). It is easy to compute \( g_n|_{x=\alpha} \) through the spline interpolation of \( \psi(x) \), which we discuss in chapter 3. Once all matrix elements are ready, the next task of this model is to solve quadratic optimization problem with both equality and inequality constraints represented by (2.30). It can be solved as an optimization problem without constraints by introducing the penalty function and a Lagrangian multiplier, where the penalty function is an artificially made function. It remains zero, when the constraint is not violated, and it increases rapidly to a very large number when the constraint is violated. A conjugated gradient method, with a restart scheme, is used in solving the unconstrained optimization problem. The details of these numerical schemes are left to the next chapter.

Equation (2.31) is an inequality which can be rewritten as

\[
-n_i^{\omega} \psi_{\omega}^{(e)} + g_n|_{x=\alpha} \leq 0
\]

(2.74)

where \( \alpha_c \) is the node number on contact region \( \Gamma_c \) as shown in Figure (2.3). We define

\[
g_{\alpha_c}^{\omega}(\nu) = \max \{0, g_{\alpha_c}(\nu)\}
\]

(2.75)

34
Figure 2.3. The contact boundary condition in finite elements

where

$$g_{ac}(v) = -n^i_{v(x^a)}v^a_i + g_n|x^a$$  

(2.76)

The penalty function takes the form $c_p \sum_{a=1}^{m} (g_{ac}^+(v))^2$, where $c_p$ is a very large number to make $c_p \sum_{a=1}^{m} (g_{ac}^-(v))^2$ grow very rapidly compared with the quadratic functional, once the constraint condition is not satisfied. The other equality constraints lead to penalty function being described as:

$$c_p \sum_{a_1}^{(v_1^{a_1} - U_1^{a_1})^2} + c_p \sum_{a_2}^{(v_2^{a_2} - U_2^{a_2})^2}$$  

(2.77)

where $\alpha_1$ represents the node number on which the displacement constraint of the $x$ component is applied as a constant $U_1^{a_1}$, and $\alpha_2$ represents the node number on which the displacement constraint of $y$ component is applied as constant $U_2^{a_2}$. The new functional $L(v)$ is obtained by summing up the original quadratic functional with the penalty function above. Solving the unconstrained optimization problem with respect to $v$ will then lead to the answer to the Signorini problem.

$$L(v) = \frac{1}{2} E^{ik}_{\alpha \beta} v^i_k v^i_\alpha - f^i_\alpha v^i_\alpha + c_p \sum_{a_1}^{(v_1^{a_1} - U_1^{a_1})^2}$$

$$+ c_p \sum_{a_2}^{(v_2^{a_2} - U_2^{a_2})^2} + c_p \sum_{a}^{(g_{ac}^-(v))^2}$$  

(2.78)
It is easy to use the conjugated gradient method to minimize $L(v)$, this particular approach is actually called the exact penalty method, this works well when the penalty parameter is not so large compared with the functional, but it sometimes suffers from ill conditioning when the convergence rates deteriorates. The drawbacks and other means of improvement are be discussed in the next chapter.

2.3.2 Contact between a Rigid Body and an Elastic Foundation without Friction.

This class of problems is called the rigid punch problem and has the opposite configuration to the Signorini problem. We can treat this problem as a special case of the Signorini problem by regarding the rigid body as stationary and making the elastic foundation approach to the rigid body under the action of a force on the floor of the elastic foundation. Due to the lack of the exact knowledge about the boundary force distribution at this floor, the solution may not be accurate. In this section we use a new scheme for the rigid punch problem. The contact body is treated as a rigid body, and the target body is modeled as elastic. A new variable of indentation distance of the rigid solid body into the elastic body $d_y$ is introduced, and the quadratic functional is written as functional of $v$ and $d_y$. It is a good approximation for the situation where the contact body is much harder than the target surface. The equivalent minimization problem is

$$
\min \left[ \frac{1}{2} a_h(v, v) - f_h(v) \right]
$$

(2.79)

$\forall v, d_y$ satisfying the contact boundary constraint

$$
\mathbf{n}|_{\psi(x)} \cdot v(x) - g_n|_x \geq 0
$$

(2.80)

on $\Gamma_c$, and the computational domain $\Omega_h$ is the elastic foundation. Here

$$
f_h(v) = f_0^i u_i^2 + F_y d_y
$$

(2.81)

$$
g_n|_x = N \left[ \psi(x) + d_y - \phi(x) \right]
$$

$F_y$ is the applied load, and $d_y$ is the indentation depth of the rigid solid body penetrating into the elastic solid body. $F_y d_y$ is the work done by external force. The boundary condition at the node $\alpha_c$ is

$$
\mathbf{n}|_{\psi(x, \alpha_c)} \cdot v_{\alpha_c} - g_n|_{x, \alpha_c} \geq 0
$$

(2.82)
where $\alpha_c$ is the contact node number of $\Gamma_c$. The system is also subject to equality constraints at the boundary nodes where the displacement is prescribed

$$u_1^{\alpha_1(k)} = U_1^{\alpha_1(k)}$$  \hspace{1cm} (2.83)

and

$$u_2^{\alpha_2(k)} = U_2^{\alpha_2(k)}$$  \hspace{1cm} (2.84)

where $\alpha_1(k)$ and $\alpha_2(k)$ are the numbers of the nodes, which have displacement constraints on the $x$ component and $y$ component, respectively. The downward normal unit vector on the rigid surface of contact is

$$n = \begin{bmatrix} \frac{\partial \omega}{\partial x} & \frac{\partial \omega}{\partial y} \end{bmatrix}, \begin{bmatrix} -1 \\ \sqrt{1 + \left(\frac{\partial \omega}{\partial x}\right)^2} \sqrt{1 + \left(\frac{\partial \omega}{\partial x}\right)^2} \end{bmatrix}$$  \hspace{1cm} (2.85)

Let

$$g_k^-(v) = \max\{0, g_k(v)\}$$  \hspace{1cm} (2.86)

where

$$g_k(v) = -\sum_{i=1}^{2} n_i^{\alpha_c(k)} \cdot u_i^{\alpha_c(k)} + g_{n_k}\left|u^{\alpha_c(k)}\right|$$  \hspace{1cm} (2.87)

and the penalty function is

$$P(v) = c_p \sum_{k=1}^{d_1} \left(u_1^{\alpha_1(k)} - U_1^{\alpha_1(k)}\right)^2 + c_p \sum_{k=1}^{d_2} \left(u_2^{\alpha_2(k)} - U_2^{\alpha_2(k)}\right)^2 + c_p \sum_{k=1}^{m} \left(g_k^-(v)\right)^2$$  \hspace{1cm} (2.88)

The rigid pouch problem is now equivalent to the unconstrained optimization problem

$$L(v) = \min \left[ \sum_{\alpha} \sum_{\beta} \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{1}{2} E_{\alpha\beta} u_k^{\beta} v_i^{\alpha} - \sum_{i=1}^{2} \sum_{\alpha} f_{\alpha} u_i^{\alpha} - F_d y_d + P(v) \right]$$  \hspace{1cm} (2.89)

### 2.3.3 Frictionless Contact between Two Elastic Solid Bodies

The problem of the contact between two smooth elastic bodies without friction is treated as an extension to the Signorini problem. The problem was solved [29, 26] by variational
inequality as in the Signorini problem, except that the integration is over two domains representing two bodies. The equivalent optimization problem is

$$\min \left[ \frac{1}{2} a_{h1}(v, v) + \frac{1}{2} a_{h2}(\bar{v}, \bar{v}) - f_{h1}(v) - f_{h2}(\bar{v}) \right]$$  \hspace{1cm} (2.90)$$

subject to constraints which are more complicated than the Signorini problem. where $v$ and $\bar{v}$ represent the displacement vector of solid body I and solid body II respectively, $v \in V_1$ and $\bar{v} \in V_2$, $(v, \bar{v}) \in V_1 \otimes V_2$. which is the function space extended by a direct product of $V_1$ and $V_2$. Since the two contact surfaces change shape at the same time, there are both deflection and tangential shifts between these two surfaces that make it more complicated to accurately model the contact condition. We present our proposal for dealing with the tangential shift and contact condition in subsection (2.3.4). This is another new approach used by our modeling. Previous research [26, 29] ignored the tangential shift between the two surfaces and used the contact condition

$$\forall (v, \bar{v}), n|_{\psi(x)} \cdot [v(x) - \bar{v}(x)] - g_n|x \geq 0 \text{ on } \Gamma_c$$  \hspace{1cm} (2.91)$$

The normal direction and initial gap are computed from the initial surface shape of the second body

$$\hat{n}|_{\psi(x)} = \left( \frac{\frac{\partial \psi}{\partial x}}{\sqrt{1 + \left( \frac{\partial \psi}{\partial x} \right)^2}}, \frac{1}{\sqrt{1 + \left( \frac{\partial \psi}{\partial x} \right)^2}} \right)$$  \hspace{1cm} (2.92)$$

$$g_n|x = \frac{\psi(x) - \phi(x)}{\sqrt{1 + \left( \frac{\partial \phi}{\partial x} \right)^2}}$$  \hspace{1cm} (2.93)$$

Since the actual normal direction is always changing too, this approximation will not accurately reflect the stress distribution and deformation of the surface in the contact region. Another consideration to be addressed is that the node distribution on the surfaces are not exactly matched in their $x$ coordinates, Usually each contact surface node of body I sits between two neighboring contact nodes of body II in terms of $x$ coordinate position. Thus it is necessary to locate the corresponding surface point with the same $x$ coordinate on the body II in the process of creating the contact constraint. The following contact constraint condition is similar to the method of Kikuchi [26] by applying (2.91) to each contact node point:

$$\forall (v, \bar{v}), n|_{\psi(x_{\alpha})} \cdot [v(x_{\alpha}) - \bar{v}(x_{\alpha})] - g_n|x_{\alpha} \geq 0$$  \hspace{1cm} (2.94)$$
at $\alpha_c$ on $\Gamma_c$, where $\alpha_c$ is a contact node in body I. The critical problem is how to compute $g_{n|x_{\alpha_c}}$ and $\tilde{v}(x_{\alpha_c})$ from $v^i_\alpha$ and $\tilde{v}^\ell_{\alpha_c}$. We need to search for two neighboring nodes from all the nodes $\alpha'_c$ on the contact boundary $\Gamma'_c$ of body II. Suppose that $m_1(\alpha_c)$ and $m_2(\alpha_c)$ are the two neighboring nodes on $\Gamma'_c$ of body II between which, on the $x$ axis, we can locate a body I node $\alpha_c$ thus

$$\tilde{x}_{m_1(\alpha_c)} < x_{\alpha_c} < \tilde{x}_{m_2(\alpha_c)}$$  (2.95)

Define a parameter $\theta(\alpha_c)$ to describe the relative position of $\tilde{x}_{m_1(\alpha_c)}$, $\tilde{x}_{m_2(\alpha_c)}$ and $x_{\alpha_c}$.

$$\theta(\alpha_c) = \frac{\tilde{x}_{m_2(\alpha_c)} - x_{\alpha_c}}{\tilde{x}_{m_2(\alpha_c)} - \tilde{x}_{m_1(\alpha_c)}}$$  (2.96)

Using $\theta(\alpha_c)$, $\psi(x_{\alpha_c})$ and $\tilde{v}(x_{\alpha_c})$ are approximated as

$$\psi(x_{\alpha_c}) = \psi(\tilde{x}_{m_1(\alpha_c)})\theta(\alpha_c) + \psi(\tilde{x}_{m_2(\alpha_c)})\left(1 - \theta(\alpha_c)\right)$$  (2.97)

$$\tilde{v}(x_{\alpha_c}) = \tilde{v}_{m_1(\alpha_c)}\theta(\alpha_c) + \tilde{v}_{m_2(\alpha_c)}\left(1 - \theta(\alpha_c)\right)$$  (2.98)

The optimization proceeds in a manner similar to that in the last section, but our implementation of this method does not give realistic results in some meshes. The problem is most likely that the contact surface nodes of the two surfaces are not distributed evenly for some meshes. It is usual to have many surface nodes of body I sitting between two neighboring contact surface nodes of solid body II, so the constraints (2.94) will affect the relation between the neighboring nodes of solid I, This is not expected from the theory and will very likely change the nature of the problem. Eventually we gave up this direction.

