1998

# Finite element modeling of interface crack contact problems in semiconductor packages

Yavuz Bertan Bayram
*Lehigh University*

Bayram, Yavuz Bertan

# Finite Element Modeling of Interface Crack Contact Problems in Semiconductor Packages

May 13, 1998

# FINITE ELEMENT MODELING OF INTERFACE CRACK CONTACT PROBLEMS IN SEMICONDUCTOR PACKAGES

by

Yavuz Bertan Bayram

A Thesis
Presented to the Graduate and Research Committee
of Lehigh University
in Candidacy for the Degree of
Master of Science
in
Mechanical Engineering

**Lehigh University**
**May 1998**

This thesis is accepted in partial fulfillment of the requirements for the degree of Master of Science.

May 5, 1998

(Date)

Prof. Herman F. Nied
(Thesis advisor)

Prof. Charles R. Smith
(Chairman of Department)

# Acknowledgments

I would like to express my deepest gratitude to my advisor Dr. Herman Nied who patiently answered all my questions and showed me the right way for the completion of my thesis. The financial support of Semiconductor Research Corporation is greatly appreciated. I also would like to thank my family and my friends here at Lehigh for their endless support and help.

I would like to dedicate this work to my mother Asuman Bayram to whom no matter how hard I try, I cannot thank enough. Without you I wouldn't be who I am. Hope you finally rest in peace.

# Table of Contents

# List of Tables

# List of Figures

# ABSTRACT

In this study, the interface crack surface contact problems that occur in semiconductor packages have been investigated. The ultimate goal of modeling is to come up with a methodology that will be used for the fail-safe design of semiconductor packages. For this reason it is very important to model the actual problem using the least possible number of assumptions thereby simulating the real situation.

A common problem in modeling interface cracks in semiconductor packages is the interface crack surface contact problem. For certain geometries, material combinations and temperature loading, using conventional finite element methods, the interface crack surfaces pass through each other giving result to drastically wrong stress intensity factors, strain energy release rates and phase angles. For this reason, an efficient contact mechanics algorithm has been implemented into the interface crack problem. General finite element methods for contact mechanics are reviewed and special attention is given to the penalty function method to implement the necessary displacement constraints to prevent possible penetrations. Together with the enriched element formulation for modeling the crack problem, the penalty function method offers an easy to implement, accurate and efficient approach in modeling interface crack surface contacts. The correctness of the algorithm will be proved using some example problems. Numerical examples will be given for edge-cracked and center-cracked flip chip packages using several boundary conditions. The relationship between total strain energy release rates, phase angles and crack lengths will be shown. The results will be compared with the ones in which the contact constraints were not imposed.

# CHAPTER 1

# INTRODUCTION

## 1.1 SEMICONDUCTOR PACKAGES

In today's technological world, semiconductors are almost everywhere; in cars, computers, washing machines, satellites etc. They are controlling our everyday life, making it better. The advances done in this industry over the last decades have been breathtaking. From primitive computers of the 50's that were big enough to fill a room, we have reached to ultra-fast, compact personal computers. From "light-bulb" transistors we came to space-age semiconductors.

There are many different types of semiconductor packages, such as plastic encapsulated package, flip-chip package and so on. In this thesis particular attention will be given to flip-chip packages since they represent today's technological trend in the semiconductor industry. A typical flip-chip package is shown in Figure 1.1. It is mainly composed of 3 parts; the silicon die, the polymeric substrate and the underfill layer bonding the die to the substrate. The underfill layer itself contains solder balls which holds the die and the substrate together. The space between the solder balls are filled with the underfill material.

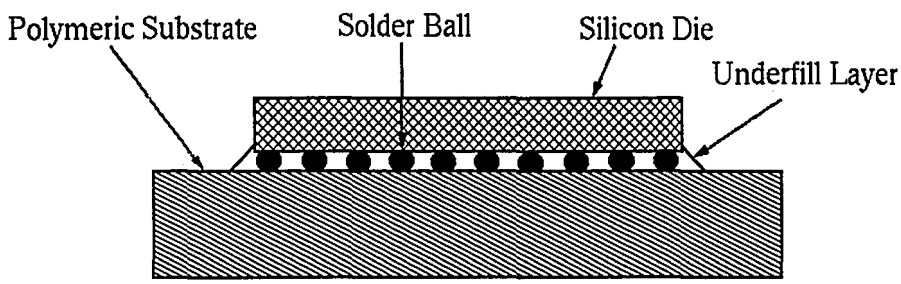


**Figure 1.1:** A typical flip-chip semiconductor package

With the growing need in faster and smaller semiconductors, new design challenges are faced. Since the silicon chip has a denser structure, the package is heated more and has more difficulty in dissipating that heat, therefore, thermal loads on the package increase. With the addition of the mismatch in the coefficients of thermal expansion of the different constituent materials, thermal stresses can cause the failure of the package.

## 1.2 OVERVIEW OF THE PROBLEM

As explained in the previous section, the silicon die and the polymeric substrate are bonded together by the solder balls present in the underfill layer. This is done by a process called *solder reflow process*, in which the temperature in the package can be quite high. Most of the adhesion problems occurs in this stage, the high temperature rise in the package can cause the silicon die to separate from the substrate. The reason for this phenomena is the mismatch in the coefficients of thermal expansions of the constituent materials. This mismatch causes high shear stresses along the interface making crack initiation and the growth of an already existing or a just initiated defect easier. The interfacial integrity of semiconductor packages is a very big reliability issue and a lot of work is being done in this subject to understand and prevent semiconductor failures.

There are two common locations within the package where either a crack can initiate or can be already present. In a fracture mechanics point of view, the corner between the underfill layer and the polymeric substrate, shown in Figure 1.2, has a relatively high strength of singularity, i.e. the stresses are very high at that location. Therefore a crack is quite susceptible to start at that location. The other common location for a crack to be present is the interface between the underfill and the substrate. It is

known that some problems can occur in filling the space between the solder balls with the underfill material. The underfill material is injected from the sides of the package after the solder reflow process. Sometimes a flaw can be left unfilled. Also environmental factors, especially moisture intake, can cause a crack to form at that interface. The substrate is a polymeric material, therefore some moisture will be absorbed from the environment. This moisture will evaporate during the solder reflow process where the temperatures are quite high. This pressurized vapor may initiate cracks along the interface or may fill the already present flaws and apply pressure on the interface which may finally cause the package to break.

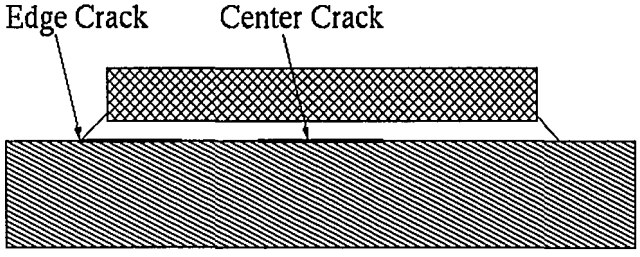


**Figure 1.2:** Common cracks encountered in flip-chip packages

As seen in Figure 1.1, the geometry of a semiconductor package can be quite complex. The problem gets even more difficult with the addition of a bimaterial interface crack, numerous boundary and loading conditions. Obviously, obtaining an analytical solution for the problem is often either inadequate or impossible. Therefore, a numerical analysis tool is necessary. Finite element method, which is by far the most popular numerical method to tackle this kind of problems, is commonly used in the past to model semiconductor packages.

A common problem in modeling interface cracks in a semiconductor package is the crack surface contact problem. For certain loads, boundary conditions, geometry and material property combinations the crack surfaces may become in contact. Conventional methods for modeling the interface crack problems fail to give correct results when contact occurs since the contact constraints are not implemented into the finite element equations to be solved, resulting in a mesh overlap. As an example, an edge-cracked flip-chip package is considered and the model is shown in Figure 1.3. Two different temperature loads have been applied for the same model with the same geometry, boundary conditions and constituent materials.

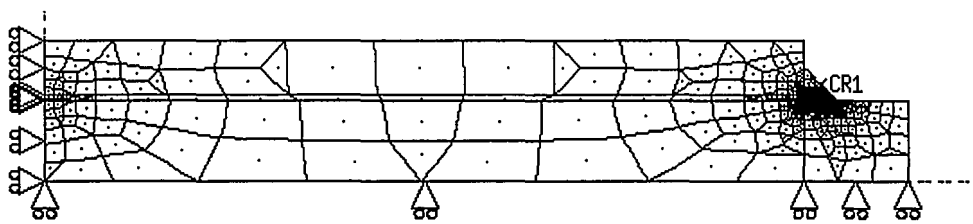


**Figure 1.3:** Edge-cracked flip-chip package model where contact occurs for $\Delta T = -10^{o}C$ but no contact occurs for $\Delta T = 10^{o}C$.

For a temperature loading of $\Delta T = 10^{o}C$ the crack surfaces opens, as shown in Figure 1.4. This kind of linear elastic crack problems have been extensively studied in the past.

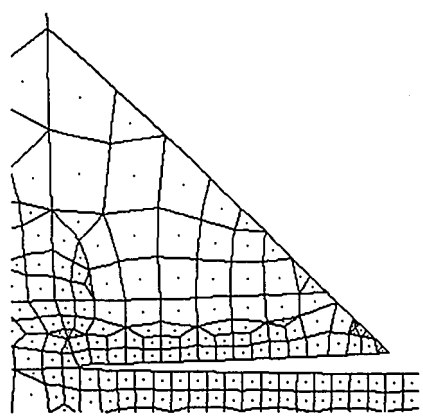

**Figure 1.4:** Close-up view on opened crack surfaces for $\Delta T = 10^{o}C$

When a temperature loading of $\Delta T = -10^{o}C$ is applied, the reverse occurs and the crack surfaces overlap with each other, as shown in Figure 1.5. If the contact constraints are not imposed on the crack surfaces and the mesh is allowed to overlap, the results found will be inevitably wrong.

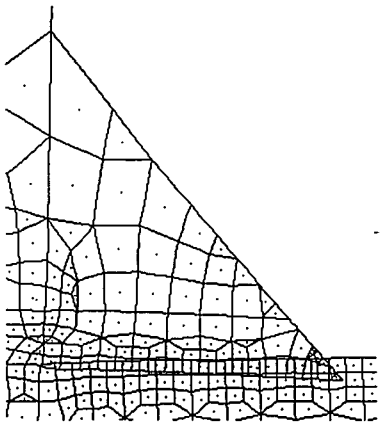

**Figure 1.5:** Close-up view on overlapping crack surfaces for $\Delta T = -10^{o}C$

For this case, the stress intensity factors becomes the negative of the ones obtained with $\Delta T = 10^{o}C$ and the total strain energy release rate will be the same, but these are obviously not correct. Therefore, the necessary contact constraints are needed to be implemented into the interface crack problem to obtain a methodology for fail-safe design.

As it will be shown in the numerical examples chapter, Chapter 3, for certain boundary conditions and geometries, contact may occur for certain crack lengths and may not occur for others. Therefore in modeling one cannot be sure beforehand whether contact will occur or not and action should be taken in order to detect possible contacts and thereby prevent possible penetrations which are physically inadmissible.