Our approach treats the problem of two elastic solid bodies under frictionless contact as a special case of the two elastic solid body under frictional contact with the Coulomb friction coefficient equal to zero. This will be discussed in the next section. We will provide an approach which iteratively couples the results for two elastic solid bodies taking into consideration the tangential shift between the two surfaces.

### 2.3.4 Contact Between an Elastic Solid and a Rigid Surface with Friction

The problem of frictional contact between elastic solid and rigid surface is a prerequisite procedure in our scheme of solving frictional contact between two elastic solid bodies.
Frictional effects add complexity to the problem of contact between elastic bodies. One of the difficulties in modeling frictional contact behavior is that we are still lacking in sufficient understanding of the microscopic nature of friction [38, 45]. The basic Coulomb frictional law came from the observation of a macroscopic phenomena and was designed to describe the relative motion of rigid bodies. The law basically include: i) the frictional force is proportional to the normal load; ii) the frictional force is independent of the apparent contact area of the sliding surfaces; iii) sliding friction is independent of the sliding velocity. The origin of the Coulomb frictional law may be traced back to Leonardo Da Vinci two hundred year before the foundations of classical mechanics was established. To our surprise, it has been used for several centuries. Classical mechanics and continuum mechanics developed from the time of Newton to the time of Hertz without seriously questioning the basic Coulomb frictional law. Even till now, there is no universally accepted physical explanation for this friction law. Neither is there guarantee that this law is applicable in the pointwise case. It was considered as a grey area of classical physics [53]. In the last two decades, many researcher started to investigate the friction law theoretically and experimentally. Several microscopic pictures of friction were attempted. The one considered most standard is to regard the macroscopic smooth surface as rough on the microscopic scale [26, 41, 45]. Thus the real surface is not flat and smooth but consists of many asperities. The real contact region only occurs at the tip of the asperities and the total actual contact area is smaller than the macroscopic contact area and is proportional to the compressive force between the two surfaces. Because of roughness, The area of real atomic contact between bodies is usually a very small fraction of the apparent area of contact. For example, if a cubic steel block with a side of length 10cm is placed on a steel substrate, the area of real contact is of the order 0.001cm$^2$, i.e, only a fraction $10^{-5}$ of the apparent area. Several papers from Persson [46, 47, 48] have a detailed description of the rough surface contact and the real contact area. The bulk solid deforms elastically, but plastic yielding happens at the tip of the asperities in the process of contact, eventually forming a juncture by welding the asperities of the two surfaces. Slippage of asperities occurs when the tangential traction is greater than the shear strength of the juncture, which is proportional to the actual area of contact. This microscopic model of friction will lead to a nonlocal friction law which demands that the tangential friction at
a contact surface point is dependent on the statistical average of normal traction in the
neighboring region. All these details are not completely confirmed by physicists, since the
attempts to construct a model of surface interaction at atomic scale were not successful.
Actually a very complex chemical and physical process at the atomic level occurs on the
surface during the sliding friction. Cieplak, Smithon and Robbins [49] made a conclusion
from their perturbation theory calculation that the friction of solid monolayer on a bulk
solid has a very different form from that observed between macroscopic solids. There is no
threshold force or static friction to initiate sliding. This is not confirmed yet, but a Atomic
sliding friction measured by atomic force microscope (AFM) shows that the friction force
shows sawtooth behavior and hysteresis under the reversal of the scanning direction. [50].
By using the friction force microscope (FFM) [52, 51], it is possible to examine the fric-
tional properties of an approximate point contact at the nanometer scale. The observed
frictional behavior differs significantly from the behavior expected from the macroscopic
friction laws introduced above. In particular, it is found that frictional forces are pro-
tional to the true area of contact, which is generally not proportional to the loading force.
Consequently, laws (i) and (ii) are no longer valid at the nanometer scale. but law (iii)
still holds within a wide range of sliding velocities. All these recent discoveries in atomic
level imply that the Coulomb friction law may be a comprehensive dynamic effect.

Another simple microscopic model [54] introduces an adsorbed layer of mobile atoms
between the surfaces and attributes the friction to these foreign mobile atoms trapped
between the two surfaces. These atoms lock the two surfaces in place. This model can
account for the proportionality of the friction force to the compressive force and tangen-
tial friction force being independent of normal contact area, but it generates a vanishing
coefficient of friction for incommensurate or disordered surfaces as the size of the contacts
increases. It does not include the plastic and elastic deformation of the walls.

There is also some research work attempting to incorporate the Coulomb frictional
law into the theory of elasticity. Rice [55] investigated the steady frictional sliding at the
interface of dissimilar materials with Coulomb friction acting at the interface within the
picture of continuum mechanics. It was shown by him that steady sliding with Coulomb
friction was ill-posed for an arbitrarily small value friction due to the unstable growth
of interfacial disturbances of all wavelengths. It was also reported by Caroli [56] that
Figure 2.4. (a) An illustration of a self-healing crack proposed by Gerde and Marder. Note the dimensions in the figure are enlarged relative to the real situation. (b) The triangular grids of atomistic lattice points used in the model of Gerde and Marder.

frictional sliding with the slip pulse at the interface obeying a simple local Coulomb friction law is impossible.

Recently, Gerde and Marder [38] have proposed an atomistic lattice model for friction. They have succeeded in finding solutions based on self-healing cracks traveling along the interface. These solutions may be physically reasonable, since self-healing cracks are observed in earthquakes [57]. This traveling self-healing crack is schematically shown in Figure 2.4a. The crack is basically a surface fracture taking place at the interface of two elastic media. In the model, the separation of this fracture forms a ridged shape and propagates along the interface with a certain constant speed. This type of surface fracture is different from the type of fracture created when two surfaces are pulled apart and kept far away. The two surfaces of the self-healing crack remain in close proximity and can come together again and bond or heal in their new laterally shifted positions. This re-binding sets in some distance downstream of the crack edge. In the middle of the crack, the two surfaces are separated from each other by a gap. As the crack travels, the tip of the front edge of the crack will continue to detach and the gap at the back edge will shrink and close. Once the crack has travelled through part of the block, that part of the block will be displaced, through slipping, from its original position.

Das [57] tried to explain the Coulomb friction law by such self-healing cracks traveling
through the interface by continuum mechanics. The results derived from his theory were found to be unphysical. Gerde and Marder resolved this problem by combining information from an atomistic lattice model with continuum mechanics. Their atomistic lattice model treats the contact solids as systems of lattice points, each point is filled with a classical particle. Gerde and Marder started their investigation with a two dimensional zero temperature model for simplicity. The thermal oscillation is thus ignored for each particle. Each particle is at its equilibrium position before deformation and the equilibrium positions of these particles are arranged in triangular grids as shown in Figure (2.4b). Each particle is modeled as an oscillator with a small dissipation driven by the direct coupling force between neighboring particles in the lattice. The system of equations of the lattice are derived from Newton Second Law with a small dissipation term, which is linearly proportional to the velocity of particles. The driving force is derived from the direct coupling force of the neighboring particles. Since the stress at the tip of the crack predicted by the above atomistic model shares the same form with the continuum calculation, Gerde and Marder conducted a search for possible self–healing crack solutions. There is an adjustable time parameter, $s$, representing the time difference between left bond ruptures and right bond ruptures for atoms along the interface. This parameter is separated into $s_f$ at the front tip and $s_b$ at the back tip. For a certain self–healing crack with a fixed length $l$ traveling with a velocity $v$ and causing a slip $\Delta u$, the atomistic model predicts an oscillating stress at the tip.

Using a connection with the continuum calculation, Gerde and Marder found solutions for the stresses $\sigma_{xy}(\infty)$ and $\sigma_{yy}(\infty)$ at infinity which cause this kind of traveling crack. Those stresses are derived in the continuum mechanics using the $s$ parameters and structure of the atomistic model. The authors made some artificial adjustments to $s_f$ and $s_b$ so that the stresses $\sigma_{xy}(\infty)$ and $\sigma_{yy}(\infty)$ derived relative to the front and back tips were the same. By choosing different velocities, they made an exhaustive search of all possible $\sigma_{xy}(\infty)$ and $\sigma_{yy}(\infty)$. The results suggest that the minimum $\sigma_{xy}(\infty)$ is roughly linearly proportional to $\sigma_{yy}(\infty)$, in approximate accordance with the Coulomb prediction of a linear correlation.

According to Gerde and Marder the sliding is modeled as a propagation of self–healing cracks and the speed of sliding is dependent on the number of traveling self–healing cracks.
generated at the same time. Static friction can also be explained within the context of this model using self-healing cracks. For any given compressive stress $-\sigma_{yy}(\infty)$, there is a minimum shear stress $\sigma_{xy}(\infty)$ that allows the cracks to begin. The upper block cannot slide for any shear stress smaller than $\sigma_{xy}(\infty)$. The minimum shear force $\sigma_{xy}(\infty)$ to initiate a propagating crack is roughly proportional to the compressive stress $-\sigma_{yy}(\infty)$ pushing the surfaces together. This proportionality is the Coulomb friction law. The coefficient of proportion enables another way to calculate friction coefficients.

Gerde and Marder mentioned that the validity of their calculation was checked with a molecular-dynamics calculation. Whether or not self-healing cracks actually underlie frictional sliding will have to be settled by experimentation. Real surfaces are certainly not flat enough at the microscopic scale for the self-healing cracking description of friction to be complete. A model with the assumptions of zero temperature, flat surfaces and only two dimensions is relatively easy to implement. Gerde and Marder intend to present more realistic generalizations as they develop. They are optimistic about the generalization of these positive results generated for ideal flat surfaces to more realistic rough surfaces.

So far we have discussed various microscopic models supporting the Coulomb friction law from different perspectives. The Coulomb friction law is still the most used model and is regarded as the most reliable one, and many macroscopic observations of rigid solids also support it, although there is no universally accepted explanation of this basic law of friction. We use the Coulomb friction law in our finite element model.

Friction and the contact problem are usually classified in three categories [26]: quasi-static dry friction problems, dynamical friction problem, and wear and plow problems. Our modeling is designed to target static and quasi-static dry friction problems, that means polished metallic surfaces are pressed together and displaced relative to each other slowly enough to regard each instantaneous state as a static equilibrium state. The normal loading is not big enough to produce the gouging and plowing on the surface. In the picture of the self-healing crack model, the condition for quasi-static dry friction problems requires the loading increase slowly and that kinematic motion is very slow compared to the rate of generation and propagation speed of self-healing crack.

Our modeling is an investigation based on the classical Coulomb frictional law, and more information about the nature of friction can be obtained from checking what the
model can imply. Since the Coulomb frictional law approximates the macroscopic world quite well plus numerous microscopic supporting models as mentioned above, there is no reason to abandon this frictional law. As it is a belief in physicist community that nature is more likely to behave according to the principle of simplicity and harmony, we choose this simple model as a starting point. It may be modified once a more reliable frictional law is discovered and confirmed in the future.

The friction coefficient depends on the microscopic roughness, which is related to asperities, atomic structure and chemical properties of the oxide layers of the two metal surfaces. How does the friction coefficient depend on the above elements is outside the scope of this research; here we just choose it as a positive constant number less than one.

If the strain and stress field distributions plus the boundary is known, it is easy to compute the traction force \( \sigma_n \) on the surface by Cauchy stress from the theory of elasticity

\[
\sigma_n = \sigma \cdot n \tag{2.99}
\]

where \( \sigma \) is the Cauchy stress tensor and \( n \) is the unit vector normal to the boundary surface. Normal traction can be written as \( \sigma_n n = (\sigma_n \cdot n)n \) with \( \sigma_n = \sigma \cdot n = n \cdot (\sigma \cdot n) \).