## 1.3 PREVIOUS WORK

Interface crack contact problems or interface crack closure problems have been the focus of many research in the past. Recently, the interest in this area increased due to

the desire in understanding the modes of failure in multi material structures. Analytical investigations in this area dates back to the work of Williams [1], who used an asymptotic analysis for traction free crack surfaces to find out that the displacements behaved in an oscillatory manner near the crack tip, causing crack surface penetration. Erdogan [2] found that for a non homogeneous interface with cracks, the extend of this oscillatory zone is about $10^{-6}$ of the crack length. England [3] and Rice and Sih [4] also encountered the oscillatory behavior of the displacements near the crack tip. To overcome the physically inadmissible oscillatory displacements at the crack tip, Atkinson [5] proposed two models. In the first one, the interface is replaced by a strip of finite thickness and the crack is put in this new strip. The strip is homogeneous and its modulus of elasticity is different than the moduli of the two materials. Since the crack is in a homogeneous strip, the previous oscillatory behavior didn't occur. This model hasn't gain much attention. In the second model, the interface layer has a varying modulus, being equal to the modulus of the surrounding bodies at the extremities. Since the modulus is continuous, the oscillatory behavior is again eliminated. This is a more realistic model since combinations of materials will produce diffuse interfaces. In the Comninou model [6, 7, 8], the crack surfaces are allowed to get into contact, the contact zone is not known and has to be solved as a part of the problem. Besides analytical work, not many finite element work has been done in this area. Van der Zande and Grootenboer [9] calculated mode II stress intensity factors for an interface crack with contact zones using finite element methods. Dattaguru [10, 11] used finite element approach to predict the mode II dominance at the crack tip, as suggested by Comninou. Vankatesha [12] employed a modified crack closure integral technique and finite elements to calculate the strain energy release rates in a bimaterial problem with large crack-tip contact zones. Vankatesha even observed the crack-tip surface penetration with a very refined mesh and

using 4 noded linear elements. Liu [13] used a 3-D finite element analysis to model plastic encapsulated packages including large scale contact and large deformation effects.

# CHAPTER 2

# FINITE ELEMENT FORMULATION

## 2.1 OVERVIEW OF THE FINITE ELEMENT METHOD

Finite Element Method is a very popular and widely used method to analyze complicated problems that are either too difficult to solve with analytical methods or no known solution exists at hand. Generally, analytical methods are efficient for simple geometries and boundary conditions, but in real life most of the problems in hand are quite complex because of the irregular geometries, different material combinations and properties, time and/or space dependence of properties, etc. Finite Element Method offers a reliable and efficient tool to tackle complex problems. Although it's used in various disciplines like heat transfer, fluid mechanics, electronics, only solid mechanics formulation will be used in this work.

In a general point of view, it's a numerical tool to transform a set of differential equations into a set of linear algebraic equations and solve for the unknowns at certain points in the continuum called *node*. The continuum in question is discretized using some *elements* composed of *nodes*. The choice of elements to be used in discretization is left to the user and to the capacity of the finite element code but they mainly differ in the order of interpolation functions used within the element to describe the variation of the unknown to be solved, which are in this case the displacements. For example a quadratic element assumes the displacements to vary as a second degree polynomial while a cubic element assumes a $3^{rd}$ degree variation.

### 2.1.1 FINITE ELEMENT EQUATIONS

The algebraic equations which are also called the *finite element equations* are obtained by minimizing the potential energy in the system. For an elastic body the potential energy can be expressed as:

$$\Pi_p = \int_v \left(\tfrac{1}{2}\{\epsilon\}^T[E]\{\epsilon\} - \{\epsilon\}^T[E]\{\epsilon^0\} + \{\epsilon\}^T\{\sigma^0\}\right)dV$$

$$- \int_v \{u\}^T\{F\}dV - \int_s \{u\}^T\{T\}dS - \{D\}^T\{P\} \qquad (2.1)$$

where

| | |
|---|---|
| $\{u\}$ | displacement field |
| $\{\epsilon\}$ | strain field |
| $\{\epsilon^0\}$ | initial strain field |
| $\{\sigma^0\}$ | initial stress field |
| $[E]$ | material property matrix |
| $\{F\}$ | body forces |
| $\{T\}$ | surface tractions |
| $\{D\}$ | nodal displacements |
| $\{P\}$ | nodal external forces |

Upon discretization of the domain using shape functions (interpolating functions), the minimization of $\Pi_p$ with respect to the nodal displacements gives:

$$\frac{\partial \Pi_p}{\partial D} = 0 \qquad \Rightarrow \qquad [K]\{D\} = \{R\} \qquad (2.2)$$

where $[K]$ is the global stiffness matrix and $\{R\}$ is the global load vector. Eq 2.2 represents the well-known most general form of the finite element equations. It is obvious that trying to compute the stiffness matrix and load vector of a model in the

global sense would most probably be impossible or inefficient. In most of the cases, $[K]$ and $\{R\}$ are calculated in an element-by-element basis and then assembled to form the global stiffness matrix and global load vector, which will then be solved for the unknown displacements.

The element stiffness matrix $[k]$ is known to be:

$$[k] = \int [B]^T [E][B] \tag{2.3}$$

where $[B]$ can be expressed as

$$[B] = \begin{bmatrix} \frac{\partial}{\partial x} & 0 & 0 \\ 0 & \frac{\partial}{\partial y} & 0 \\ \frac{\partial}{\partial y} & \frac{\partial}{\partial x} & 0 \end{bmatrix} [N] \tag{2.4}$$

and $[N]$ is the shape function matrix.

The element load vector $\{r\}$ can be expressed as:

$$\{r\} = \int [B]^T [E]\{\epsilon^0\} dV - \int [B]^T \{\sigma^0\} dV$$

$$+ \int [N]^T \{F\} dV + \int [N]^T \{T\} dS \tag{2.5}$$

Note that the integrals have to be calculated numerically. Most often they are calculated with the Gaussian integration technique which is the most popular one. For cubic elements that are used in this work, 4 point Gaussian integration is sufficient. It will be shown later in Appendix C that in order to get good results, higher Gaussian integration points are necessary for enriched elements.

Once the element stiffness matrices and load vectors are found, the global finite element equations are assembled and solved. In this work, the frontal solution technique is used as the solution algorithm.

### 2.1.2 FRONTAL SOLUTION TECHNIQUE

Solution of the finite element equations is the step in the finite element analysis that takes the most computation time, therefore an intelligent choice has to be made. There are several solution techniques and frontal solution technique is one of the best among them.

Once the element stiffness matrices and load vectors are calculated, the common way is to assemble the global stiffness matrix and load vector using the connectivities of the elements and then solve the whole system of equations at the same time. For this purpose there are several available solution techniques such as LU decomposition, Gaussian elimination etc. The frontal solution technique, which was first developed by Irons [14], offers a distinct advantage: the assembly and solution takes place at the same time. Therefore there is no more need to construct the whole big global matrix and load vector, saving a lot of memory space. Considering the fact that most often physical problems (not enough memory, disk space) are more important that CPU time consumed by the code, this feature enables the solution of larger models using less hardware.

To understand this solution technique, one should understand the concepts of *active node* and *deactivation* of a node. The frontal solver will operate on an element-by-element basis. It will start with the 1$^{st}$ element and finish with the last one. The element stiffness and load vector will be calculated then assembled. If the node will

not appear in the following elements, the information corresponding to that node will be eliminated from the "front" and it will be written to a disc file. An *active* node is a node whose stiffness and load terms are still in the memory. A *deactivated* node is a node whose corresponding terms are eliminated from the equations and written to a disc file. To clarify these concepts, Figure 2.1 shows a simple case with linear triangular elements.

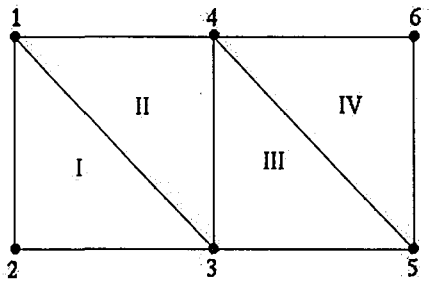


**Figure 2.1:** Simple example for the frontal solver

The element connectivities of the elements is shown in Table 2.1:

| element # | connectivity |
|---|---|
| I | 1, 2, 3 |
| II | 3, 4, 1 |
| III | 3, 5, 4 |
| IV | 5, 6, 4 |

**Table 2.1:** Connectivities of the elements in the simple example for the frontal solver

It can be seen that the node 1 only appears in the elements I and II, node 2 only in I, node 3 in I, II and III and so on. The first and last elements the nodes appears are shown in Table 2.2.

| node # | first element | last element |
|---|---|---|
| 1 | I | II |
| 2 | I | I |
| 3 | I | III |
| 4 | II | IV |
| 5 | III | IV |
| 6 | IV | IV |

**Table 2.2:** First and last elements in which the nodes appear

This table truly shows the *life* of a node in the *front*. The node 1 becomes active in element I and deactivated in II, the node 3 becomes active in I and is deactivated in III, and so on. It is clear that the information for a certain node stays in the memory only for a certain time, as long as that node is active. When it gets deactivated, the information is saved in disc for backward substitution, the space occupied by that node in the *front* will be filled with another active node, to save some more memory space.

Note that in the frontal solution technique there is no limit whatsoever on the size of the element stiffness matrices. This enables the use of different kinds of elements within the same model without a problem. Whatever the size of the stiffness matrix or load vector, everything will pass through the same steps and there will be no problem as long as the elements are *connected* correctly. This feature will be useful in contact analysis when dealing with target elements with stiffness matrices larger than 24x24 for example. The only restriction is that the element stiffness matrices has to be symmetric.

When all the elements are handled, a backward substitution process extracts the information written to disc and calculates the required unknowns, which are the nodal displacements in our case.

## 2.2 FINITE ELEMENT FORMULATION - CRACK PROBLEM

Catching field singularities is a great deal of concern in computational mechanics. By field singularity it is meant that one or many values of the field is infinite at certain locations. For example crack tips have stress singularities. Several techniques exist to correctly catch the singularity near a crack tip. Enriched element formulation, which is used in this work, is one of the most accurate ones.

The enriched element formulation dates back to the work of Benzley [15]. Although the technique can be extended to other forms of singularities as well, only crack tip singularities will be discussed in this work. In enriched element formulation, the main objective is to calculate the stress intensity factors $K_I$, $K_{II}$ and then to find the other important quantities such as total strain energy release rate and phase angles, which are calculated using the stress intensity factors. Therefore, once the stress intensity factors are found correctly, everything else of interest is found in a straight forward fashion. Using enriched formulation, $K_I$ and $K_{II}$ are implemented into the finite element equations and they are solved at the same time as the displacements.

At this point, two definitions need to be made. *Enriched elements* are defined as the elements in the model that are touching the crack tip node, i.e. the crack tip node is in their connectivity. *Transition elements* are defined as the elements surrounding the enriched elements. Transition elements are needed because the displacement interpolation within these two elements are different from each other. In other words, transition elements helps the smooth passage from the enriched interpolation to the usual cubic interpolation. The way in which enriched and transition elements are placed around the crack tip can be seen in Figure 2.2.

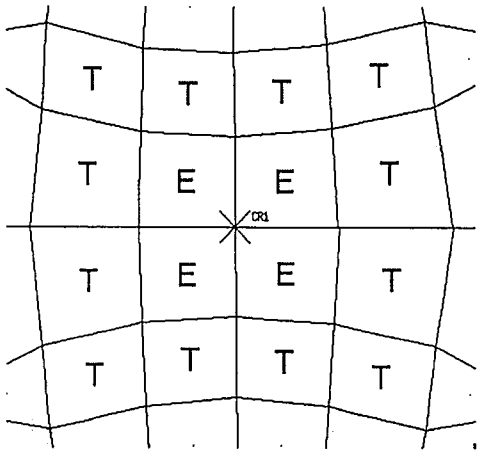


**Figure 2.2:** Placement of special crack tip elements around the crack tip

For enriched elements, the asymptotic displacement fields are discretized as:

$$u(s,t) = \sum_{j=1}^{nnpe} N_j(s,t)u_j + Z_0(s,t)K_I\left\{f_1(s,t) - \sum_{j=1}^{nnpe} N_j(s,t)f_{1j}\right\} + Z_0(s,t)K_{II}\left\{g_1(s,t) - \sum_{j=1}^{nnpe} N_j(s,t)g_{1j}\right\} \quad (2.6)$$

$$v(s,t) = \sum_{j=1}^{nnpe} N_j(s,t)v_j + Z_0(s,t)K_I\left\{f_2(s,t) - \sum_{j=1}^{nnpe} N_j(s,t)f_{2j}\right\} + Z_0(s,t)K_{II}\left\{g_2(s,t) - \sum_{j=1}^{nnpe} N_j(s,t)g_{2j}\right\} \quad (2.7)$$

where

| | |
|---|---|
| $K_I$ | = mode I stress intensity factor |
| $K_{II}$ | = mode II stress intensity factor |
| $Z_0$ | = zeroing function evaluated at (s,t) |
| $nnpe$ | = number of nodes per element |
| $f_1, f_2, g_1, g_2$ | = asymptotic displacement functions |

These equations can be further simplified by defining $f_1^*$, $f_2^*$, $g_1^*$ and $g_2^*$ as:

$$\begin{aligned} f_1^*(s,t) &= f_1(s,t) - \sum_{j=1}^{\text{nnpe}} N_j(s,t) f_{1j} \\ f_2^*(s,t) &= f_2(s,t) - \sum_{j=1}^{\text{nnpe}} N_j(s,t) f_{2j} \\ g_1^*(s,t) &= g_1(s,t) - \sum_{j=1}^{\text{nnpe}} N_j(s,t) g_{1j} \\ g_2^*(s,t) &= g_2(s,t) - \sum_{j=1}^{\text{nnpe}} N_j(s,t) g_{2j} \end{aligned} \tag{2.8}$$

Finally

$$u(s,t) = \sum_{\text{j=1}}^{\text{nnpe}} N_j(s,t) u_j + Z_0(s,t) f_1^*(s,t) K_I + Z_0(s,t) g_1^*(s,t) K_{II} \tag{2.9}$$

$$v(s,t) = \sum_{\text{j=1}}^{\text{nnpe}} N_j(s,t) v_j + Z_0(s,t) f_2^*(s,t) K_I + Z_0(s,t) g_2^*(s,t) K_{II} \tag{2.10}$$

Note that the asymptotic displacement expressions $f_1$, $f_2$, $g_1$, $g_2$ are evaluated at the 12 nodes and also at the isoparametric coordinates s and t where the displacement is desired. One can see easily that at the nodes starred expressions $f_1^*$, $f_2^*$, $g_1^*$, $g_2^*$ becomes zero and the same expressions for enriched elements and normal cubic elements are found.

$Z_0$ is the zeroing function which smoothes the interelement compatibility between enriched elements and the surrounding elements. This function has the value of 1 everywhere in the enriched element, takes a value between 1 and 0 in the transition elements and is 0 everywhere in usual isoparametric elements. The variation of the zeroing function in the transition element can also be decided by the user, although generally a linear variation in terms of isoparametric coordinates is used.

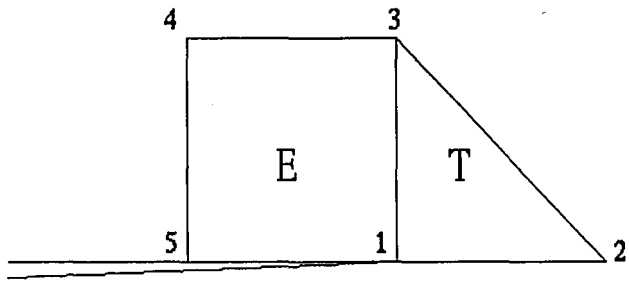


**Figure 2.3:** Illustration to show the zeroing function $Z_0$

The zeroing function can be better understood with the help of Figure 2.3, which shows an enriched and a transition element at the crack tip. The zeroing function $Z_0$ will take the value of 1 everywhere in the enriched element, i.e at the nodes 1, 3, 4, 5 and also between them. It will then take a number between 1 and 0 in the transition element. Let's assume $Z_0$ varies linearly in the transition element, then it will be 1 at nodes 1 and 3, 0 at node 2 which touches a normal element and varies linearly between 1 and 0 for other points in the transition element.

Let's assume we are dealing with 12 noded quadrilateral cubic elements. In an element-by-element basis, it can be seen that for each enriched element, other than 24 unknown displacements, 2 unknown stress intensity factors are introduced into the equations. This enables the calculation of the stress intensity factors at the same time

as the displacements. In a way, the crack tip singularity is *embedded* into the finite element equations. The stress intensity factors will be represented in the element as the degrees of freedom of a "*virtual node*". If $n$ physical nodes and a crack exits, a virtual node is added making the number of total nodes $n+1$. The unknowns corresponding to the physical nodes are the displacements whereas the unknowns associated with that virtual node will be the mode I and mode II stress intensity factors. One should note that the conventional finite element formulation is exactly the same in enriched element formulation, with the addition of a few stiffness and load terms. Again the stiffness matrix and load vector of the constituent elements will be calculated. The system of resulting finite element equations will be solved for the unknown $2n$ displacements and 2 stress intensity factors. The main difference is really the number of unknowns solved.

## 2.3 FINITE ELEMENT FORMULATION - CONTACT PROBLEM

Computational contact mechanics analysis is somewhat more difficult than static linear elastic analysis in the sense that the problem is truly nonlinear, like plasticity or large deformation analysis. What makes it nonlinear is quite different than the later ones. In plasticity material nonlinearity is present, in large deformation analysis the nonlinear strain terms neglected in linear elastic case are taken into consideration. The contact analysis is nonlinear because the stiffness matrix is dependent on the displacements and the forces. This is because of the fact that generally the contact zone is not known a priori. The contact zone is known only if the coefficient of friction is so high that all the contactor nodes are in sticking contact and initially all the contactor nodes are touching the target surfaces. If all nodes are in sticking contact initially and at the end, the nodes stick to the point on the target

surface that they touch and remains at the same isoparametric coordinate all the time. For every other cases, the problem is nonlinear and therefore should be handled incrementally. Another reason for nonlinearity is friction, one does neither know the correct friction forces present in the system a priori nor the correct contact state at contactor nodes. These will be found by iteration. The incremental procedure will be discussed later in this chapter.

In the finite element formulation, the contact problem will be inserted in the set of equations in the form of constraints. For this purpose, two common methods are used, namely Lagrange Multipliers and Penalty Function Method. Whichever formulation is used to add the necessary constraints, the formulation mainly consists of 6 parts :

- ▶ Detecting of all new contacting nodes,
- ▶ Forming necessary constraint equations in the form of matrices,
- ▶ Forming the finite element equations ,
- ▶ Solving for the unknowns displacements,
- ▶ Calculating the contact forces at contactor nodes and target element nodes,
- ▶ Determining of the correct contact state at contactor nodes.

These steps are performed in the given order for each iteration in each increment until convergence occurs, i.e. the incremental displacements (or the incremental forces) goes to zero. These steps will be discussed in detail in the following sections.

### 2.3.1 COMMON METHODS AND PREVIOUS WORK

Basically, two popular methods exists in adding the necessary constraints to the system, these are the penalty function method and the Lagrange multipliers method. These two methods will be presented in detail in this thesis, with special attention given to the penalty function method. There are also some other formulations besides these two, which are in fact modified forms of either of them. Perturbed Lagrange multipliers method [16], augmented Lagrange multipliers method [17,18] and the Uzawa algorithm [19] can be given as examples.

Numerous work has been done in the past on the finite element modeling of the contact problem, in most of them either the Lagrange multipliers or the penalty function method have been used. The first use of the Lagrange multipliers in the computational contact problems dates back to the work of Hughes et al. [20], who investigated the problems of elastic contact and elastic impact. Node-to-node contact was used in the contact zone. Okamoto and Nakawaza [21] investigated frictional contact with the use of unsymmetric stiffness matrix. Using an incremental approach, Guerra and Browning [22] showed that both Lagrange multipliers and penalty function method can be used in imposing the required displacement constraints. A general two-dimensional contact algorithm has been developed by Bathe and Chaudhary [23] using Lagrange multipliers. Node-to-node contact was not necessary and an elaborate algorithm to decide on the contact forces and the contact status has been proposed. First papers on using penalty function method in contact analysis used some "gap elements" at the contact zone. White and Enderby [24] and Stadter and Weiss [25] used gap elements whose stiffnesses were varying to account for penetrations and separations. Mazurkiewicz and Ostachowicz [26], Zolti and [27], Ostachowicz [28] used gap elements including frictional effects. The concept of

tangential stiffness was introduced, being the product of the gap element stiffness and the coefficient of friction. Their method was susceptible of "locking" when a high penalty function was used. Cheng and Kikuchi [29] used a non-Coulombian friction model together with penalty function method. Problems using gap elements with the large displacement formulation, such as the poor Jacobians of the highly distorted gap elements have been investigated by Padovan et al [30]. Yagawa et al [31] applied the penalty function method to fracture mechanics problems using 3-D brick elements. Motthershead et al [32] proposed a method to handle curved surfaces using isoparametric coordinate transformation, successful using both Lagrange multipliers method and penalty function method. The formulation used in this paper is used as the main focus in this thesis.

### 2.3.2 CONTACT SEARCHING

Contact searching is one of the most important steps in computational contact mechanics. Before getting into the contact searching algorithms, the concepts of *contactor node* and *target element* has to be explained. Contact always occurs between two surfaces in the continuum, therefore the contact zone can be either a point or a line. These two surfaces will be defined while creating the geometry, to prevent the program from checking every possible surface combinations. One of the surfaces is called *contactor surface* and the other being called *target surface.* The definition of these two surfaces, whether contactor or target is quite arbitrary. Consequently, the nodes on the contactor surfaces are called *contactor nodes* and the elements on the target surfaces are called *target elements.* Figure 2.4 shows contactor nodes and a target element of a simple model with 12 noded quadrilateral elements.

Contact searching can be described as checking contactor node - target element pairs in order to detect new penetrating nodes, then calculating the isoparametric coordinates s and t (also called corrected coordinates) on the surface of the target element to which the contactor node will be "pushed" in the next iteration.

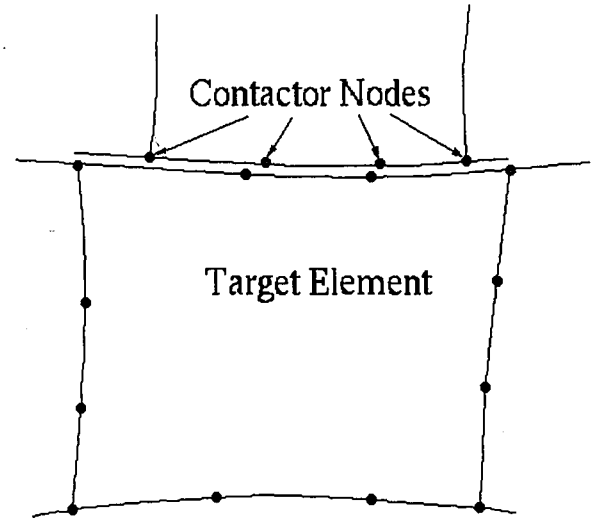


**Figure 2.4:** Definition of contactor nodes and target elements

Precise detection of new contacting nodes is essential in order to satisfy compatibility condition (no contactor node can penetrate into a target element) up to desired accuracy. Calculation of the corrected coordinates is necessary in forming the necessary constraint matrices, no matter which constraint formulation is used.

In earlier work, generally the contact searching was performed using a line search technique. Figure 2.5 shows a target element and a contactor node C, $C_i$ being outside the element in the $i^{th}$ iteration and $C_{i+1}$ being in the element in the $(i+1)^{th}$ iteration. A new contact is said to occur when the cross product $\overrightarrow{CA} \times \overrightarrow{CB}$ changes sign, or in other words

$$\left(\left(\left(\overrightarrow{C_iA}\right) \times \left(\overrightarrow{C_iB}\right)\right) \cdot \vec{k}\right)\left(\left(\left(\overrightarrow{C_{i+1}A}\right) \times \left(\overrightarrow{C_{i+1}B}\right)\right) \cdot \vec{k}\right) \leq 0 \qquad (2.11)$$

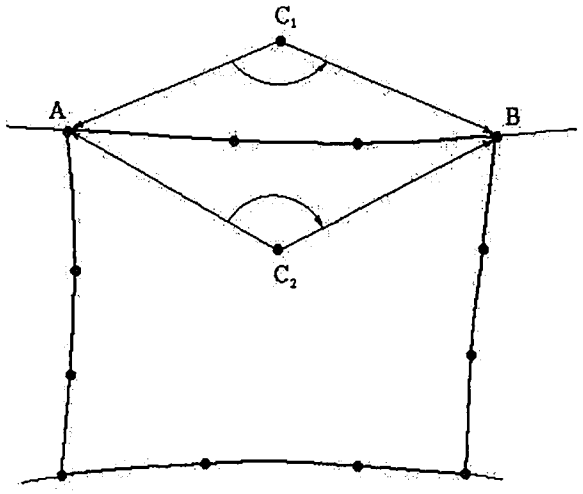


**Figure 2.5:** Line search algorithm model

The main handicap of this method is its inefficiency in modeling curved surfaces which are encountered in many cases like Hertzian contact of cylinders, etc. In order to model curved surfaces, one must use a very fine mesh near the contact zone, if it's known. If it's not known, the whole target surface has to be refined which will put a lot of useless elements and increase CPU time considerably. Also, this technique assumes that the contactor node has followed a linear path in Cartesian coordinates between $i^{th}$ and $(i+1)^{th}$ iterations. In order to satisfy this assumption, one should use a lot of load increments. This, again, will increase greatly the CPU time spent in solving the problem.