The tangential traction of the contact surface can be expressed as

\[
\sigma_t = \sigma_n - (\sigma_n \cdot n)n \tag{2.100}
\]

\( \sigma_n \) and \( \sigma_t \) are needed later in frictional contact boundary conditions. The classical formulation of the frictional contact problem using Coulomb friction is

\[
\frac{\partial}{\partial x_j} \sigma_{ij}(u) + f_i = 0 \tag{2.101}
\]

where \( \sigma_{ij}(u) = E_{ijkl}u_{k,l} \) and the boundary conditions are

\[
u_i = 0 \text{ on } \Gamma_d \tag{2.102}
\]

\[
\sigma_{ij}n_j = t_i \text{ on } \Gamma_F \tag{2.103}
\]

\( \Gamma_d \) is the boundary of the displacement constraint, and \( \Gamma_F \) is the boundary of the traction. In addition, there are frictional contact conditions on \( \Gamma_c \), the contact boundary:

\[\text{if } u_n < g_n \text{ on } \Gamma_c, \sigma_n(u) = 0 \text{ and } \sigma_t(u) = 0;\]
if \( u_n = g_n \) on \( \Gamma_c \), \( \sigma_n (u) < 0 \) \( u_t = 0 \) when \( |\sigma_t| < -\mu (\sigma_n \cdot n) \)
and there exists a real number \( \lambda > 0 \), such that
\[
u_t = -\lambda \sigma_t \text{ when } |\sigma_t| > -\mu (\sigma_n \cdot n)\]

This contact condition means that the normal traction in the contact region has to be compressive, and the traction in the no contact region of the boundary has to be zero. The impending sliding only occurs when the ratio of the tangential traction over the normal traction is greater than the friction coefficient. \( \lambda \) is a real number relating the relative sliding displacement of the surface to the tangential traction. The difficulty of this formulation that prevents us from solving it is that the contact boundary has two main unknown regions, one of which satisfies \( u_n < g_n \) on the \( \Gamma_c \), the other of which satisfies \( u_n = g_n \) on \( \Gamma_c \). The free boundary between them is also unknown. Even within the second region, there are two subregions undefined with unknown free boundaries between them, one is no slipping region with
\[
|\sigma_t| < -\mu (\sigma_n \cdot n) \quad \text{(2.104)}
\]
the other region has slippage and satisfies
\[
|\sigma_t| = -\mu (\sigma_n \cdot n) \quad \text{(2.105)}
\]
Without specifying the free boundary of these regions, there seems very little hope to solve the above equation (2.101) analytically or numerically. If we use a variational approach, however which is a weak formulation of boundary problem, it can be solved numerically without knowing the boundary. It is proved in [22] that the solution of the above frictional contact problem is equivalent to the variational inequality formulation. We introduce in chapter 1 that if \( u \) is the displacement field of the loaded solid body under frictional contact, the inequality will be satisfied
\[
a (u, v - u) + J (u, v) - J (u, u) \geq f (v - u) \quad \text{(2.106)}
\]
for any \( v \in \mathbb{K} \), \( \mathbb{K} = \{ v : v = 0 \text{ on } \Gamma_d \text{ and } u_n - g_n \leq 0 \text{ on } \Gamma_c \} \). This inequality stipulates that any virtual work done by an external force is not sufficient to overcome the virtual
elastic potential energy and frictional potential energy increase on the boundary \( \Gamma_c \), that is
\[
a(u, \delta v) + J(u, \delta v) \geq f(\delta v) \tag{2.107}
\]

Equation (2.106) is proved in [26] to be equivalent to the frictional contact problem, but there is no general proof of the existence and uniqueness of the solution regarding this variational inequality. This is due to the behavior of functional \( J(u, v) \), which is not differentiable, nor has any direct numerical method been presented previously to solve this problem. All previous solutions rely on the hit and miss technique. One way to solve this problem is to divide it into two simple special cases, and solve it by an iterative method. This is the method our modeling follows. The procedure for dealing with these two special cases follows:

**Special Case I:** The tangential stress \( \sigma_t \) on the contact is given, while the normal traction is unknown. The problem is then a special Signorini problem satisfying the inequality
\[
a(u, v - u) \geq f_1(v - u) \tag{2.108}
\]
with
\[
f_1(v) = f_0^i u_0^i + \int_{\Gamma_c} \sigma_t v_t ds \tag{2.109}
\]

We choose to evaluate the integration \( \int_{\Gamma_c} \sigma_t v_t ds \) on the contact region by dividing the contact region into units, then summing the integrand over all the units. We choose to divide each contact element into \( p' \) parts, then evaluate \( \int_{\Gamma_c} \sigma_t v_t ds \) by summation over all these \( p' \) units. All the real contact elements contribute to this integral. The higher \( p' \) is, the more accurate the summation is. \( n_1, n_2 \) and \( \sigma_t \) are computed from the \( v \) of previous special case II problem, which is discussed in next section. At the first step when the computation starts, the variable \( \sigma_t \) is chosen as zero, and \( n_1, n_2 \) are computed from \( v = 0 \).

The tangential displacement is
\[
v_t = v \cdot n_t \tag{2.110}
\]
where \( n_t = [-n_2, n_1] \), while \( v_x, v_y \) are computed from the current \( u_0^i \). The solution of (2.108) is equivalent to solving the minimization problem
\[
\min \left[ \frac{1}{2} a(v, v) - f_1(v) \right] \forall v, v \cdot n - g_n \leq 0 \text{ on } \Gamma_c \tag{2.111}
\]
written in terms of finite element methods

$$\min \left( \frac{1}{2} \v^T K \v - F_1^T \v \right) \forall \v, \ A\v \leq g \text{ on } \Gamma_c$$

(2.112)

We can solve this by the same method as the Signorini problem. After special case I is solved, the real contact area $\Gamma_{rc}$ where $n \cdot v - g_n = 0$ is known, then we can move to the second part of the procedure.

**Special Case II.** It starts with the results of special case I, assuming the real contact surface $\Gamma_{rc}$ to be fixed and that normal stress $\sigma_n$ as well as $n_1$ and $n_2$ on it are known from the $\psi^i$ of previous results. The problem is treated with the known normal stress from the computational results of special case I, so $J(u, v)$ becomes

$$j(v) = -\int_{\Gamma_c} \mu (\sigma_n \cdot n) |v| ds$$

(2.113)

and (1.61) becomes

$$a(u, v - u) + j(v) - j(u) \geq f_2(v - u), \forall v$$

(2.114)

where

$$f_2(v) = f^i_2 v^i + \int_{\Gamma_c} \sigma_n u_n ds$$

(2.115)

Solving (2.114) is equivalent to solving the minimization problem

$$\min \frac{1}{2} [a(v, v) - f_2(v) + j(v)]$$

(2.116)

for $\forall v$ with $n \cdot v - g = 0$ on $\Gamma_c$. In order to make (2.116) differentiable, we modify $j(v)$ by regularizing $|v|

$$|v| \rightarrow \phi(v) = \begin{cases} 
-\frac{v - \frac{1}{2}}{2\epsilon} & v > \epsilon \\
\frac{1}{2\epsilon} v^2 & |v| \leq \epsilon \\
-\frac{v - \epsilon}{2} & v < -\epsilon
\end{cases}$$

(2.117)

so

$$\frac{\partial \phi(v)}{\partial v} = \begin{cases} 
1 & v > \epsilon \\
\frac{1}{\epsilon} v & |v| \leq \epsilon \\
-1 & v < -\epsilon
\end{cases}$$

(2.118)

Suppose the contact nodes are $\alpha_1^c, \alpha_2^c, \ldots, \alpha_n^c$, and we choose to divide each contact element into $p'$ parts and then evaluate $f_2(v)$ and $j(v)$ by summation over all these $p'$ units,
eventually all the contact elements are counted. The finer the divisions the more accurate
the integration is

\[
\begin{align*}
\psi (v) &= \int_{\Gamma_c} -\mu (\sigma_n \cdot n) |v_r| ds \\
&= \sum_{ce} \left( \sum_{p=1}^{P'} [-\mu (\sigma_n \cdot n)]_p l_p \phi (v_{pp}) \right)
\end{align*}
\]

\(\sum_{ce}\) is actually a summation over the real contact elements. The components \(n_x|_p\) and
\(n_y|_p\) of \(n|_p\) are calculated by spline interpolation from the results of \(v\) in special case I. In
order to evaluate \(\psi (v)\), we evaluate \(|v|_p\) by

\[
\phi (|v|_p) = \phi (|v - (v \cdot n) n|_p)
\]

\[
= \begin{cases} 
|v - (v \cdot n) n|_p - \frac{\varepsilon}{2} & |v - (v \cdot n) n|_p \cdot n > \varepsilon \\
\frac{1}{2 \varepsilon} |v - (v \cdot n) n|_p^2 & |v - (v \cdot n) n|_p \leq \varepsilon \\
-|v - (v \cdot n) n|_p - \frac{\varepsilon}{2} & |v - (v \cdot n) n|_p \cdot n < -\varepsilon
\end{cases}
\]

with finite element approximation

\[
\pi (u) = \left( \frac{1}{2} v^T K v - F^T_2 (v) + j (v) \right)
\]

subject to the equality constraint at the node \(\alpha_c\) of real contact boundary \(\Gamma_{rc}\)

\[
n_{1,\alpha}^c \cdot u_{1}^c + n_{2,\alpha}^c \cdot u_{2}^c - g_n|_{\alpha_c} = 0
\]

and the equality constraints \(u_{\alpha_1}^1 = U_{\alpha_1}^1\) and \(u_{\alpha_2}^2 = U_{\alpha_2}^2\), where \(\alpha_1\) denote node which
have the displacement constraint in the \(x\) component, \(\alpha_2\) for the node which have the
displacement constraint in the \(y\) component, and \(\alpha_c\) for the real contact node satisfying
(2.122). The first two terms of \(\pi (u)\) can be written as

\[
\frac{1}{2} v^T K v = \sum_{\alpha} \sum_{\beta} \sum_{i=1}^{2} \sum_{k=1}^{2} \frac{1}{2} E^{ik} v_{\alpha}^\beta v_{\alpha}^\alpha
\]

\[
F^T_2 (u) = f^\alpha_{\alpha} u_{\alpha}^\alpha + \sum_{ce} \left[ \sum_{p=1}^{P} [- (\sigma_n \cdot n)]_p |v \cdot n|_p l_p \right]
\]
Eventually we need to minimize the following functional \( F(v) \) with a very large penalty coefficient to find the solution for special case II.

\[
F(v) = \left( \frac{1}{2} v^T K v - F_2^T(v) + j(v) \right)
\]
\[
+ \sum_{\alpha\epsilon} c_p \left( n_1^{\alpha\epsilon} \cdot v_1^{\alpha\epsilon} + n_2^{\alpha\epsilon} \cdot v_2^{\alpha\epsilon} - g_{n\alpha\epsilon} \right)^2
\]
\[
+ c_p \sum_{\alpha_1} (v_1^{\alpha_1} - U_1^{\alpha_1})^2 + c_p \sum_{\alpha_2} (v_2^{\alpha_2} - U_2^{\alpha_2})^2
\]

from this \( \nabla_v F(v) \) can be calculated analytically

\[
\frac{\partial \left[ \frac{1}{2} v^T K v \right]}{\partial v_{\alpha\epsilon}'} = \sum_{\alpha} \sum_{i=1}^2 \frac{1}{2} E_{\alpha\epsilon}^i v_i^\alpha + \sum_{\beta} \sum_{k=1}^2 \frac{1}{2} E_{\alpha\beta}^k v_k^\beta
\]

\[
\frac{\partial F_2^T(v)}{\partial v_{\alpha\epsilon}'} = \left[ f_{d\delta} \delta_{\alpha\epsilon} \delta_{iv} + \sum_{\epsilon p} \sum_{p=1}^{p'} \left( \sigma_{\alpha \epsilon} \cdot n \right)_p \left( \frac{\partial v_{x}}{\partial v_{\alpha\epsilon}'} n_1 + \frac{\partial v_{y}}{\partial v_{\alpha\epsilon}'} n_2 \right)_p \right]
\]

\[
\frac{\partial j(v)}{\partial v_{\alpha\epsilon}'} = \sum_{\epsilon p} \left( \sum_{p=1}^{p'} \left( \sigma_{\alpha \epsilon} \cdot n \right)_p \frac{\partial \phi(v_{\epsilon p\epsilon})}{\partial v_{\alpha\epsilon}'} \right)
\]

The natural coordinates \( L = (L_1, L_2, L_3) \) are used to locate the boundary point \( p \) depending on what kind of contact element it is on. This is shown in Figure (2.5). If it is on the side of \( jk \),

\[
[L_1]_p = 0, [L_2]_p = 0.1
\]

50
If it is on the side of $ik$,
\[ [L_1]_p = 0, [L_2]_p = 0 \quad (2.129) \]

If it is on the side of $ij$,
\[ [L_1] = 0, [L_2]_p = 1 - [L_1]_p \quad (2.130) \]

The displacement of $p$ is $[v_x]_p$ and $[v_y]_p$:
\[ [v_x]_p = v_1^i [N_1]_p + v_1^j [N_j]_p + v_1^k [N_k]_p + v_1^l [N_l]_p + v_1^m [N_m]_p + v_1^n [N_n]_p \quad (2.131) \]
\[ [v_y]_p = v_2^i [N_1]_p + v_2^j [N_j]_p + v_2^k [N_k]_p + v_2^l [N_l]_p + v_2^m [N_m]_p + v_2^n [N_n]_p \quad (2.132) \]

where the $[N_1]_p, [N_j]_p, [N_k]_p, [N_l]_p, [N_m]_p, [N_n]_p$ are the shape functions evaluated at the point $p$. The $i, j, k, l, m, n$ label the 6 nodes of the element in which the boundary point is located. The reduced problem can be solved by an unconstrained optimization method.