To remedy these handicaps, an isoparametric formulation, which will be used in this thesis, has been proposed by Mottershead et al. [32] for contact searching. Before each iteration, all possible contactor nodes will be checked for penetration with all possible target elements. For this purpose, the Cartesian coordinates of the

contactor node will be transformed into isoparametric coordinates s and t of the target element in question. If $-1 \leq s \leq 1$ and $-1 \leq t \leq 1$ then penetration has occurred and action must be taken to "push" this contactor node back to the target element surface at the next iteration. Note that even if the contactor node is on the surface of the target element, this is considered as penetration and constraint matrices has to be formed such that the contactor node will stay on the surface at the end of the next iteration.

The position of a contactor node can be described using the target element shape functions as :

$$x_c = \sum_{i=1}^{\text{nnpe}} N_i(s_c, t_c)\, x_i \tag{2.12}$$

$$y_c = \sum_{i=1}^{\text{nnpe}} N_i(s_c, t_c)\, y_i \tag{2.13}$$

where the subscript c denote contactor node and the summation goes from i = 1 to the number of nodes per element, nnpe. We have to solve this equation for the unknown values $s_c$ and $t_c$ which will be the isoparametric coordinates of the contactor node with respect to the target element. For linear elements this equation is linear and the solution is straight forward, but for quadratic and cubic elements it's nonlinear and therefore it can be solved using a Newton-Raphson iterative approach.

Starting from an initial guess $(s_0, t_0)$, linearizing this nonlinear equation in the proximity of the initial guess, we can write :

$$x_c - x_0 = \sum_i^{\text{nnpe}} \left( \frac{\partial N_i}{\partial s}(s_0, t_0) x_t \triangle s_1 + \frac{\partial N_i}{\partial t}(s_0, t_0) x_t \triangle t_1 \right) \tag{2.14}$$

$$y_c - y_0 = \sum_i^{\text{nnpe}} \left( \frac{\partial N_i}{\partial s}(s_0, t_0) y_t \triangle s_1 + \frac{\partial N_i}{\partial t}(s_0, t_0) y_t \triangle t_1 \right) \tag{2.15}$$

where $x_0$ and $y_0$ represents the initial guess in Cartesian coordinates and the derivatives of the shape functions are evaluated at the initial guess $(s_0, t_0)$. This can be rewritten as :

$$\begin{pmatrix} x_c - x_0 \\ y_c - y_0 \end{pmatrix} = \boldsymbol{J}^{\mathrm{T}}(s_0, t_0) \begin{pmatrix} \triangle s_1 \\ \triangle t_1 \end{pmatrix} \tag{2.16}$$

Solving this for $\triangle s_1$ and $\triangle t_1$ :

$$\begin{pmatrix} \triangle s_1 \\ \triangle t_1 \end{pmatrix} = (\boldsymbol{J}^{\mathrm{T}}(s_0, t_0))^{-1} \begin{pmatrix} x_c - x_0 \\ y_c - y_0 \end{pmatrix} \tag{2.17}$$

where again the inverse of the Jacobian matrix is evaluated at the first initial guess $(s_0, t_0)$. The initial guess will then be updated as

$$\begin{pmatrix} s_1 \\ t_1 \end{pmatrix} = \begin{pmatrix} s_0 + \triangle s_1 \\ t_0 + \triangle t_1 \end{pmatrix} \tag{2.18}$$

For the next iteration, the equation is

$$\begin{pmatrix} \triangle s_2 \\ \triangle t_2 \end{pmatrix} = (\boldsymbol{J}^{\mathrm{T}}(s_1, t_1))^{-1} \begin{pmatrix} x_c - x_1 \\ y_c - y_1 \end{pmatrix} \tag{2.19}$$

The iteration continues until a predefined accuracy has been achieved, i.e.

$$\left\| \begin{pmatrix} x_c - x_k \\ y_c - y_k \end{pmatrix} \right\| \leq tolerance \tag{2.20}$$

Starting with $(s_0 = 0, t_0 = 0)$, this iteration converges quite rapidly, if it converges. It has been observed that if the contactor node is too far away from the target element then the iteration diverges. Therefore in addition to the tolerance, a limit to the number of iterations is necessary. This divergence behavior does not cause any problems since it only occurs when the penetration is already impossible. This can be fixed by employing an "envelope" technique, i.e. only the contactor nodes that are "close" to the target element will be checked for penetration. By "close" it is meant that the contactor node falls inside a circle, whose center is on the target element and the radius is specified by a certain number (let's say 5) times the average target element size.

Once the contactor node has been identified as in penetration, now the corrected isoparametric coordinates $s$ and $t$, i.e. the coordinates on the surface of the target element that the contactor node will be pushed to during the next iteration, has to be determined. The technique is dependent on the contact status of the contactor node. The determination of the contact status is explained in detail in section 2.3.6. If the contactor node is in sticking contact with the target element, then a line is drawn in isoparametric plane joining $(s_{i-1}, t_{i-1})$ , coordinates of the contactor node in $(\mathrm{i}-1)^{\mathrm{th}}$ iteration, and $(s_i, t_i)$, coordinates in $\mathrm{i}^{\mathrm{th}}$ iteration, as shown in Figure 2.6. The coordinates of the point of intersection between this line and the surface of the target element, corresponding to the point $a$, is the corrected isoparametric coordinates. This point really represents the position of first contact as the contactor node penetrates into the target element.

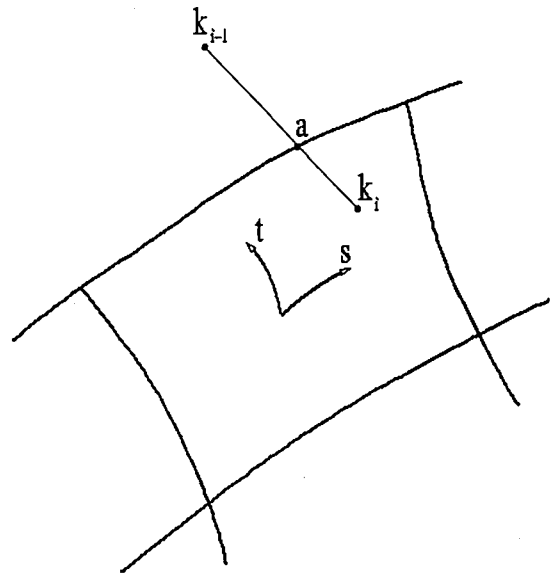


**Figure 2.6:** Illustration to determine the corrective coordinates, in sticking contact

A special case arises when the contactor node is in sliding contact with the target element. Since the node is constrained to move only in a direction tangent to the target element surface, gaps may open up or overlaps occur. These should be taken care of by including them into the corrective displacements. In sliding, the procedure is to draw a line in isoparametric plane from the current position of the contactor node perpendicular to the target surface, as shown in Figure 2.7. The coordinates of the point of projection, which corresponds to the point *a* on the figure, is the corrected isoparametric coordinates. In other words, the contactor node will be projected back to the surface in the next iteration. Note that even though a gap may occur, the contactor node is still assumed to be in contact with the target element since separation can be predicted only according to the contact forces, as explained in section 2.3.6. It should also be noted that, when a contactor node slides on the surface of a target element, it can pass to the surface of an adjacent target element. Obviously a good bookkeeping algorithm is necessary to correctly assign contactor nodes with target elements penetrated.

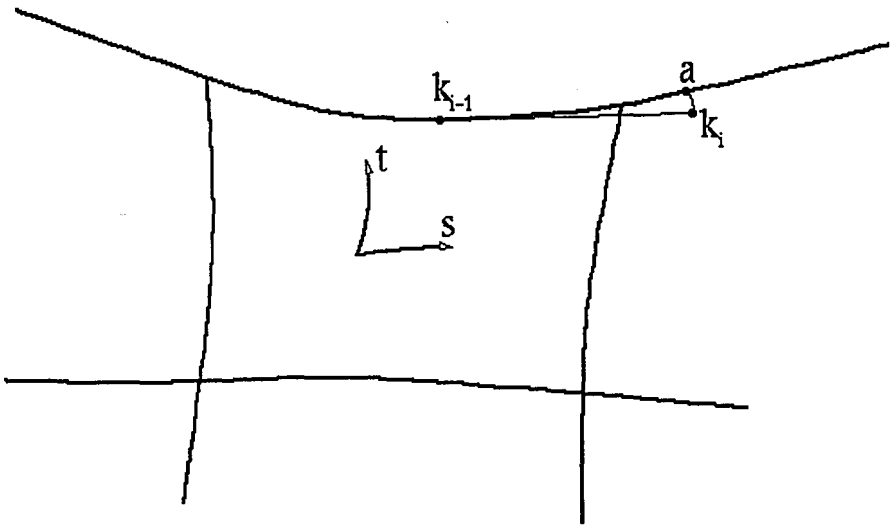


**Figure 2.7:** Illustration to determine the corrective coordinates, in sliding contact

Whether the contactor node is in sticking contact or in sliding contact, in order to find the corrected isoparametric coordinates to which the node will be pushed back at the next iteration, one should know from which surface the contactor node will penetrate into the target element. An algorithm is developed for this purpose: during the mesh generation, one should also specify the contactor and target surfaces which are susceptible of contact. From this information, the mesh generator will find the nodes on the contactor surfaces and assign them as *contactor nodes*, will find the elements on the target surface and assign them as *target elements*. Together with this information, a "surface information" will be generated for each contactor and target element. This information defines the surface from which contact may occur. The convention of the surface information assigning is shown in Figure 2.8 for a 12 noded quadrilateral element and a 10 noded triangular element.

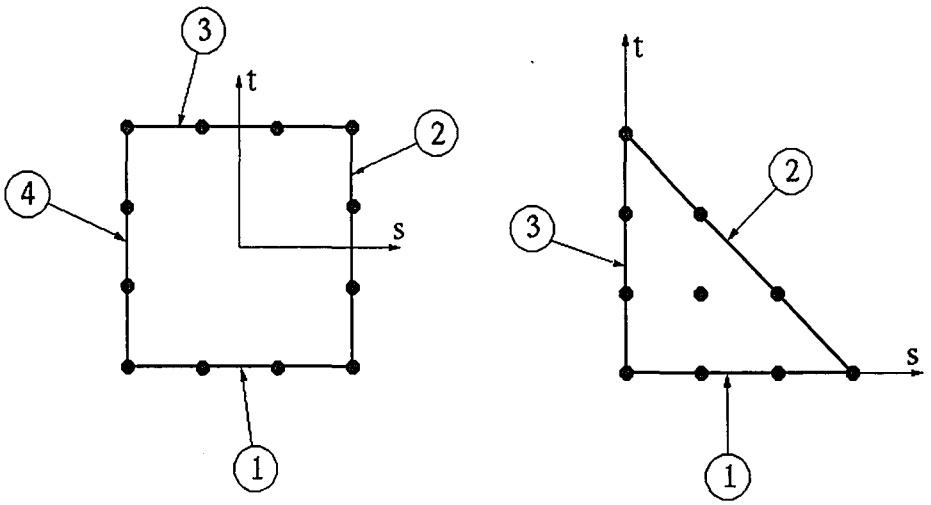


**Figure 2.8:** Convention of "surface information" for 12 and 10 noded elements.

The numbers in the circles are the surface information corresponding to each side. As it can be seen, this surface information is assigned according to which surface in isoparametric plane does that surface in Cartesian plane corresponds to. The code knows at all times that for 1, $t = -1$ on the surface if the element is a quadrilateral and $t = 0$ if the element is triangular, for 2 $s = 1$ for a quadrilateral and $s + t = 1$ for a triangular and so on. This provides an additional equation that will be used in finding the corrected isoparametric coordinates.

An exception arises when the target element in question touches the target element at only one point. This case is shown in Figure 2.9. In the $i^{th}$ iteration, the contactor node $k$ is not in contact. In the following iteration it penetrates into the element I. The corrected isoparametric coordinates are needed to push the element back to the surface, but element I has no surface on the target surface. It touches the target surface at only a point. If the contactor node is assigned to be in contact with element I, this will give wrong results. If one draws an isoparametric line connecting the positions of the contactor node in the $i^{th}$ and $(i+1)^{th}$ iterations, it will intersect

element II on its surface that is on the target surface. Therefore node $k$ should be assigned to be in contact with element II.

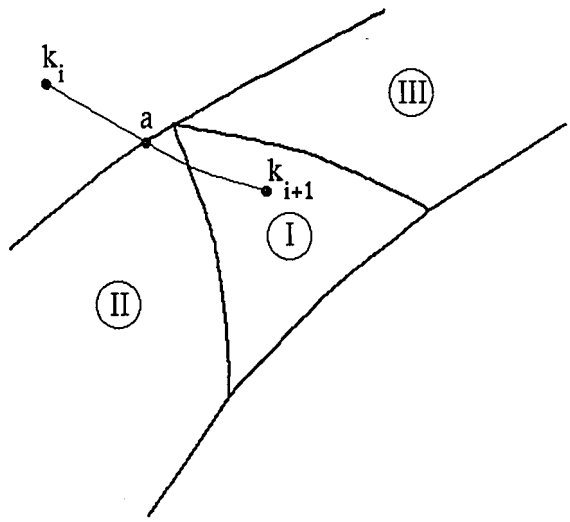


**Figure 2.9:** Illustration to determine the corrected coordinates, when target element touches the target surface at only one point

This exception is handled in the code as follows: during mesh generation, the minus node number of the touching node is assigned as the surface information of the elements that touch the target/contactor surface at only a point. Thereby whenever an element is encountered with a negative surface information, the code automatically knows that it is an element touching the target/contactor surface at only a point and therefore checks the adjacent elements (like element II) for contact.

## 2.3.3 CONSTRAINT EQUATIONS

Once all the penetrating nodes are determined and corresponding corrected coordinates calculated, these data has to be implemented into the finite element equations. For this purpose constraint equations, with which the overlaps will be

removed, will be formulated. Consider the case where a contactor node penetrates into a target element at $i^{th}$ iteration.

The corrected coordinates $s$ and $t$ being determined, the overlap has to be compensated in the $(i+1)^{th}$ iteration by returning the node to $(s,t)$. This can be done by applying a *corrective displacement vector* $\boldsymbol{\delta}$, which is opposite to the vector of overlaps. In Figure 2.10, the vector $\overrightarrow{a_i a_{i+1}}$ is the displacement vector of the corrected isoparametric point, i.e.

$$\overrightarrow{a_i a_{i+1}} = \sum_i N_i(s,t)\vec{u_i}$$

Similarly $\overrightarrow{k_i k_{i+1}}$ is the displacement vector of the contactor node and $\overrightarrow{k_i a_i}$ is the corrective displacement vector $\boldsymbol{\delta}$, i.e

$$\overrightarrow{k_i k_{i+1}} = \vec{u_c}$$

$$\overrightarrow{k_i a_i} = \boldsymbol{\delta}$$

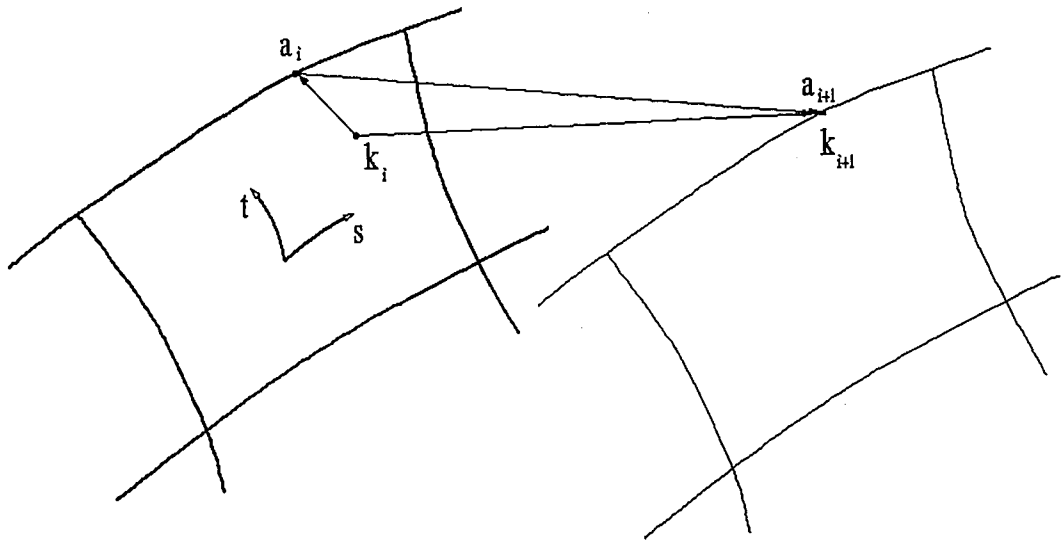


**Figure 2.10:** Applying a corrective displacement to push the node back to the surface

Therefore the components of $\delta$ becomes :

$$\delta_x = u_c - \sum_i N_i(s,t)u_i \qquad (2.21)$$
$$\delta_y = v_c - \sum_i N_i(s,t)v_i$$

which can be put into the form

$$\delta_i - \begin{Bmatrix} u_c \\ v_c \end{Bmatrix}_{i+1} + \begin{bmatrix} N_1 & 0 & N_2 & 0 & \dots & \dots & N_{12} & 0 \\ 0 & N_1 & 0 & N_2 & \dots & \dots & 0 & N_{12} \end{bmatrix} \begin{Bmatrix} u_T \\ v_T \end{Bmatrix}_{i+1} = 0 \qquad (2.22)$$

or

$$\delta_i - \begin{Bmatrix} u_c \\ v_c \end{Bmatrix}_{i+1} + [\boldsymbol{I} \otimes \mathrm{Ns}]_i \begin{Bmatrix} u_T \\ v_T \end{Bmatrix}_{i+1} = 0 \qquad (2.23)$$

where $\boldsymbol{I}$ is the identity matrix, $\otimes$ is the Kronecker product.

For sliding contact, only the tangential displacement of the contactor node should be constrained, therefore the constraint equation will be

$$\mathbf{n}^T\delta_i - \mathbf{n}^T \begin{Bmatrix} u_c \\ v_c \end{Bmatrix}_{i+1} + \mathbf{n}^T[\boldsymbol{I} \otimes \mathrm{Ns}]_i \begin{Bmatrix} u_T \\ v_T \end{Bmatrix}_{i+1} = 0 \qquad (2.24)$$

where $\mathbf{n}$ is the normal vector evaluated at the corrected isoparametric coordinates. Combining both equations, one can formulate both sticking and sliding contact constraint equations in the following form:

$$\delta_i - \boldsymbol{L}_i \triangle \boldsymbol{u}_{i+1} \qquad (2.25)$$

where $L_i$ is called the *constraint matrix*, $\triangle u_{i+1}$ represents the incremental displacements that will be solved in the (i+1)$^{th}$ iteration, $\delta_i$ is the vector containing all of the corrective displacements.

### 2.3.4 FINITE ELEMENT EQUATIONS

The well known form of the finite element equation derived by minimizing the potential energy is:

$$Ku = F \tag{2.26}$$

where $K$ is the stiffness matrix, $F$ is the consistent force vector and $u$ the resulting displacements. This formulation is valid only when the stiffness matrix $K$ is not a function of displacements nor the loads. But in the case of contact mechanics, the components of $K$ turns out to be functions of both displacements and loads, making the equation 2.26 nonlinear. Instead, an incremental approach should be taken. Seeking equilibrium iteratively, one can write the potential energy under the incrementally applied load:

$$\Pi^p = \tfrac{1}{2}(u^p)^T K^p u^p + (u^p)^T \sum_{h=1}^{p-1} K^h u^h - (u^p)^T \sum_{h=1}^{p-1} f^h \tag{2.27}$$

Minimizing we get

$$0 = \sum_{h=1}^{p} K^h u^h - \sum_{h=1}^{p} f^h \tag{2.28}$$

Since the stiffness is a function of displacements and loads, we can approach equilibrium iteratively as

$$\boldsymbol{K}_i^h \triangle \boldsymbol{u}_{i+1}^p = \sum_{h=1}^{p} \boldsymbol{f}^h - \sum_{h=1}^{p-1} \boldsymbol{K}^h \boldsymbol{u}^h - \boldsymbol{K}_i^p \sum_{j=1}^{i} \triangle \boldsymbol{u}_j^p \tag{2.29}$$

Now, we have to implement the constraint equations into 2.29. For this purpose two popular methods exist, namely *penalty function method* and *Lagrange multipliers method.*

In penalty function method, the one used in this thesis, minimizing the potential energy with respect to the displacements one gets :

$$\begin{aligned}\left[\boldsymbol{K}_i^p + \alpha \boldsymbol{L}_i^T \boldsymbol{L}_i\right] \triangle \boldsymbol{u}_{i+1}^p = & -\sum_{h=1}^{p-1} \boldsymbol{K}^h \boldsymbol{u}^h - \boldsymbol{K}_i^p \sum_{j=1}^{i} \triangle \boldsymbol{u}_j^p + \alpha \boldsymbol{L}_i^T \boldsymbol{\delta}_i \\ & + \sum_{h=1}^{p} \boldsymbol{f}^h + \sum_{h=1}^{p-1} \boldsymbol{c}^h + \sum_{j=1}^{i} \boldsymbol{c}_j^p \end{aligned} \tag{2.30}$$

In this equation, the first two terms in the right hand side represent the vector of element stress resultants, the third term represent the vector of forces to be applied at contactor nodes for the removal of overlaps. Fourth term is the incrementally applied external loads (which is the sum of the consistent nodal forces, reaction forces etc.) while fifth and sixth terms are the contact forces. On the other hand, the second left hand side term in the bracket can be seen as a matrix of stiffnesses to constrain the penetrating contacting nodes. $\alpha$ is the penalty parameter and plays an important role in satisfying compatibility. If $\alpha$ is infinite, then the compatibility is satisfied

perfectly, but it will make the system ill-conditioned. If it's too small then the penetrations will be unacceptably large and convergence will be too slow. Therefore $\alpha$ has to be chosen from somewhere in between. Experience shows that $10^7$ to $10^{10}$ is a good range. For each load increment (designated by p) this equation is iterated until the penetrations are negligibly small, i.e.

$$\boldsymbol{\delta}_i - \boldsymbol{L}_i \triangle \boldsymbol{u}_{i+1}^p \approx 0 \tag{2.31}$$

In the Lagrange multipliers method, the minimizing of the potential energy with respect to the displacements and together with the compatibility equation will give the following system of equations:

$$\begin{bmatrix} \boldsymbol{K}_i^p & \boldsymbol{L}_i^T \\ \boldsymbol{L}_i & 0 \end{bmatrix} \begin{Bmatrix} \triangle \boldsymbol{u}_{i+1}^p \\ -\boldsymbol{\lambda}_{i+1} \end{Bmatrix} = \begin{Bmatrix} -\sum_{h=1}^{p-1} \boldsymbol{K}^h \boldsymbol{u}^h - \boldsymbol{K}_i^p \sum_{j=1}^{i+1} \triangle \boldsymbol{u}_j^p \\ 0 \end{Bmatrix} + \begin{Bmatrix} \sum_{h=1}^{p} \boldsymbol{f}^h \\ \boldsymbol{\delta}_i \end{Bmatrix} + \begin{Bmatrix} \sum_{h=1}^{p-1} \boldsymbol{c}^h + \sum_{j=1}^{i} \boldsymbol{c}_j \\ 0 \end{Bmatrix} \tag{2.32}$$

Where again $\boldsymbol{K}$ is the usual stiffness matrix, $\boldsymbol{L}$ is the matrix of contact constraints, $\boldsymbol{f}$ is the incrementally applied external forces, $\boldsymbol{c}$ is the contact forces present at the contactor and target nodes, and $\boldsymbol{\lambda}$ is an array of Lagrange multipliers. Physically, $\boldsymbol{L}_i^T \boldsymbol{\lambda}_i$ represents the nodal contact forces for the removal of overlapping contactor nodes.