We are aware that the regularization of $|v|$ is written as (2.133) in [26, 29]. This function is discontinuous.
\[ \phi (v) = \begin{cases} 
 v - \frac{\varepsilon}{2} & v > \varepsilon \\
 \frac{1}{\varepsilon^2} v^2 & |v| \leq \varepsilon \\
 -v + \frac{\varepsilon}{2} & v < -\varepsilon 
\end{cases} \quad (2.133) \]

If this is not a printing error, it will cause errors in the computation results.

### 2.3.5 Contact between Elastic Solids with Friction

Once an algorithm to solve frictional contact between an elastic solid and a rigid surface is ready, we are able to solve frictional contact between two elastic solids by the iterative method plus the relaxation method. Since it is a classical problem to find the displacement field of a solid if all the boundary condition are known, We can solve it by a normal finite element method or by minimization of total potential energy with a penalty method all subject to the equality constraints,
\[ \min \left[ \frac{1}{2} E^{ik} \alpha^i \nu^i \nu^j - f_{ij} \nu^j \right] \quad (2.134) \]

with
\[ \nu^{i0} = U_1^0 \quad \text{and} \quad \nu^{i2} = U_2^0 \quad \text{on the} \, \Gamma_d \quad (2.135) \]
where
\[ f_\alpha^* = \int_{\Omega^k} f_i \phi_\alpha \, dx + \int_{\Gamma^k_p} t_i \phi_\alpha \, ds \]  \hspace{1cm} (2.136)

Since there is no inequality constraint, it is much faster to solve this kind of problem than to solve the Signorini problem. It can be done with part of the routine of the Signorini problem. For the convenience, it will be referred as the 'regular elasticity problem' or the 'solid II (\(\Omega_2\)) problem' in the following paragraph.

We present an alternative approach to this problem. The motivation for this approach is addressed in [63, 64]. The loading force applied to the solid bodies is increased incrementally, and the incremental step is small enough to treat the state of systems at each step as an equilibrium state in the process of increasing loading from zero to its final magnitude. This assumption does not work for dynamic problems, but it is reasonable for static or quasi-static problems where equilibrium is approximately maintained through a slow increase in loading.

The major part of the scheme is the same as Kikuchi [26] and Refaat [29]. The finite element method is used to implement the variational inequality to study the two dimensional frictional contact problem. The loading is increased by increasing the surface traction density incrementally. The solution procedure for each loading step is made up of many cycles of the following four steps:

1. Regard \(\Omega_2\) as a rigid body and solve the frictional contact problem of an elastic solid \(\Omega_1\) in contact with a rigid surface, which is the contact boundary \(\Gamma_2\) of \(\Omega_2\). Here, \(\Gamma_2\) is represented by the function \(\psi(x)\) initially. There is no displacement of surface particles on \(\Gamma_2\), as it is regarded as rigid. This procedure is the same as the last section and involves searching for \(\sigma_n\) and \(\sigma_t\) iteratively. Once the solution of displacement in \(\Omega_1\) is obtained, it is easy to compute the stress, then find the traction on the real contact area of surface \(\Gamma_{c1}\).

2. Calculate the traction force acting on contact boundary \(\Gamma_2\) of solid \(\Omega_2\). It has the same magnitude as the traction force acting on the real contact area of the surface \(\Gamma_{c1}\) by solid \(\Omega_2\), but it is in the opposite direction in accordance with Newton's Third Law.
3. Look upon $\Omega_2$ as an independent elastic solid with boundary displacement constraints on $\Gamma_{c2}$ and tractions on $\Gamma_{F2}$, then use the surface traction force acting on the $\Gamma_{c2}$ boundary, as calculated in the previous step and apply the method of the 'regular elasticity problem' as mentioned above to solve the displacement field of $\Omega_2$.

4. The final step is to renew the function representing $\Gamma_{c2}$ by the solution of the displacement of $\Omega_2$ in the last step.

We repeat this four step procedure until the solutions of the displacement field converge, then increase the loading by another step and start the four step procedure again. This continues until the final load value is reached and convergent solutions are obtained at that loading. With this scheme, each solid is computed independently. The solutions of the two solids are connected by the surface traction and the surface deformation. Solid $\Omega_2$ is regarded as rigid and elastic alternately. Since the second solid is regarded as rigid however in solving the first step, the relative tangential displacement of particles on the two contact surfaces is evaluated as the difference between consecutive iterations of tangential displacement of the surface particles on $\Omega_1$, i.e $v_{\varepsilon}^{(n)} - v_{\varepsilon}^{(n-1)}$. The convergence of the displacement field is tested using

$$\varepsilon^{(n)} = \frac{\|v^{(n+1)} - v^{(n)}\|}{\|v^{(n)}\|}$$

and iteration continues until

$$\varepsilon^{(n)} < \varepsilon_0$$

where $\varepsilon_0$ can be chosen as close to zero as reasonably possible, depending on the accuracy and time limits of computation.

Practice suggests this method converges most time, if the maximum strain estimated from Hertz Theory is less 1%. This strain is greater than the elastic strain limit(0.2%) for most metal materials. The results show that tangential shift plays an important role in frictional contact between two elastic bodies. The examples we have used to conduct numerical experiments were chosen to allow comparison with existing theoretical results. A stable solution exists, but It is shown that this solution is sensitive to many conditions such as loading step number and loading incremental step, boundary conditions, total geometry and dimension of solids. Such factors have been ignored by Hertz theory and
other subsequent works. Since this scheme is only designed for elastic frictional contact problems whose final loading force is reached by a quasi-static process, we choose to use the words 'the step numbers of loading and loading increment of each step' instead of loading history. Load history implies a transient process, which is associated with dynamic problems. We cannot regard the state of solid as equilibrium at every time point of this transient process for the dynamic problem. This is out of the scope of our research.
Chapter 3

Numerical Utilities

3.1 Introduction

The computational schemes of previous chapters are implemented by making use of some numerical tools. This chapter introduces all the numerical tools we have employed. They include

- Cubic spline interpolation.
- A conjugated gradients method with restarting for solving unconstrained optimization problems.
- A multiplier method for solving constrained optimization problems.
- A penalty method for solving constrained optimization problems.
- A relaxation method to guarantee a stable convergence.

3.2 Cubic Spline Method (1-D)

To describe the geometry of the surface using the coordinates of the nodes on the contact surface, spline interpolation method is used to generate a smooth function approximating the surface curve from the coordinates of a certain number of surface node points. The particular spline method used in current modeling is a cubic spline \[58\]. Cubic spline function is a spline function \(s(x)\) of order 4. That means it satisfies two conditions for a grid \( \{a = x_0 < x_1 < x_2 < \ldots < x_n = b\} \).
1) $s(x)$ is a polynomial of degree less than 4 on each subinterval $[x_{i-1}, x_i]$ ($i = 1, 2...n$).

2) $s^{(r)}(x)$ is continuous on $[a, b]$ for $0 \leq r \leq 2$.

The first point means we may assume

$$s(x) = a_i + b_i x + c_i x^2 + d_i x^3$$  \hspace{2cm} (3.1)

for $x$ in $[x_{i-1}, x_i]$, ($i = 1, 2...n$). There are $4n$ unknown coefficients

$$\{a_i, b_i, c_i, d_i \text{ where } i = 1, 2, ..n\}$$  \hspace{2cm} (3.2)

To specify these $4n$ coefficients from the coordinates of the grid points, $4n$ equations are needed. $n + 1$ equations can be created from the coordinates of the grid points

$$s(x_i) = y_i \quad (i = 0, 1, 2...n)$$  \hspace{2cm} (3.3)

We obtain $n - 1$ equations from the continuity condition on $s^{(0)}(x)$ at grid points

$$s^{(0)}(x_i + 0) = s^{(0)}(x_i - 0) \quad \text{where} \quad (i = 1, 2...n - 1)$$  \hspace{2cm} (3.4)

We obtain $n - 1$ equations from the continuity condition on $s^{(1)}(x)$

$$s^{(1)}(x_i + 0) = s^{(1)}(x_i - 0) \quad \text{where} \quad (i = 1, 2...n - 1)$$  \hspace{2cm} (3.5)

We obtain $n - 1$ equations from the continuity condition on $s^{(2)}(x)$

$$s^{(2)}(x_i + 0) = s^{(2)}(x_i - 0) \quad \text{where} \quad (i = 1, 2...n - 1)$$  \hspace{2cm} (3.6)

Thus the degree of freedom left for this system is $4n - 3(n - 1) - (n + 1) = 2$. There are several choices to restrict these two degrees of freedom. The method used in this modelling to restrict these two degrees of freedom is called ‘Not a knot’. We introduce coefficients

$$M_i = s''(x_i)$$  \hspace{2cm} (3.7)

for $(i = 0, 1, 2...n)$. The function $s''(x)$ is linear because $s(x)$ is cubic polynomial, so it is straightforward to interpolate $s''(x)$ by $M_i$

$$s''(x) = \frac{(x_{i+1} - x)}{h_i} M_i + \frac{(x - x_i)}{h_i} M_{i+1}.$$  \hspace{2cm} (3.8)
for \( x \in [x_i, x_{i+1}] \), \( i = 0, 1, 2, ... n - 1 \), where \( h_i = x_{i+1} - x_i \). It is easy to find \( s(x) \) and \( s'(x) \) by integration with two unspecified constant \( C \) and \( D \).

\[
s'(x) = -\frac{1}{2} (x_{i+1} - x)^2 \frac{M_i}{h_i} + \frac{1}{2} (x - x_i)^2 \frac{M_{i+1}}{h_i} + C \tag{3.9}
\]

\[
s(x) = \frac{1}{6} (x_{i+1} - x)^3 \frac{M_i}{h_i} + \frac{1}{6} (x - x_i)^3 \frac{M_{i+1}}{h_i} + C \left( x_{i+1} - x \right) + D \left( x - x_i \right) \tag{3.10}
\]

Substituting \( s(x_i) = y_i \) and \( s(x_{i+1}) = y_{i+1} \) reduces (3.10) to

\[
C = \frac{y_i}{h_i} - \frac{h_i M_i}{6} \quad \text{and} \quad D = \frac{y_{i+1}}{h_i} - \frac{h_i M_{i+1}}{6} \tag{3.11}
\]

so the first derivative in \([x_i, x_{i+1}]\) is

\[
s'(x) = \frac{-(x_{i+1} - x)^2 M_i + (x - x_i)^2 M_{i+1}}{2h_i} + \left[ \frac{y_{i+1}}{h_i} - \frac{y_i}{h_i} \right] \tag{3.12}
\]

\[
+ \frac{h_i}{6} \left[ M_i - M_{i+1} \right]
\]

and the first derivative in \([x_{i-1}, x_i]\) is

\[
s'(x) = \frac{-(x_i - x)^2 M_{i-1} + (x - x_{i-1})^2 M_i}{2h_{i-1}} + \left[ \frac{y_i}{h_i} - \frac{y_{i-1}}{h_{i-1}} \right] \tag{3.13}
\]

\[
+ \frac{h_{i-1}}{6} \left[ M_{i-1} - M_i \right]
\]

Since \( s'(x_i^-) = s'(x_i^+) \)

\[
\frac{h_{i-1}}{6} M_{i-1} + \left( \frac{h_i}{3} + \frac{h_{i-1}}{3} \right) M_i = \frac{h_i}{6} M_{i+1}
\]

\[
= \left[ \frac{y_{i-1}}{h_{i-1}} - \frac{y_i}{h_i} \right] - \left[ \frac{y_i}{h_{i-1}} - \frac{y_{i-1}}{h_i} \right] \quad \text{for} \ i = 1, 2, ... n - 1.
\]