The problem with the Lagrange multipliers method is that the new global stiffness matrix obtained after adding the virtual stiffnesses for contact analysis has

zero terms on the leading diagonal making it indefinite. Therefore the frontal solution technique cannot be used and a special solution algorithm is necessary. Since the present code, Frac2D, into which the contact mechanics has been added uses the frontal solution technique, Lagrange multipliers method has been automatically discarded. This method has also another disadvantage, the number of unknowns to be solved during each iteration varies and with a quite complicated model, the number of unknowns will be very high. Note that the Lagrange multipliers are being solved together with the displacements and there are two multipliers for each node in contact, due to the two degrees of freedom associated with each node. Also, penalty function method is far easier to implement into an existing code. For these reasons, penalty function method has been chosen to be used in this work.

### 2.3.5 CONTACT FORCES

The calculation of the contact forces and the decision in the correct contact status at each contactor node is very important to get the correct solution to the problem in hand. There are a few different ways to find the contact forces. For penalty function method, the general way of calculating them is :

$$\sum_{h=1}^{p-1} \boldsymbol{c}^h + \sum_{j=1}^{i+1} \boldsymbol{c}_j^p = \sum_{h=1}^{p-1} \boldsymbol{c}^h + \sum_{j=1}^{i+1} \boldsymbol{c}_j^p + (\alpha \boldsymbol{L}_i^T \boldsymbol{\delta}_i - \alpha \boldsymbol{L}_i^T \boldsymbol{L}_i \triangle \boldsymbol{u}_{i+1}^p) \qquad (2.33)$$

and for Lagrange multipliers method, the common way is:

$$\sum_{h=1}^{p-1} \boldsymbol{c}^h + \sum_{j=1}^{i+1} \boldsymbol{c}_j^p = \sum_{h=1}^{p-1} \boldsymbol{c}^h + \sum_{j=1}^{i+1} \boldsymbol{c}_j^p + \boldsymbol{L}_i^T \boldsymbol{\lambda}_{i+1} \qquad (2.34)$$

The penalty function approach has the disadvantage of a convergence problem that may occur if the Babuska-Brezzi stability condition is violated, which happens when the penalty parameter is too large. Good results will generally be obtained with a small penalty parameter for which the convergence will be slow. The later one does not have the convergence problem mentioned above. In this thesis, the method proposed by Motterhead et al. [32] will be used. The contact forces after each iteration will be calculated by subtracting the applied external forces from the element stress resultants, i.e.:

$$\sum_{h=1}^{p-1} \boldsymbol{c}^h + \sum_{j=1}^{i+1} \boldsymbol{c}_j = -\sum_{h=1}^{p} \boldsymbol{f}^h + \sum_{h=1}^{p-1} \boldsymbol{K}^h \boldsymbol{u}^h + \boldsymbol{K}_i^p \sum_{j=1}^{i+1} \triangle \boldsymbol{u}_j^p \qquad (2.35)$$

This method does not possess the convergence problem of the conventional penalty function approach thus enabling the use of large penalty parameters. As the iterations proceed, it's clear that the increment in the contact forces will vanish together with the incremental displacements $\triangle \boldsymbol{u}_{i+1}^p$.

### 2.3.6 DETERMINATION OF THE CONTACT STATUS

The contact status between a contactor node and a target element can have three values, namely sticking contact, sliding contact and separation. The correct state will be decided upon the contact forces calculated previously. Care should be taken to not decide on the contact status based upon nodal contact forces only, as this can lead to erroneous results.

For this reason, the correct contact state will be decided upon *"element facial forces"* which are basically the sum of the nodal contact forces coming from the

contactor nodes of each element. If the contactor node is shared between two elements then the contact force will be shared too. Lets call $q_n$ and $q_t$ the normal and tangential components of the element facial force respectively. Note that $q_n$ will be positive if it's direction is in the outward normal to the surface. The contact state at a contactor node $m$ is:

- Sticking, if the contactor node has just got into contact with a target element.
- Sliding, if either
  - $|q_t| > \mu |q_n|$ for all elements containing $m$
  - $|q_t| > \mu |q_n|$ for some elements containing $m$ and $q_n > 0$ for others.
- Separation, if all the $q_n$'s of the elements possessing the node $m$ is positive, i.e. $q_n > 0$ for all elements containing $m$.

Once the correct state of contact has been determined for a contactor node, the nodal contact force may be adjusted accordingly. If the status has passed from sticking to sliding, the tangential component $q_t$ will be set to $\mu q_n$, i.e. it will be zero if $\mu$ is zero. If the contact status is separation, then nodal contact force at that node will be set to zero since a separated node cannot have any contact force. As the contact forces at contactor nodes are adjusted, these will be distributed on the target surface using the cubic shape functions

### 2.3.7 SOLVING THE FINITE ELEMENT EQUATIONS

For the reasons mentioned in the section 2.3.4, penalty function method has been chosen to implement the necessary constraints into the finite element equations. Therefore the equations to be solved is Eq. 2.30. In the global sense, these equations are not difficult to form and the solution algorithm will do the rest. On the other hand,

forming and solving the equations on an element-by-element basis is not easy. The equations are symmetric in the global sense, and it has been proved that this holds true in an element-by-element basis as well. This satisfied one of the main conditions in using the frontal solver which is that the element stiffness matrices have to be symmetric. Therefore instead of forming the global matrices $\boldsymbol{K}$ and $\boldsymbol{L}^T\boldsymbol{L}$, one can first find the conventional element stiffness matrix $\boldsymbol{k}$, then a local constraint matrix $\boldsymbol{l}$, which will be formed for each contacting contactor node, will be multiplied by its transpose to get a "contact stiffness" matrix which will be added to $\boldsymbol{k}$, and the result will be sent to the frontal solver. In the local sense $\boldsymbol{l}^t\boldsymbol{l}$ is symmetric therefore the new element stiffness matrix is also symmetric.

During this process, the contactor node should be "tied" to the target element in some way. This is necessary because for each contactor node, the local constraint matrix $\boldsymbol{l}$ contains terms corresponding to the target element nodes as well as the contactor node itself. Note that $\boldsymbol{l}$ is formed for each contactor node - target element pair using the equations 2.23 and 2.24. Virtually connecting the contactor node to the target element is done as follows: the contactor node number is added to the next available spot in the element connectivity and the nnpe (number of nodes per element) of that target element is increased by one. This is done for every contactor node that is in contact with that target element, whether in sticking or sliding contact. An example is shown in Figure. 2.11. Although the contactor node 33 is not physically attached to the target element, it is virtually attached by the means of the connectivity matrix which is shown in Table 2.3 :

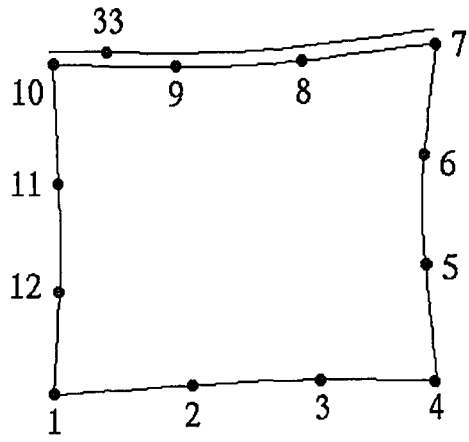


**Figure 2.11:** Simple example of connecting the contactor node to the target element

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 33 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|

**Table 2.3:** Connectivity of a target element in contact

It should be mentioned that this is just a trick for the frontal solver to assemble the stiffness terms and solve the equations correctly. This addition of another node into the connectivity increases the size of the stiffness matrix by 2. If 4 contactor nodes are in contact with a target element then the size of the stiffness matrix becomes 32 by 32, for a 12 noded quadrilateral. This is not a problem for the frontal solver, as long as the connectivity and nnpe value are setup properly. Note that this enlargement of the stiffness matrix to accommodate the contact constraints is only applied to target elements. Since the number of contacting nodes assigned to an element is not fixed, the size of the stiffness matrix varies during the iterations, which is not the case for linear elastic, plastic or large deformation problems.

## 2.4 IMPLEMENTATION INTO THE CRACK PROBLEM

In the enriched formulation, it is known that the stress intensity factors are calculated at the same time as the displacements. In other words, $K_I$ and $K_{II}$ are interpreted as unknowns of the finite element equations. For this purpose a virtual node number is assigned to each crack, the number is found by increasing the total number of nodes by one. For example if we have 300 nodes and a crack in the model, then the crack will take the virtual node number of 301, it will appear in the connectivity matrices as 301. Consequently $K_I$ and $K_{II}$ will be the x and y component equivalents of virtual node 301. In this way, the crack is virtually attached to that element. The same algorithm is followed when contact is also present in the problem. The total number of nodes will not change since the penalty function method does not add any unknowns to be solved into the finite element equations. The only difference will be the nnpe of some elements. Let's say a 12 noded quadrilateral element is at the same time a target and a crack tip element (whether enriched or transition). In the local sense, the $13^{th}$ position in the connectivity will always be assigned to the crack tip virtual node. The contactor nodes in contact will be added to the connectivity starting from the $14^{th}$ position and so on. If the element is not enriched, the contactor nodes will be put starting from the $13^{th}$ position. To summarize, the contact problem will not add any additional unknowns into the crack problem, it will only effect the size of the local unknowns associated with a target element and this is for the purpose of implementing the necessary contact constraints into the equations. As an example, let's take again the example in Figure 2.11 and let's make the target element an enriched element. Let's say there are 200 nodes in the model, therefore the crack will take the virtual node number of 201. The connectivity of the element will become as shown in Table 2.4:

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 201 | 33 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|

**Table 2.4:** Connectivity of an enriched target element in contact

The implementation of the crack problem into the finite element formulation of the contact problem presented in section 2.3 is quite straight forward. When a crack is present in the model, some special element stiffness and load terms will be calculated, as done in linear elastic crack problem case, and they will be added to the previously formed modified element stiffness matrices and load vectors. The connectivity and nnpe value of the elements being already adjusted if necessary, the frontal solver is called to solve for the unknown quantities. The contact problem is truly nonlinear, the solution is reached incrementally. Therefore what is solved in Eq. 2.30 is not the displacements and the stress intensity factors but the incremental displacements and incremental stress intensity factors. These will be added up to obtain the final values for the displacements and the mode I and mode II stress intensity factors.

Care should be taken in forming the necessary constraint matrices $\boldsymbol{l}$, the formulation presented in 2.3 applies for isoparametric elements, i.e. the displacements can the interpolated using shape functions only. But in enriched formulation, for the displacements within an enriched element, other than the usual shape function terms there are also some asymptotic displacement functions, as shown previously in Eq. 2.9 and 2.10. Therefore when a target element in contact is also an enriched element, like in the case of crack surfaces in contact, the formulation should be changed as follows:

$$u(s,t) = \sum_{j=1}^{nnpe} N_j(s,t)u_j + Z_0(s,t)f_1^*(s,t)K_I + Z_0(s,t)g_1^*(s,t)K_{II} \qquad (2.9)$$

$$v(s,t) = \sum_{j=1}^{nnpe} N_j(s,t)v_j + Z_0(s,t)f_2^*(s,t)K_I + Z_0(s,t)g_2^*(s,t)K_{II} \qquad (2.10)$$

Using Eq. 2.9 and 2.10, Eq 2.21 becomes:

$$\delta_x = u_c - \sum_i N_i(s,t)u_i - Z_0(s,t)f_1^*(s,t)K_I - Z_0(s,t)g_1^*(s,t)K_{II} \qquad (2.36)$$
$$\delta_y = v_c - \sum_i N_i(s,t)v_i - Z_0(s,t)f_2^*(s,t)K_I - Z_0(s,t)g_2^*(s,t)K_{II}$$

Note that the isoparametric coordinates s and t corresponds to the corrected coordinates found during contact searching.