Thus we have \( n - 1 \) equations for \( n + 1 \) unknown \( M_i(i = 0, 1, 2..n) \). Since

\[
s''(x) = \frac{M_{i+1} - M_i}{h_i} \tag{3.15}
\]

for \( x \in [x_i, x_{i+1}] \), we obtain two new equations by imposing two new conditions that the third derivatives at \( x_1 \) and \( x_{n-1} \) is continuous. From

\[
s'''(x_i^-) = s'''(x_i^+),
\]

we get

\[
-M_0 + \left( 1 + \frac{h_0}{h_1} \right) M_1 - \frac{h_0}{h_1} M_2 = 0. \tag{3.17}
\]
From

\[ s''(x_{n-1}^0) = s''(x_{n-1}^+0) \]  

we get

\[ -M_{n-2} + \left(1 + \frac{h_{n-2}}{h_{n-1}}\right) M_{n-1} - \frac{h_{n-2}}{h_{n-1}} M_2 = 0 \]  

(3.19)

Solving the system of equations (3.14), (3.17) and (3.19) leads to a tridiagonal problem

\[ AM = D, \]

where

\[
A = \begin{pmatrix}
-1 & 1 + \frac{h_0}{h_1} & -\frac{h_0}{h_1} & 0 & 0 & 0 \\
\frac{h_0}{6} & \frac{h_1 + h_0}{3} & \frac{h_1}{6} & 0 & 0 & 0 \\
0 & \ldots & \ldots & \ldots & 0 & 0 \\
0 & 0 & \frac{h_{n-1}}{6} & \left(\frac{h_0}{3} + \frac{h_{n-1}}{3}\right) & \frac{h_0}{6} & 0 \\
0 & 0 & 0 & \ldots & \ldots & \ldots \\
0 & 0 & 0 & -1 & \cdot 1 + \frac{h_{n-2}}{h_{n-1}} & -\frac{h_{n-2}}{h_{n-1}}
\end{pmatrix}
\]  

(3.20)

\[
M = \begin{pmatrix}
M_0 \\
M_1 \\
\ldots \\
M_i \\
\ldots \\
M_n
\end{pmatrix}
\]

\[
D = \begin{pmatrix}
0 \\
\frac{1}{h_1}(y_2 - y_1) - \frac{1}{h_0}(y_1 - y_0) \\
\ldots \\
\frac{1}{h_i}(y_{i+1} - y_i) - \frac{1}{h_{i-1}}(y_i - y_{i-1}) \\
\ldots \\
0
\end{pmatrix}
\]  

(3.21)

Once \( M \) is obtained by Gauss elimination methods, the cubic spline function \( s(x) \) at any interval of \([x_i, x_{i+1}]\) for \( i = 0, 1, 2, \ldots n - 1 \), and it's first derivative \( s'(x) \) can be calculated as

\[
s(x) = \frac{(x_{i+1} - x)^3 M_i + (x - x_i)^3 M_{i+1}}{6h_i}
+ \frac{1}{h_i} [y_i (x_{i+1} - x) + y_{i+1} (x - x_i)]
- \frac{h_i}{6} [M_i (x_{i+1} - x) + M_{i+1} (x - x_i)]
\]  

(3.22)

\[
s'(x) = \frac{-3 (x_{i+1} - x)^2 M_i + 3 (x - x_i)^2 M_{i+1}}{6h_i}
+ \frac{1}{h_i} [y_{i+1} - y_i] - \frac{h_i}{6} [M_{i+1} - M_i]
\]  

(3.23)
3.3 The Conjugated Gradient Method

Most computational time is spent on the calculation of a non-constrained optimization problem, so it is very important to choose an efficient numerical method to solve this problem. A conjugated gradient method (CGM) is employed to optimize the above quadratic functional (2.30). The detail and theoretical basis for the (CGM) can be found in [58, 59]. Here we show how it is used in the current project. The basic formulation of the problem is to solve

$$\min \{ f(x) \} \text{subject to } x \in R^n \quad (3.24)$$

by marching step by step along a conjugated gradient direction. \( f(x) \) is usually a quadratic function as

$$f(x) = \frac{1}{2} x^T A x + b^T x + c \quad (3.25)$$

where \( A \) is a positive definite matrix. Since the gradient \( \nabla_x f(x) \) of a function \( f(x) \) can be calculated by analytical methods, the conjugated gradient direction is defined as

$$d_k = -\nabla f(x_k) + \beta_k d_{k-1} \quad (3.26)$$

where \( \beta_k \) is defined as

$$\beta_k = \frac{[\nabla f(x_k)]^T [\nabla f(x_k)]}{[\nabla f(x_{k-1})]^T [\nabla f(x_{k-1})]} \quad (3.27)$$

Each \( d_k (k = 1, 2, \ldots) \) is in conjugacy with every other one. This means \( d_i^T A d_i = 0 \) for \( i \neq j \). The next search point can be found by the line search method, which searches for the real number \( \alpha_k \) which minimizes \( f(x_k + \alpha_k d_k) \). For a thorough introduction to line search methods see [58, 59]. The above procedure is repeated until the condition

$$||\nabla f(x_k)|| < \varepsilon \quad (3.28)$$

is satisfied, where \( \varepsilon \) is a criterion for stopping the iterations. To reach an exact optimization, \( \varepsilon \) should be zero theoretically, but it is impossible to reach zero due to errors in digital calculations. We can still reach a very small number within a realistic time period. For example it takes about less than one hour to finish a optimization problem of dimension 800. Theoretically, marching \( n \) steps along conjugated gradient direction can reach the minimum, but it is impossible in reality due to the resolution of digital machines, which
truncates decimals. The more step the iteration, the more loss of the conjugacy in $d_k$. When $||\nabla f(x_k)||$ is close to zero, the convergence rate deteriorates, $||\nabla f(x_k)||$ starts to wander around and even goes up. However there is a way called restart procedure to improve it. There are many choices of restart procedure [59, 60]. The one chosen for this modeling is that when $k > n$ the program restarts with a steepest descent method, that is to search for the minimizer $x_{k+1}$ of the functional by a line search

$$\min_{\alpha_k} f(x_k - \alpha_k \nabla f(x_k))$$  \hspace{1cm} (3.29)

The numerical experiment suggests that this restart procedure is effective. Usually it just takes one or two restarts for the work to be finished. The specific line search method chosen also affects the computation speed and the convergence rate. There are various line search methods to choose, the method used in this model is called ‘Brent method’ [58]. The Brent method requires less function evaluation than the most widely used Golden search method if the function is smooth enough. The Golden search is known for the smallest number of steps to reach minimum. The Brent method makes use of quadratic interpolation to approximate the function by a quadratic function through three points, then uses the minimum of this quadratic function to approximate the minimum of the function. Suppose $g(x)$ is the function to be optimized, and $x^*$ is the real location which optimizes the function $g(x)$. If a quadratic function goes through $(a, g(a)), (b, g(b))$ and $(c, g(c))$, then

$$\lambda = \frac{1}{2} \frac{g(a) (c^2 - b^2) + g(b) (a^2 - c^2) + g(c) (b^2 - a^2)}{g(a) (c - b) + g(b) (a - c) + g(c) (b - a)}$$  \hspace{1cm} (3.30)

is the optimizer of the quadratic function, which is the first approximation to $x^*$. If $\lambda < b$ then substitute $(a, \lambda, b)$ for $(a, b, c)$ and repeat the above procedure. If $\lambda > b$ then substitute $(b, \lambda, c)$ for $(a, b, c)$ and repeat the above procedure. Our numerical experiment convinces us that the Brent method is quite efficient if the function is smooth in the neighborhood regions of $x^*$. The Golden search method is more efficient however for functionals which are not smooth enough for the Brent method.
3.4 A Multiplier Method for the Constraint Optimization Problem

A multiplier method is used to solve the constrained optimization problem. The original multiplier method for an equality constrained problem (ECP) is formulated as

$$\min f(x) \text{ subject to } h(x) = 0$$  \hspace{1cm} (3.31)

where $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ are given functions, noting that

$$h = (h_1(x), h_2(x), \ldots, h_m(x))$$  \hspace{1cm} (3.32)

The vector $x^*$ is a strict local minimum and a regular point of $f(x)$ if $x^*$ and its associated $\lambda^*$ satisfy

$$x'\nabla^2_{xx}L_0(x^*, \lambda^*) z > 0$$  \hspace{1cm} (3.33)

for all $z \neq 0$, with $\nabla h(x^*)' z = 0$, where $L_0 = f(x) + \lambda' h$. The converse is also true.

We just provide the procedures of multiplier method in here; for more detail see [59].

Let $L_c$ be the augmented Lagrangian function

$$L_c(x, \lambda) = f(x) + \lambda' h(x) + \frac{1}{2}c |h(x)|^2$$  \hspace{1cm} (3.34)

If the parameters $\lambda$ and $c$ are fixed, we can optimize $L_c(x, \lambda)$ by (CGM) method. Starting from an initial multiplier parameter $\lambda = \lambda_0$, and an initial penalty $c = c_0$, we continue the following iterations. In every iteration, given a multiplier vector $\lambda_k$ and a given penalty parameter $c_k$, we can minimize $L_{ck}(x_k, \lambda_k)$ by (CGM) to find a $x^*_k$, then renew $\lambda$ by

$$\lambda_{k+1} = \lambda_k + c_k h(x^*_k)$$  \hspace{1cm} (3.35)

and rechoose $c_{k+1}$ with $c_{k+1} \geq c_k$ at each iteration. The optimization point can be approached very closely without increasing $c_k$ to infinity by iterations of the above algorithm.

The problem involved in contact problems is however an inequality constraint problem. The multiplier method for inequality constraint problem is similar to the method for equality constraint problem. The problem is formulated as

$$\min f(x) \text{ subject to } h(x) = 0, \ g(x) \leq 0$$  \hspace{1cm} (3.36)
where \( f : \mathbb{R}^n \rightarrow \mathbb{R}, \quad h : \mathbb{R}^n \rightarrow \mathbb{R}^m \) and \( g : \mathbb{R}^n \rightarrow \mathbb{R}^r \) are given functions \( h = (h_1, h_2, \ldots, h_m), \quad g = (g_1, g_2, \ldots, g_r) \). The inequality constraint problem can be converted into an equality constraint problem by introducing \( z = (z_1, z_2, \ldots, z_r) \) such that
\[
\min f(x) \quad \text{subject to} \quad h_1(x) = \ldots = h_m(x) = 0,
\]
with \( g_1(x) + z_1^2 = g_2(x) + z_2^2 = \ldots = g_r(x) + z_r^2 = 0 \)
\[\text{(3.37)}\]

From this, the augmented Lagrangian function is
\[
\bar{L}_e(x, \lambda, z, \mu) = f(x) + \lambda' h(x) + \frac{1}{2} c |h(x)|^2 + \sum_{j=1}^{r} \left\{ \mu_j \left[ g_j(x) + z_j^2 \right] + \frac{1}{2} c |g_j(x) + z_j^2|^2 \right\}
\]
\[\text{(3.38)}\]

There is an easy way of minimizing (3.38). The minimization can be carried out with respect to \( z \) for a fixed point \((x, \lambda, \mu)\) first, which is equivalent to minimizing
\[
\min_{u_j \geq 0} \left\{ \mu_j [g_j(x) + u_j] + \frac{1}{2} c |g_j(x) + u_j|^2 \right\}
\]
\[\text{(3.39)}\]

first with respect to a \( u_j \); the solution is
\[
\dot{u}_j = - \left[ (\mu_j/c) + g_j(x) \right]
\]
\[\text{(3.40)}\]

Since there is a restriction \( u_j = z_j^2 \geq 0 \), the exact solution of problem is
\[
u_j^* = \max \{ 0, -([\mu_j/c] + g_j(x)) \}
\]
\[\text{(3.41)}\]

Then
\[
g_j(x) + u_j^* = \max \{ g_j(x), -([\mu_j/c]) \}
\]
\[\text{(3.42)}\]

Define \( g_j^*(x, \mu_j, c) = g_j(x) + u_j^* = \max \{ g_j(x), -([\mu_j/c]) \} \), and by extensions
\[
g^*(x, \mu, c) = \begin{bmatrix}
  g_1^*(x, \mu_1, c) \\
  \vdots \\
  g_r^*(x, \mu_r, c)
\end{bmatrix}
\]
\[\text{(3.43)}\]

Thus by minimizing
\[
\min_{x, \lambda, \mu} \bar{L}_e(x, \lambda, z, \mu) = f(x) + \lambda' h(x) + \frac{1}{2} c |h(x)|^2 + \sum_{j=1}^{r} \left\{ \mu_j g_j^*(x, \mu_j, c) + \frac{1}{2} c |g_j^*(x, \mu_j, c)|^2 \right\}
\]
\[\text{(3.44)}\]
with respect to \((x, \lambda, \mu)\), we can minimize \(\tilde{L}_c (x, \lambda, z, \mu)\). The Lagrangian parameter is updated again by

\[
\lambda_i^{(k+1)} = \lambda_i^{(k)} + c_k h_i \left( x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)} \right)
\]

(3.45)

\[
\mu_j^{(k+1)} = \mu_j^{(k)} + c_k \max \left\{ g_j (x), - \left( \mu_j^{(k)} / c \right) \right\}
\]

(3.46)

Once \((x, \lambda, \mu)^{(k)}\) is found, where \(i = 1, \ldots, m\) and \(j = 1, \ldots, r\), the new penalty parameter \(c_{k+1}\) is chosen to be \(c_{k+1} \geq c_k\).