Eq 2.36 can be put into the form

$$\delta_i - \begin{Bmatrix} u_c \\ v_c \end{Bmatrix}_{i+1} + \begin{bmatrix} N_1 & 0 & N_2 & 0 & \dots & \dots & N_{12} & 0 \\ 0 & N_1 & 0 & N_2 & \dots & \dots & 0 & N_{12} \end{bmatrix} \begin{Bmatrix} u_T \\ v_T \end{Bmatrix}_{i+1}$$
$$+ \begin{bmatrix} Z_0(s,t)f_1^*(s,t) & Z_0(s,t)g_1^*(s,t) \\ Z_0(s,t)f_2^*(s,t) & Z_0(s,t)g_2^*(s,t) \end{bmatrix} \begin{Bmatrix} K_I \\ K_{II} \end{Bmatrix} = 0 \qquad (2.37)$$

or

$$\delta_i - \begin{Bmatrix} u_c \\ v_c \end{Bmatrix}_{i+1} + [I \otimes \mathrm{Ns}]_i \begin{Bmatrix} u_T \\ v_T \end{Bmatrix}_{i+1}$$
$$+ \begin{bmatrix} Z_0(s,t)f_1^*(s,t) & Z_0(s,t)g_1^*(s,t) \\ Z_0(s,t)f_2^*(s,t) & Z_0(s,t)g_2^*(s,t) \end{bmatrix} \begin{Bmatrix} K_I \\ K_{II} \end{Bmatrix} = 0 \qquad (2.38)$$

where $\boldsymbol{I}$ is the identity matrix, $\otimes$ is the Kronecker product.

For sliding contact, only the tangential displacement of the contactor node should be constrained, therefore the constraint equation will be

$$\mathbf{n}^T\boldsymbol{\delta}_i - \mathbf{n}^T\begin{Bmatrix} u_c \\ v_c \end{Bmatrix}_{i+1} + \mathbf{n}^T[\boldsymbol{I}\otimes\mathrm{Ns}]_i\begin{Bmatrix} u_T \\ v_T \end{Bmatrix}_{i+1} + \mathbf{n}^T\begin{bmatrix} Z_0(s,t)f_1^*(s,t) & Z_0(s,t)g_1^*(s,t) \\ Z_0(s,t)f_2^*(s,t) & Z_0(s,t)g_2^*(s,t) \end{bmatrix}\begin{Bmatrix} K_I \\ K_{II} \end{Bmatrix} = 0 \qquad (2.39)$$

where $\mathbf{n}$ is the normal vector evaluated at the corrected isoparametric coordinates. Combining both equations, one can formulate both sticking and sliding contact constraint equations in the following form:

$$\boldsymbol{\delta}_i - \boldsymbol{L}_i\triangle\boldsymbol{u}_{i+1} \qquad (2.40)$$

The rest of the formulation is exactly the same as described in section 2.3.

# CHAPTER 3

# NUMERICAL EXAMPLES

## 3.1 INTRODUCTORY EXAMPLE

In this section, as a simple introductory example, a block in contact with another block will be examined. The dimensions of the blocks are each 20 by 5 inches. The model can be seen in Figure 3.1. The modulus of the blocks are 30e6 psi and the Poisson's ratios are 0.30. The problem is basically two similar blocks in contact with each other. The bottom block is constrained to move in the horizontal direction. Pressure loads act on the top and side surfaces of the top block, each of magnitude 1000 psi/in, the directions being shown on the model. Plane strain is assumed. The problem will be examined for both sticking and frictionless sliding contact conditions and the deformed shapes will also be shown. Note that sticking condition can be easily imposed on every node in a model by simply assigning a coefficient of friction so high that no contactor node can pass from sticking contact to sliding contact.

Without any contact algorithm, the problem would be indefinite because there's no constraint on the top block except the contact constraints. Therefore without these constraints, the stiffness matrix would become singular, causing infinite displacements. The contact constraints prevent the stiffnesses from becoming singular and prevents the two blocks from passing through each other.

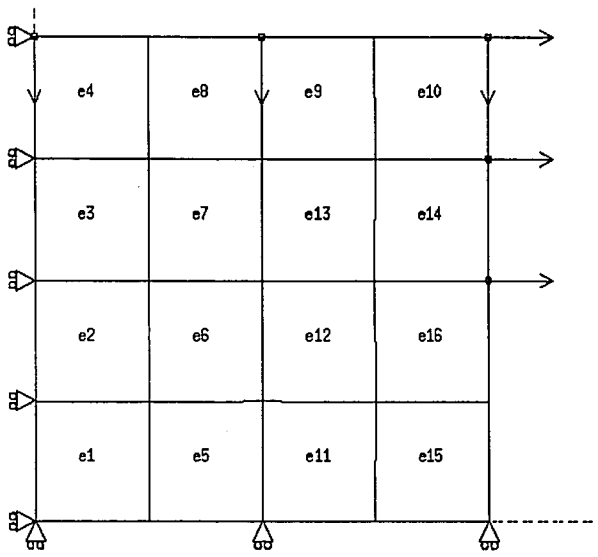


**Figure 3.1:** Model for two blocks in sticking/sliding contact

A good way of checking the sticking contact problem between the two blocks is the "whole block" or "tied block" problem. Since the contactor nodes are prevented from sliding on the bottom block surface, the contactor nodes will *stick* to the first point of contact on the target surface. The model has been prepared in order to get node-to-node contact, therefore the contactor nodes and the corresponding target nodes will have exactly the same displacements. It is the same problem as the one with the contactor nodes and the corresponding target nodes being tied, or even as the problem where there's a unique block of width 20" and of height 10" and having the nodes at exactly the same locations as the contact problem.

In Figure 3.2, the stresses coming from the contactor and target surfaces are shown. They are practically equal on the left side of the graph but are quite different on the other side. As they are averaged, the stresses are in good agreement with the known solution (or the whole body solution), as shown in Figure 3.3. This difference between the stresses coming from the two surfaces are believed to be caused by the

stress smoothing algorithm that was already present in the code. This issue is explained in detail in Appendix B.

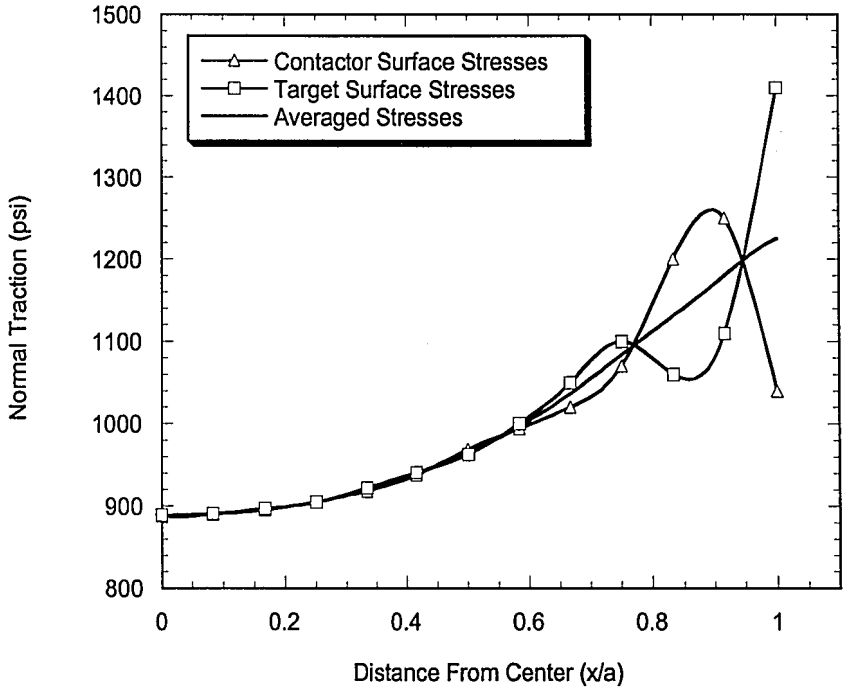


**Figure 3.2:** Stresses for the introductory example (a is the half width of the blocks)

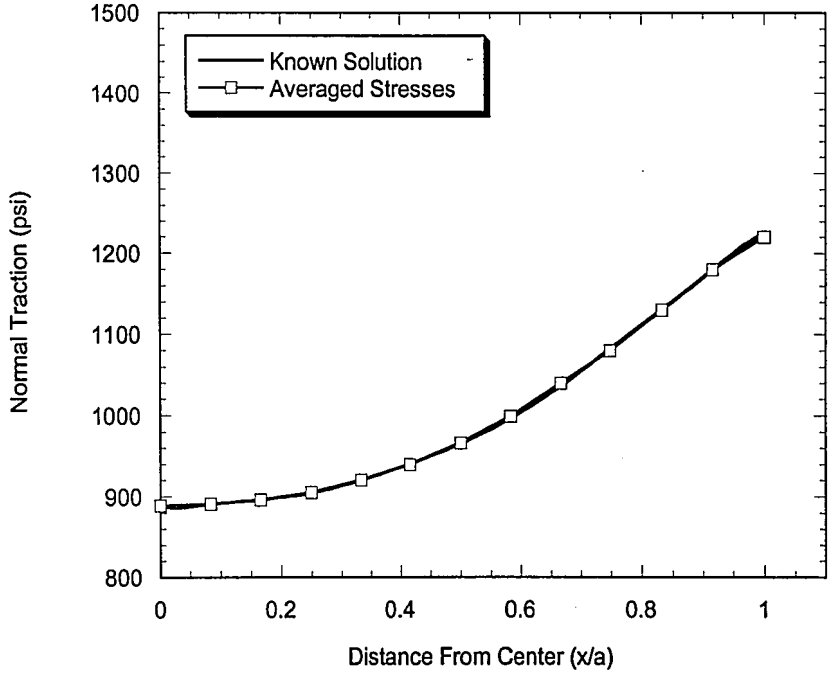


**Figure 3.3:** Comparison between averaged stresses and the known solution

In Figures 3.4 and 3.5, the deformed shapes of the sticking and frictionless sliding problems are plotted respectively with a scale factor of 500, i.e the displacements are multiplied by 500 in order to visualize the deformations.

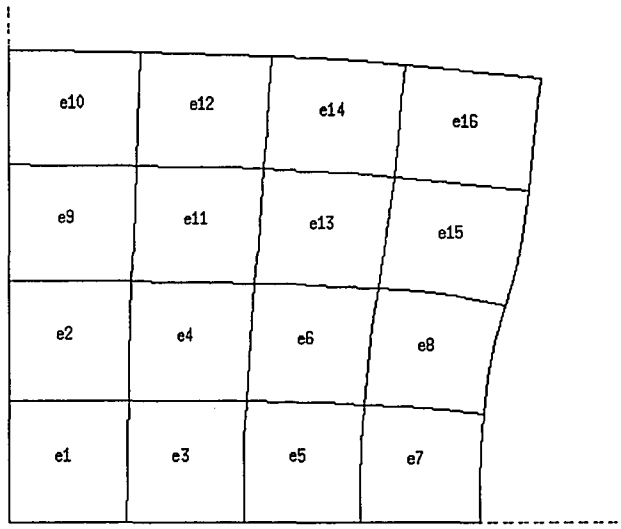


**Figure 3.4:** Deformed shape for the sticking block problem

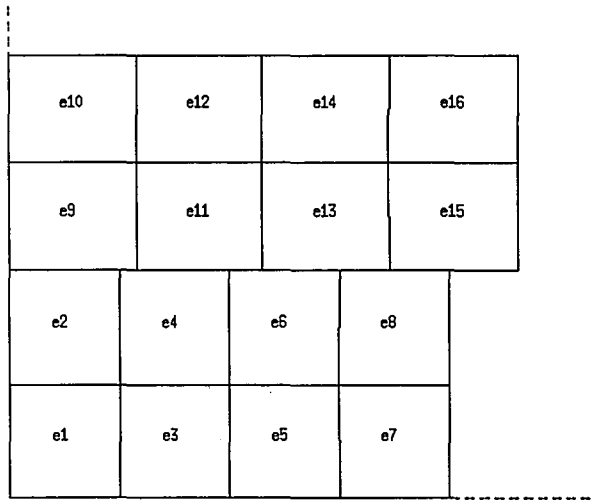


**Figure 3.5:** Deformed shape for the frictionless sliding block problem

It can be seen that, for the sticking problem no sliding is allowed therefore the right end of the contact surface has to bend down with respect to the left end. But in frictionless sliding problem, the contactor surface is free to slide on the target surface,

there are no forces in the x direction acting on the contact surfaces and consequently both the left and right sides of the contact surface deforms almost equally.