### 3.5 Penalty Method

Some authors [29] use projected gradient methods and Newton methods to solve constraint optimization problems, while Newton methods have a fast convergence rates, they suffer from instability in most cases when the initial starting point is not close enough to the optimization point. The penalty method is the method used in the friction procedure of this project and could be used in solving the Signorini problem too. It is stable and converges, but it also suffers from ill conditions when the penalty parameter \(c\) is a very large number. The details about the ill condition are described in [59]. From the inequality constraint

\[
\sum_{i=1}^{2} n_i^i \alpha_c v_i^\alpha - g_n \alpha_c \geq 0
\]

(3.47)

where \(\alpha_c\) is the node number of contact nodes, a function is introduced

\[
g_{\alpha_c} (v) = \max \left\{ 0, g_{\alpha_c} (v) \right\} = \max \left\{ 0, \sum_{i=1}^{2} n_i^i \alpha_c v_i^\alpha + g_n \alpha_c \right\}
\]

(3.48)

Using (3.48) the penalty function is defined as

\[
P (v) = c_p \sum_{\alpha_1} \left( v_{\alpha_1}^1 - U_{\alpha_1}^1 \right)^2 + c_p \sum_{\alpha_2} \left( v_{\alpha_2}^2 - U_{\alpha_2}^2 \right)^2 + c_p \sum_{\alpha_c} \left( g_{\alpha_c}^\alpha (v) \right)^2
\]

(3.49)

where \(c_p\) is the large penalty parameter. The unconstrained problem to be solved as optimization of the original function \(F (v)\) plus the \(P (v)\),

\[
\min L_{c_p} (v) = \min [F (v) + P (v)]
\]

(3.50)

with a gradient

\[
\nabla_v L_{c_p} (v) = \nabla_v F (v) + \nabla_v P (v)
\]

(3.51)
This is the exact penalty method without using a multiplier. Numerical experiments suggest that the ill condition can be avoided without giving up rapid convergence. This is done by choosing the initial penalty parameter $c_k$ to be a relatively small number. Then, after a local minimizer is found, choose $c_{k+1} = 10c_k$ and repeat the same procedure by choosing the local minimizer of the last step as a start point. Eventually we can compute $c_k$ as large as $10^{10}$. However the relatively secure method is the multiplier method of section 3.3.

3.6 Relaxation Coefficients

A criterion such as

$$
\varepsilon^{(n)} = \frac{||v^{(n+1)} - v^{(n)}||}{||v^{(n)}||}
$$

(3.52)

is computed in the iteration of frictional contact of two elastic solids to decide whether or not to terminate the iteration process. Numerical experiments tell us that $\varepsilon^{(n)}$ does not always go to zero directly, sometime it wanders back and forth with $\varepsilon^{(n+1)} > \varepsilon^{(n)}$, but the general tendency is to approach zero. If a relaxation coefficient $0 < \omega < 1$ is introduced to relax the marching of displacement of solid II

$$
\tilde{v}^{(n+1)} = \tilde{v}^{(n)} + \omega \left( \tilde{v}^{(n+1)} - \tilde{v}^{(n)} \right)
$$

(3.53)

and then proceed with the two body iterations, $\varepsilon^{(n)}$ can be reduced towards zero without destabilizing the procedure. When $\omega = 1$, there is no difference from before, while $\omega$ is a real number much smaller than 1.00 the convergence is stable, but $\varepsilon^{(n)}$ drops very slowly. Numerical experiments suggest that the best choice of $\omega$ depends on the specific problem.

3.7 Flowchart

The flowchart of the algorithm for computing the effect of frictional contact between two elastic solid bodies is given in Figure (3.1).
Figure 3.1. The flowchart of our algorithm for modelling frictional contact of two elastic solid bodies. The input parameters are Young's modulus, the Poisson ratio, load force, boundary condition and the files containing the nodes and elements. The applied load is increased at the Load Increase box. The Signorini box refers to the procedure to solve the Signorini problem for solid I ($\Omega_1$), while the Friction box refers to the procedure to solve the friction problem with a prescribed normal traction. The procedure to evaluate the normal contact stress of the contact surface is implemented at box Stress, while Solid II relates to the procedure to solve the solid II ($\Omega_2$) problem. The final output is in terms of displacements and stresses. In the flowchart we use question marks to indicate conditional switches. They ask the following: (a) Has the solid I problem converged? (b) Has the normal stress on the contact surface converged? (c) Has the solid II problem converged? (d) Does the applied load force equal the entered load force?
Chapter 4

Examples of Numerical Experiment

4.1 Introduction

The formulations of the last two chapters have been implemented using the C++ language. This program uses the mesh data created from commercial software package ANSYS. That package generates the mesh for a given geometry and provides the information about the list of nodes, elements, boundary and displacement constraints for exporting. Each node is numbered in sequence by a natural number followed by its $x$, $y$, $z$ coordinates. We choose six node complex isoparametric elements with the six nodes labeled as $(i, j, k, l, m, n)$. Each element is also numbered using natural numbers, followed by the node numbers of its $(i, j, k, l, m, n)$ constituents. The list of nodes, elements, boundary nodes and displacement constraints are the inputs of our modeling program and also are exported by the ANSYS software. Luck McAven [64] provides a new method to generate an efficient mesh, which could also be used with this computational scheme naturally. Since the shape functions of the complex isoparametric elements are 2nd order polynomials, see section (2.1), the stress is taken to vary linearly with this kind of element. If there is some drastic change within the region of one element, the program just shows a straight line representing linear change. Refinement of the mesh is necessary to show subtle changes within in the elements, but that increases memory and computational demands to finish the job in a limited time. The numerical experimental results using currently available resources are presented in this chapter. Widely applications and more accurate results can be obtained with better computing resources, but the current results are enough to show some features
and draw some conclusions.

Investigations in this chapter compare the results of the smooth contact problem with the results of Hertz theory based calculations in terms of deformed surface shape, contact semi-width, contact pressure distribution, and the stress distribution underneath the contact surface. Those are the factors which modern engineers deem most significant. It is shown that numerical results of the smooth contact problem are very close to Hertz theory.

The numerical data for the frictional contact problem display different features from Hertz theory predictions. Tangential traction is more effective than what Hertz theory predicts, the normal traction is distributed closer to the edge with a peak higher than normal Hertz pressure in frictional contact between elastic solids. The modeling results also behave differently with different geometries, dimensions, friction coefficient and load incremental steps. The contact stress and transverse displacement are found to be sensitive to the load force distribution.

4.2 Modeling Frictionless Contact between an Elastic Solid and a Rigid Surface

The first example: A two dimensional plane strain problem can closely approximate the problem with cylindrical symmetry. This example deals with the ideal problem of an elastic cylinder in contact with a rigid cylindrical smooth surface with their axes parallel to each other. Since the computational domain is symmetrical, we only use a quarter of the domain in order to concentrate computation resources on the contact region. Nodes are highly concentrated on the contact region, so the element should be refined enough to show the trend of the change in contact stress. The loading pressure is evenly distributed on the top of elastic solid, see Figure (4.1).

The results of this modeling are very close to the Hertz theoretical results. It is necessary to mention that the Hertz results are calculated using a reduced radius \( R \)

\[
\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} \quad (4.1)
\]
Figure 4.1. The first example: An elastic half cylinder in contact with a rigid surface. Only a quarter of the cross section of the cylinder is used in the computation. A symmetry boundary condition is applied on symmetry axis.
and reduced modulus of elasticity $E$ by

$$\frac{1}{E} = \frac{1 - \sigma_1^2}{E_1} + \frac{1 - \sigma_2^2}{E_2}$$  \hspace{1cm} (4.2)$$

We choose the applied total loading to be $1.72 \times 10^6 N/m$, $E_1 = 2.0 \times 10^{11} Pa$, and the Psoisson ratio $v = 0.3$. Since the second solid body is rigid, $E_2$ is infinite. The numerical model gives an almost identical distribution of pressure with the Hertz theory prediction, see Figures (4.2) and (4.3). The maximum normal pressure is the same as the Hertz result of $6.0082 \times 10^6 N/m^2$ and the contact semi--width is very close to the Hertz result of $1.82249 \times 10^{-3} m$. It is easy to see that the approach is just the $y$ component displacement of the node on the top surface of the elastic body, which is $1.74697 \times 10^{-5} m$, but the approach diverges by Hertz theory. The approach differs for different particle in the solid, the closer a particle is to the contact region, the smaller is its approach. Hertz theory computes the approach of a particle situated at infinity, which is not a realistic situation for a solid of finite dimension.

The only difference between the Hertz theory results and the results of this numerical modeling can be found at the edge of the contact region. One reason for this is due to the fact that the 6 node complex elements limit the approximation of stress to vary linearly within an element. This error makes material appear harder than it is. Another reason is on the side of Hertz theory, which uses the semi--infinite media without considering the
Figure 4.3. Comparison of the stress under the surface between the Hertz theory and our numerical calculation. Curve 1: Hertz Theory principle compressive stress, curve 2: numerical principle compressive stress, curve 3: numerical principle shear stress, curve 4: Hertz theory principle shear stress. The stress is relative to the maximum stress predicted by Hertz theory.

graphy of the contacting solid body and the pressure distribution. However at least numerical results suggest properties closer to what Hertz theory predicts, so that the modeling is on the right track. In addition, numerical modeling can provide approach and transverse displacement of surface particles on the contact region. This will be shown in the section 4.3. We need to bear in mind that the assumptions of rigid surface and frictionless contact are always idealistic. A more general situation by considering frictional contact and an elastic surface for the second solid body is to be investigated in section 4.2.

The second example is illustrated in Figure (4.4). The geometry of this example differs from the first example, but the loading condition is similar. Figures (4.5) and (4.6) show the results of comparison between the first example and the second example in terms of the normal pressure and the stress underneath the contact surface. Both examples have the same radius of surface curvature for the elastic solid and the rigid surface except that the elastic solid in the second example has a smaller size, it is made of part of domain of the elastic solid in the first example. It is shown that their difference in the surface
contact pressure and contact surface shape are so small that it is reasonable to treat them as being identical. Only transverse displacement and stress underneath the contact surface are different, see Figures (4.5) and (4.6).

So far we can see the influence of the geometry and the size of solid bodies on the contact pressure. The influence on surface shape and stress distribution within the solid is almost negligible for frictionless contact between an elastic solid body and a rigid solid body. The obvious difference is in the $z$ component of particle displacement on the contact surface $u_z$. This is easy to understand from intuition. The second example is a thinner block, which is weaker in terms of resistance to deformation than the first example. But the difference in the slippage does not affect the stress distribution significantly however, because of the smooth contact surface. This is why Hertz and previous workers ignored details regarding the distribution of loading and specific geometry on the rest part of the

Figure 4.4. The second example. An elastic solid in contact with a rigid surface.
Figure 4.5. Comparison of the stress underneath the Contact surface between the first and second examples. curve 1: numerical principal compressive stress of example 2; curve 2: numerical principal compressive stress of example 1; curve 3: numerical principal shear stress of example 2; curve 4: numerical principal shear stress of example 1. The stress is relative to the maximum stress predicted by Hertz theory.

Figure 4.6. Comparison of the displacement $v_x$ between the first and the second examples. The $x$ and $v_x$ axes are in terms of Hertz contact half-width.
solid bodies. However this cannot be ignored for frictional contact. Once friction interferes with the deformation, the stress is sensitive to the relative transverse displacement. We will discuss this in section 4.3.

Contact stress dependence on the distribution of the loading force can be investigated by changing the loading force distribution of the first example, while maintaining the same total force. It is seen in Figure (4.7) that the difference is so small that it is acceptable to neglect the detail of loading distribution in frictionless contact between an elastic solid body and a rigid surface. The first example is applied with different kind of Loading conditions. For example loading condition I has a force density $8.6 \times 10^8 N/m^2$ pointing toward the negative direction of $y$ axis, and it is evenly distributed on the top of elastic solid surface between $x = -0.001m$ and $x = 0.001m$. Loading condition II has a force density $4.3 \times 10^8 N/m^2$ pointing toward the negative direction of $y$ axis and it is evenly distributed on the top of elastic solid surface between $x = -0.002m$ and $x = 0.002m$. For convenience we denote the loading distribution I as $8.6 \times 10^8 N/m^2 \times [-0.001 - 0.001]m$ and the loading condition II as $4.3 \times 10^8 N/m^2 \times [-0.002 - 0.002]m$ in all captions and the rest of this chapter.

It is shown in Figure 4.8 that $u_x$ of the surface particles does strongly depend on the loading distribution. The force of loading condition I is more concentrated on the middle point than that of the loading condition II, thus the transverse shift $u_x$ of the surface particles in loading condition I is smaller than that in the loading condition II. This does not influence the stress distribution for frictionless contact much, but it is found by numerical experiments discussed in the next section that this difference in transverse shift has a significant impact on the stress distribution when friction is taken into consideration.

### 4.3 Frictional Contact between two Elastic Solids

Whether or not elasticity and friction have an important impact on the contact surface is the focus of our study of this section. We illustrate the size and geometry of our third example in Figure (4.9).

The meshed elements of the third example are shown in Figure (4.10). The contacting elastic body in this example has the same radius as that of the second example, however
Figure 4.7. Comparison of contact pressure between loading condition I (Series1) and loading condition II (Series2). The x axis is in terms of Hertz contact half-width. The contact pressure is relative to the maximum pressure predicted by the Hertz theory.

Figure 4.8. Comparison v_x in contact surface between loading condition I and loading condition II. The x and v_x axes are in terms of Hertz contact half-width.
the surface is frictional and the second solid body is elastic. Numerical results based on
the algorithm in section 2.3.3 show different features from what Hertz theory predicts.

The normal pressure and tangential traction distribution of frictional contact between
two elastic solid bodies are shown in Figure (4.11). The final load distribution is \(4.3 \times 10^8 N m^{-2} \times [-0.002 - 0.002] m\).

When the loading force density is increased from \(0 N m^{-2}\) to the final force density
\(4.3 \times 10^8 N m^{-2}\) in 4 equal steps with each step increasing the force density \(1.075 \times 10^8 N m^{-2}\),
the final contact pressure shows peaks of normal pressure within the contact region. These
peaks are about the same height as the maximum Hertz pressure. When the force density
is increased from \(0 N m^{-2}\) to the final force density \(4.3 \times 10^8 N m^{-2}\) in 10 equal step with
each step adding the increment force density \(0.43 \times 10^8 N m^{-2}\), there are small peaks of
normal pressure around the edge of the contact region. For both kinds of loading increase,
one obvious peak is in the middle of the contact region with a height 30% higher than the
maximum Hertz pressure. The contact width is almost the same as what Hertz theory
predicts, but the semi-width of the 10 step loading case is a little wider than that of

75
**Figure 4.10.** The mesh of the third example with a frictional coefficient $\mu = 0.35$. The principal radius of curvature of the contact surface of elastic solid I is 0.005m. Solid II is elastic. The principal radius of curvature of the contact surface of elastic solid II is 0.01m.

The 4 step loading. How high these peaks can reach may be further investigated using a more refined mesh, but this kind of mesh condition can indicate that the final stress distribution is dependent on the step numbers of loading and loading increment of each step. It is indicated by Rabinovich [62] that the pressure around the edge is a little higher than that in the middle. We can draw the conclusion that the impact of frictional and elastic effects is definitely apparent. The stress under the contact surface also deviates from Hertz theory due to friction and elastic effects see Figure (4.12).

Our fourth example examines how stress depends on the size, geometry of the contact bodies and loading distribution for elastic frictional contact. The example is illustrated in Figure (4.13).

The contact surface curvature of the fourth example is the same as that of the third example, but the geometries of the remaining part of these two solid bodies are different. Both elastic bodies are a quarter of the cylinder. It is seen in Figure (4.14) that the geometry and size does affect the stress distribution. The stress in the region close to the edge rises above the Hertz theory results too, but there is no peak in the middle, and the surface approach is even more than that of the last example, see Figure (4.15). This is
Figure 4.11. Comparison of the surface traction of frictional contact between different step numbers of loading and loading increment of each step. curve 1: normal pressure after a 4 step load increase; curve 2: normal pressure after a 10 step load increase; curve 3: tangential traction after a 4 step load increase; curve 4: tangential traction after a 10 step load increase. The x axis is in terms of Hertz contact half-widths. The traction is relative to the maximum pressure predicted by the Hertz theory.
Figure 4.12. The final stress underneath the contact surface for the third example with a 4 step loading increase from zero to final loading condition as \(4.3 \times 10^4 N/m^2 \times [-0.002 - 0.002]m\). curve 1: numerical principal compressive stress; curve 2: Hertz theory principal compressive stress; curve 3: Hertz theory principal shear stress; curve 4: numerical principal shear stress. The stress is relative to the maximum pressure predicted by the Hertz theory.

because body I (\(\Omega_1\)) and body II (\(\Omega_2\)) in the fourth example have a larger size than those in the third example, so that both surfaces will collapse more, leading to the two contact solids approaching each other more. This is the information which cannot be provided by Hertz theory. Another phenomenon we have observed through this example is that when the size of both contact bodies increases, that means it is closer to the idealistic half space condition, which is an assumption of Hertz theory, the numerical results of the stress under the contact surface will be closer to what Hertz theory predicts. See Figure (4.16). Thus this program can repeat Hertz theory results to some extent, as the geometry conditions approach the ideal conditions assumed by Hertz theory. We noticed that the Hertz maximum pressures of the third example and the fourth example are \(9.1 \times 10^9 N/m^2\), which is 4.55% of the modulus of elasticity \(E\). The strain limit of elasticity for most metal materials is usually 0.2%. Thus these examples are only used for the purpose of illustration. The examples used by Kikuchi [26] and Refaat [29] are much more over the elasticity limit. Another observation is that the program for the frictional
Figure 4.13. The fourth example: two elastic body contact without friction.
Figure 4.14. Contact stress of the fourth example with load distribution $4.3 \times 10^8 N m^{-2} \times [-0.002 - 0.002] m$. The $x$ axis is in terms of Hertz contact half-width. The contact pressure is relative to the maximum pressure predicted by the Hertz theory.

Problems converge stably if the loading is low enough to make the maximum strain less than 1%, and it sometimes does not converge properly for cases when the maximum strain is more than 1% caused by the higher loading. This phenomena becomes more obvious as the number of loading steps increases to more than 20. The computation time increases as well.

4.4 Slippage and Frictional Contact

It has been shown that the transverse displacement of a surface particle is very sensitive to the loading distribution, and also to the geometry and size of the solid body. In this section we make a comparison of the transverse displacement $u_z$. Figure (4.17) shows a comparison between smooth contact and frictional contact under a loading distribution $4.3 \times 10^8 N m^{-2} \times [-0.002 - 0.002] m$. Figure (4.18) shows a comparison of $u_z$ between smooth contact and frictional contact under a loading distribution $8.6 \times 10^8 N m^{-2} \times [-0.001 - 0.001] m$. Figure (4.19) shows a comparison of $u_z$ between two loading distributions in the case of frictional contact.

These comparisons tell us that the $u_z$ displacement, which is closely related to the
Figure 4.15. Comparison of the surface shapes of the third and fourth examples. The $x$ and $y$ axes are in terms of Hertz contact half-width.
Figure 4.16. Comparison of stress underneath the surface between Hertz theory results and our numerical results for the fourth example. Curve 1: numerical principal compressive stress; Curve 2: Hertz theory principal compressive stress; Curve 3: numerical principal shear stress; Curve 4: Hertz theory principal shear stress. The stress is relative to the maximum pressure predicted by the Hertz theory.

slippage, is sensitive to the loading condition. The friction between the surfaces prevents the surface particles on the two surfaces from moving transversely relative to each other. The surface particles have a tendency to move to the center of the contact region. The more scattered the loading force, the stronger the tendency to move toward the center, which will affect the friction on the surface. Since the transverse displacements of the surface particles are not evenly distributed, as shown in the Figure (4.17) (4.18) (4.19), so the stress accumulation is also uneven. Once the tangential stress is over the its limitation, which is proportional to the normal traction, the relative slip between two surface occurs. After slipping the stress is released. What we computed above is just a final configuration for the quasi-static situation. Stress is under a process of repeated accumulation and slip release during the time of loading increase. In a dynamic situation this accumulation will proceed before an equilibrium state is reached. It is even possible for the stress accumulation and slip release to happen simultaneously making the friction problem very complicated. Results of the stress distribution from this modeling cannot represent what happens during the dynamic process.
Figure 4.17. Comparison of the $v_x$ between frictionless and frictional surface under the loading distribution $4.3 \times 10^8 N m^{-2} \times [-0.002 - 0.002] m$. Curve 1: target surface is frictional and elastic; Curve 2: target surface is frictionless and rigid. The $x$ and $v_x$ axes are in terms of Hertz contact half-width.
Figure 4.18. Comparison of $v_x$ between frictionless and frictional surface under the loading distribution $8.6 \times 10^8 \text{Nm}^{-2} \times [-0.001 - 0.001]\text{m}$. Curve 1: target surface is frictional and elastic; Curve 2: target surface is frictionless and rigid. The $x$ and $v_x$ axes are in terms of the Hertz contact half-width.

Figure 4.19. Comparison of the $v_x$ between different loadings for the same frictional surface. The $x$ and $v_x$ axes are in terms of Hertz contact half-width.
4.5 An Experimental Example

The rigid punch program is tested in this section using the experiment conducted by Brémond and Durelli [61] as an example for comparison. Kikuchi [26] chose this example to test his model. The disagreement between his computational results and experimental reports made him change the Young modulus of his modeling to a different value from the experiment reports in order to generate matching results. It is shown here that a quite close agreement with the experimental results, using the same modulus of elasticity and Poisson ratio as in the experiment, can be achieved through our modeling. Figure (4.20) is the schematic diagram of the experimental setup and parameters.

The contact solid bodies I and II are cylindrical and rectangular shape respectively. Solid body I is made of plexiglass, which was regarded as a rigid pouch by Brémond and Durelli [61]. Actually most plexiglass has a tensile strength around 103MPa, which is
much softer than the normal metal. Brémond and Durelli did not mention the exact type of plexiglass used in their experiment. Solid body II is made of polyurethane rubber, which is transparent, has a modulus of elasticity $E = 0.3daNmm^{-2}$ and a Poisson ratio $\nu = 0.48$. Brémond and Durelli measured the deformation of the surface of body II, the $u_x, u_y$ curve and $\xi_{xy}$ component of strain on the surface through the measurement of the deformations of the printed grids. They computed the stress, which should be Piola-Kirchoff stress according to their definition. Since the experiment is conducted with a quite large deformation, the difference between Cauchy stress and Piola-Kirchoff stress becomes obvious in this example. It is shown that our modeling results from the frictionless rigid pouch program are quite close to the experiment, with an indentation approach of $3.2 \times 10^{-3}m$, where the experimental result is a little above $3.0 \times 10^{-3}m$. The deformation of elements is shown in Figure (4.21) for terms $u_x$ and $u_y$, see Figure (4.22). The stress distribution on the contact surface is shown in Figure (4.23).

It is necessary to mention that this experiment is actually a large deformation example more like a plane stress case as Kikuchi mentioned, so the difference between Cauchy and Piola stress becomes obvious. See Figure (4.24).

It is still not possible to reliably measure the stress of the contact region experimentally till now, especially for small deformations of metallic material. I have not yet found any experimental reports which have been conducted with metal material. Most papers only deal with numerical modelling using an ideal normalized modulus of elasticity. The applied pressure is large enough to overcome the elasticity limit. The results of our approach imply that the contact stress distribution in realistic circumstances is sensitive to many factors, including size, geometry, loading distribution and loading history. The relative slippage of the interaction surface tends to accumulate the stress in a particular concentrated place, so that the stress at that particular point of the contact surface is much higher than what Hertz theory predicts. This may be the reason for surface fatigue of the parts, however since the friction model we used here is a Coulomb model, we can not prove that there is a unique solution to the variational inequality. Thus the numerical phenomenon of high stress at some point, as the example in this chapter gives is very possibly one solution of frictional elastic contact. Reality is so complicated that there are other elements such as heat, multiple contacts, and dynamical and three-dimensional effects being ignored.
Figure 4.21. The deformation of the finite elements grid. The solid lines represent the mesh after deformation, while the dashed lines represent the mesh before deformation.
Figure 4.22. Displacement of the contact surface.
Figure 4.23. Numerical result of contact stress. The x axis is in terms of Hertz contact half-width. The contact stress is relative to the maximum pressure predicted by the Hertz theory.
Figure 4.24: Stress under the surface. The stress is relative to the maximum pressure predicted by the Hertz theory.
far we can push our model is largely dependent on the computational resources. The real world may be much more subtle than what Hertz theory describes.

Excessive computational time is involved in the above numerical computation, and time increases with total nodes of the domain (see Table 1). The results in the table are for numerical experiments run on an Intel Pentium 500MHz machine with Windows 98 operation system. Much faster run times could be achieved in a more powerful system.

<table>
<thead>
<tr>
<th>Number of nodes per body</th>
<th>Frictionless problem</th>
<th>Frictional problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>300</td>
<td>30 minutes</td>
<td>3 hours</td>
</tr>
<tr>
<td>500</td>
<td>45 minutes</td>
<td>10 hours</td>
</tr>
<tr>
<td>600</td>
<td>2 hours</td>
<td>15 hours</td>
</tr>
<tr>
<td>800</td>
<td>4 hours</td>
<td>30 hours</td>
</tr>
</tbody>
</table>

Table 4.1. The estimated computation time for different number of nodes per body.
Chapter 5

Conclusions and Proposals for Future Investigations

5.1 Summary of New Methods in this Modeling

We have addressed the problems of previous researchers [26, 29] in the former chapters, and provided our methods and schemes for making improvements. The implementation of these methods leads to some new physical conclusions. We are going to summarize these methods in this section and present our conclusions in the next sections.

1) For the Signorini problem, which is a problem of contact between an elastic solid and a rigid surface, Kikuchi [26] solved the variational inequality by the Riesz projection methods, and Refaat [29] solved it by a projected-gradients method. We use multiplier methods with penalty functions plus optimization methods by the Conjugated Gradients Method. Our numerical experiment shows that this method is clear and easy to implement.

2) For the rigid pouch problem, we introduce an indentation displacement with a moving rigid surface and a modified contact boundary condition, eventually solving the problem by the optimization of total potential energy with inequality constraints. Our result is closer to the experiments than what is reported in [26]

3) For the frictional contact problem, we again use exact penalty methods plus optimization by the Conjugated Gradients Methods, while Refaat [29] used Newton–
Raphson Methods.

4) For all the problems, a spline interpolation method is used to approximate the contact surface by a continuous polynomial function. Kikuchi and Refaat [26, 29] modelled the contact surface by straight-line joining the neighboring contacts nodes, which is not smooth if the node is not dense enough at the boundary.

5) We repeatedly use a Conjugated Gradient subroutine.

6) The transformation between Piola–Kirchoff stress and Cauchy stress is implemented in our model. This was ignored by previous researchers. We have found, see chapter 4, that the difference between them is apparent, especially for large strain deformation.

7) For the elastic frictional contact problem, we introduce a new approach by coupling the two elastic solid solution iteratively. Kikuchi and Refaat [26, 29] ignored the relative tangential shift of the two contact surfaces, and assumed that each particle on one surface has a constant mating relationship with certain particles on the other surface. Our model solves these two solids alternatively and shows that the relative tangential shifts change the contacting relation between surface particles.

8) Kikuchi and Refaat [26, 29]'s model ignore the effects of geometry, size and the loading distribution. Our model is sensitive to the differences.

9) The first example shows that our modeling can closely reproduce the Hertz theory result at conditions which is close to the ideal conditions assumed by Hertz Theory.

5.2 Physical Conclusions

Several new conclusions can be drawn by the implementation of the above new methods and schemes of our modeling. These conclusions could lead to a reconsideration of the impact of frictional surfaces for elastic solids.

The first conclusion is that our numerical modeling can reproduce Hertz's results in an ideal situation, which is frictionless contact between an elastic solid and a rigid surface. The loading distribution, total geometry and size of the elastic solid have little influence on
the contact traction and stress. The effects of those factors are assumed to be negligible in the Hertz theory. Previous numerical research implies that the dependence of the contact traction on the loading distribution, total size and geometry are small enough to neglect. That is why references [26, 29] modeled the loading only by a total force vector acting on a point on the surface opposite to the contact surface without further investigation. The authors then carried this assumption over to the frictional contact case, treating it the same way as the frictionless contact case.

The second conclusion is that the stress and surface traction of frictional surface contact between two cylindrical elastic solid bodies deviate from the predictions of Hertz theory. The nature of the deviation depends on the specific condition of the example, which can be characterized by the loading distribution, the step numbers of loading, loading increment of each step, total geometry and the size of the solids. Such kinds of dependence were ignored before. Our numerical examples demonstrate that the pressure in the middle of the contact region will be raised by about 30% for the frictional contact above the Hertz maximum pressure. Also some peaks and valleys of normal and tangential traction take place at positions between the center and the edge. The unevenness of surface traction can significantly affect the initialization of cracking. This phenomenon happens for our example with a small size solid and a bigger loading increase step. Our results tend to be more even and closer to the Hertz results with a bigger size solid and smaller loading increase step. This shows that Hertz results will be approached asymptotically as the above parameter approaches the Hertz ideal situation. That means the principal radius of the surface curvature and the dimension of the solid are much larger than the dimension of the contact area. Hertz theory has been used for a century by engineers, and it can approximate reality well in many cases. Our current work complements and refines Hertz theory.

The third conclusion is that our modeling results show that materials are under harsher conditions than assumed by the Hertz theory. The convergence indicates the robustness of the program. Our modeling is basically not in contradiction with the previous theories and model, but it is a further refinement of them. Distinguishing between different loading and geometric conditions is necessary because the exact nature of the loading and geometry will vary greatly between different real examples. We still lack a detailed knowledge about
the loading distribution. Since all loading is transferred through contact, which is related to a free boundary problem, we can not completely know the distribution of loading force before we solve the problem. The assumed even distribution used in our example is a simple case to illustrate the problem. Our different results on various even distributions are enough to show that the influence of the loading distribution is effective.

Our friction model is based on the Coulomb friction law. Whether this law is reliable is still unclear, but at least our modeling results tell us that friction between the surfaces and the elasticity of solid II together play an important role in the distribution of stress. Our model shows that the relative tangential displacement of the two surfaces is dependent on the loading and the geometric configuration. Since the friction is caused by relative tangential displacement, the surface particles on these two surfaces change their contact relation once the slippage occurs. Stress release occurs at the same time as slippage. Stress accumulation and release happen at each loading step. The final stress remnants become uneven across the contact region. It is possible for stress at some particular place to be much higher than the smooth Hertz distribution at final equilibrium state.

The last conclusion is that the difference between Cauchy stress and Piola-Kirchoff stress is well addressed in our model. We find that the difference between these two stresses is small for small deformation. But it grows as the deformation grows. This has also been ignored by most previous studies. The theoretical work of Duvaut and Lions [22] used Cauchy stress concepts only, many examples of finite element methods of Kikuchi [26] and Refaat [29] use reference coordinates, i.e. Piola-Kirchoff stress. The loading in their examples however is high enough to exceed the elastic limit for the purpose of illustration. This most likely is the other reason for the difference between our results and theirs.

5.3 Proposal for Future Investigation

A large number of questions are still open. Is the peak of pressure at the edge real as suggested by Figures (4.11) and (4.14)? How precisely does the pressure behave in the three-dimensional case? The nonclassical and nonlinear frictional model still needs further testing by numerical computation, experimental work and engineering practice. Friction is
by nature a dynamical interaction because it is dependent upon the loading history. How close and reasonable this quasi-static approximation can be is still unclear. Further work requires more powerful computing resources. Most work in the literature is theoretical and numerical. There is little experimental work on the plane strain case, which matches our two dimensional modeling. Moreover, the real surface, which we regard as smoothly defined by a curve generated from spline interpolation of the surface node points, always has asperities on the microscopic scale. Little is known about friction at the microscopic scale [46, 47, 48, 51], so the pressure we introduced here is just the average of the pressure of the microscopic scale. The real peak at the microscopic scale may be much higher than this within the asperities. It would be interesting to test this principle for rough surfaces, but so far it seems to be a big problem from the point of view of computational resources.

Another problem is that since the real mechanical components are all in contact with two or more other components, and each contact region is under an unknown condition. Whether or not the pressure and friction on all these contact surfaces couple with each other is unknown. If so, how much they influence each other is certainly an open question. Since our model does not consider the influence of more than two contact regions, we cannot address those issues right now. Whether a solution exists or is unique is still unknown. All these problems demand more powerful computation resources if they are to be addressed numerically. Distributed computing technology, which is a new computing technique making use of available computing resource through the network, may offer a solution in the future [64].
APPENDIX

Gauss Quadrature

In order to evaluate the integration, Gauss quadrature is used

\[ \int_0^1 \int_0^{1-t} y(\xi, \eta) \, d\eta d\xi = \sum_{i=1}^{r} w_i y(\xi_i, \eta_i) \]  

(1)

For triangular elements, 12 points are chosen to compute the integration, and the table A(1) below lists the coordinates and weights of these points. These point are distributed as in Figure A (1).

<table>
<thead>
<tr>
<th>point n</th>
<th>( L_1 )</th>
<th>( L_2 )</th>
<th>weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>a</td>
<td>a</td>
<td>0.025422453185103</td>
</tr>
<tr>
<td>2</td>
<td>1 - 2a</td>
<td>a</td>
<td>0.025422453185103</td>
</tr>
<tr>
<td>3</td>
<td>a</td>
<td>1 - 2a</td>
<td>0.025422453185103</td>
</tr>
<tr>
<td>4</td>
<td>b</td>
<td>b</td>
<td>0.058393137863189</td>
</tr>
<tr>
<td>5</td>
<td>1 - 2b</td>
<td>b</td>
<td>0.058393137863189</td>
</tr>
<tr>
<td>6</td>
<td>b</td>
<td>1 - 2b</td>
<td>0.058393137863189</td>
</tr>
<tr>
<td>7</td>
<td>c</td>
<td>d</td>
<td>0.041425537809187</td>
</tr>
<tr>
<td>8</td>
<td>d</td>
<td>c</td>
<td>0.041425537809187</td>
</tr>
<tr>
<td>9</td>
<td>1 - (c + d)</td>
<td>d</td>
<td>0.041425537809187</td>
</tr>
<tr>
<td>10</td>
<td>1 - (c + d)</td>
<td>d</td>
<td>0.041425537809187</td>
</tr>
<tr>
<td>11</td>
<td>c</td>
<td>1 - (c + d)</td>
<td>0.041425537809187</td>
</tr>
<tr>
<td>12</td>
<td>d</td>
<td>1 - (c + d)</td>
<td>0.041425537809187</td>
</tr>
</tbody>
</table>

\( a = 0.063089014491502 \) \( b = 0.249286745170910 \)
\( c = 0.310352451033785 \) \( d = 0.053145049844816 \)

Table 1. The Gauss quadrature for triangular element.
Figure 1. Points used in the Gauss quadrature for triangular elements.
Bibliography


[4] E. McEwen 'Stress in elastic cylinders in contact along a generatrix' Philosophical Magazine 40 (1949) 454--.


[40] J. Boussinesq Compt. Rend 114 (1892) 1510–.


Vita Auctoris

NAME:               Yuan Yao

PLACE OF BIRTH:    Hunan, P. R. China.

YEAR OF BIRTH:     1971
EDUCATION:         1993 B.Sc. Beijing Normal University, China.