## 3.2 FLAT PUNCH ON ELASTIC FOUNDATION

In this section, a flat punch on elastic foundation problem will be modeled. Call the height and width of the elastic foundation $H$ and $2W$ and the height and width of the punch $h$ and $2w$ respectively. The dimension ratios considered are

$$\frac{h}{w} = \frac{8}{5} \qquad \frac{H}{W} = \frac{5}{4} \qquad \frac{w}{W} = \frac{5}{8}$$

as in the example in [32]. A uniform distributed load of 1 $N/mm^{-1}$ is applied on the top surface of the flat punch. Both are made of the same material, with a modulus of elasticity of 3e6 psi and Poisson's ratio of 0.35. Plane strain is assumed.

There are two models used for this problem, Figure 3.6 shows the simple and refined models respectively. The idea in using a refined model is to properly catch the singularity at the corner of the flat punch that is also predicted by linear theory of elasticity. In fact the simple model failed to catch that singularity, which will be shown in the following plots. Figures 3.7 and 3.8 shows respectively a closer look the contact surfaces for the simple and refined models.

Although there's no analytic solution for this problem, there are some finite element results to compare the results. One of them is in the reference [32], where the problem was solved for various material combinations and for some coefficients of friction different than zero. Here only sliding contact, i.e. contact with zero coefficient of friction, and only one material combination (both have equal material properties) will be investigated for simplicity purposes.

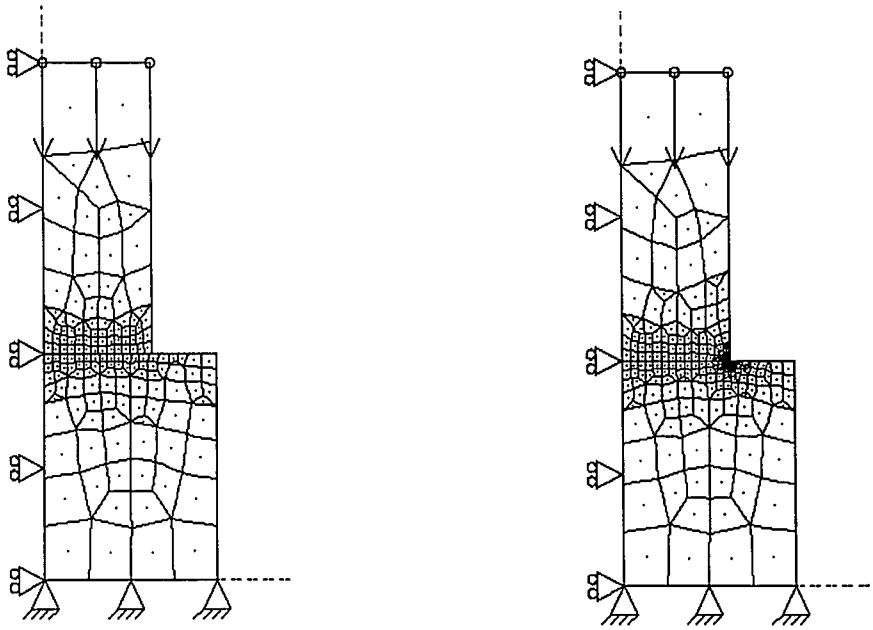

**Figure 3.6:** Simple and refined models for flat punch on elastic foundation problem

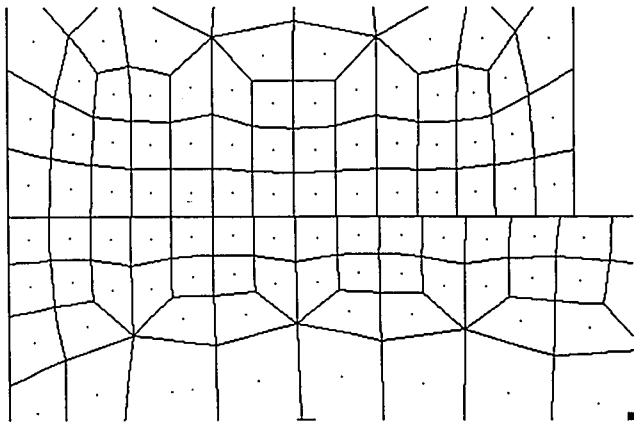

**Figure 3.7:** Contact zone mesh for the simple model

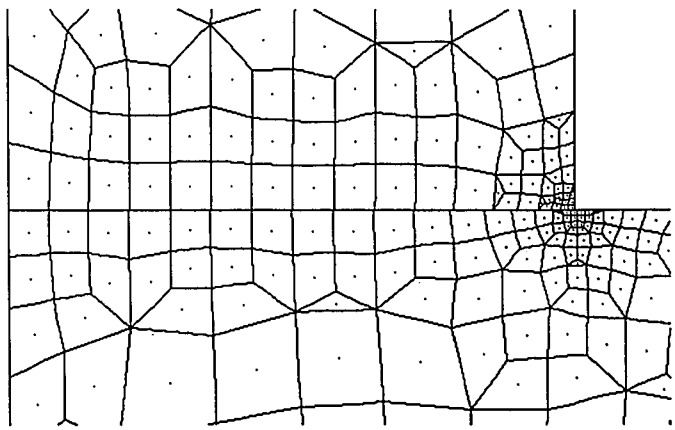

**Figure 3.8:** Contact zone mesh for the refined model

The stresses for both the simple and refined models are shown in Figure 3.9 and Figure 3.10. As it can be seen, the simple model is not satisfactory in catching the singularity at the corner of the punch. Note that again the stresses coming from the target and contactor surfaces are different from each other on the right hand side of the plot. This is again due to the stress smoothing, which is explained in Appendix B. Nevertheless, the average of the stresses coming from both surfaces are in agreement with the example solved in [32].

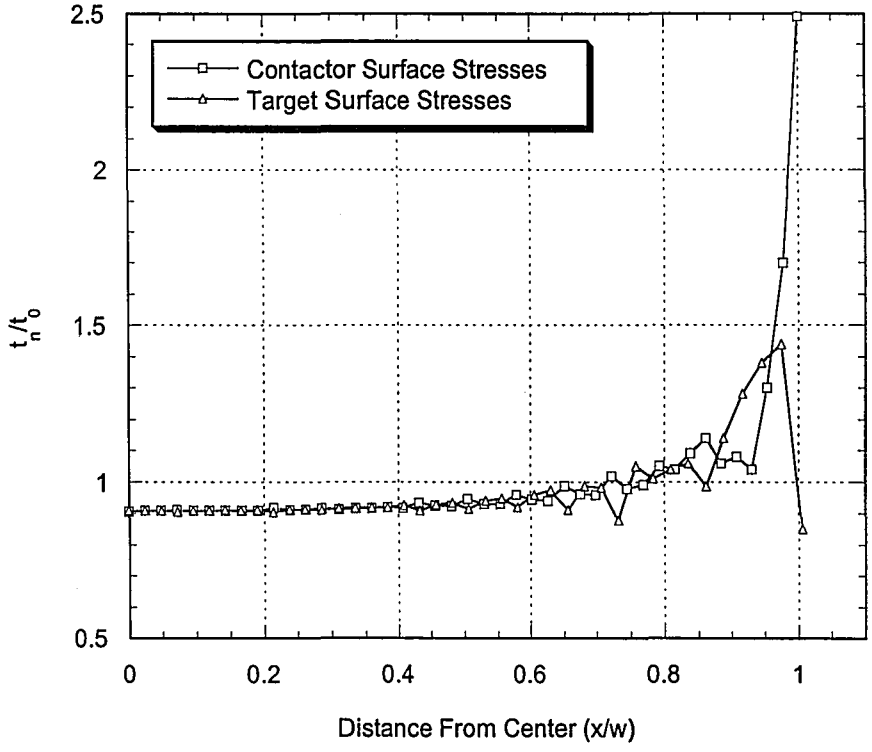


**Figure 3.9:** Stresses for the simple model of flat punch on elastic foundation problem ($t_0$ is the applied pressure, $t_n$ is the normal traction and w the width of the punch)

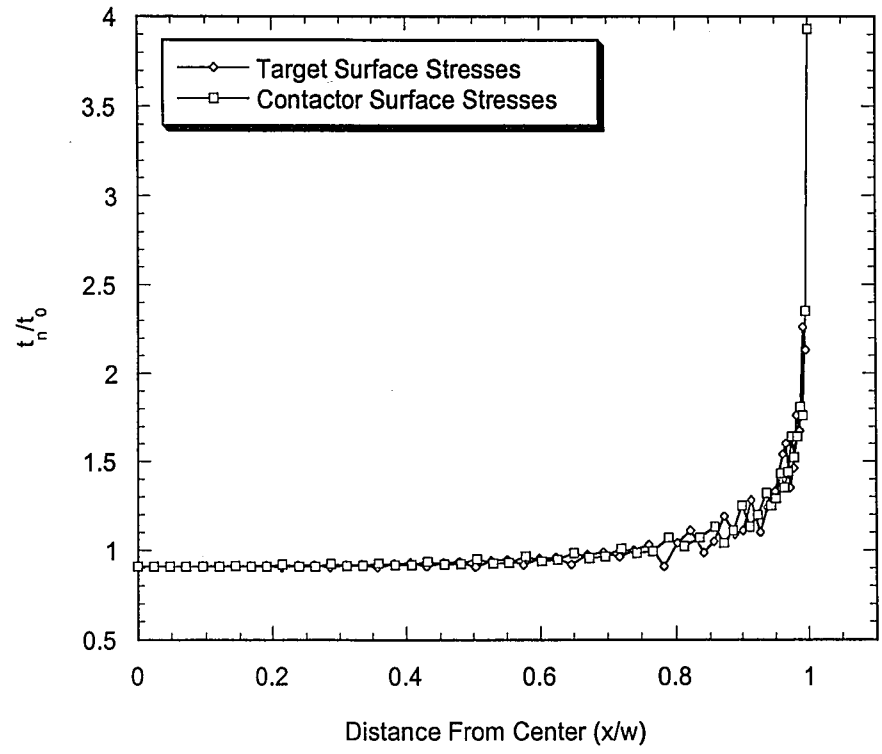


**Figure 3.10:** Stresses for the refined model of flat punch on elastic foundation problem ($t_0$ is the applied pressure, $t_n$ is the normal traction and w the width of the punch)

## 3.3 CYLINDRICAL PUNCH ON ELASTIC FOUNDATION

The cylindrical punch on elastic foundation problem is a quite interesting and popular one. It's also quite difficult to model numerically, in the sense that every kind of contact conditions occur at the same time in the same model, i.e. sticking, sliding and separation. In the previous examples, only sticking and sliding was present. Separation complicates the problem by making it difficult to keep track of the penetrating and seperating nodes.

Analytically this problem has been first analyzed by Hertz [33]. Some details of his study can be found in Appendix A. His formulation applies for spheres and

cylinders in contact with each other as well as cylinders and spheres in contact with a half medium, just by setting the radius of one of the cylinder as infinity.

Since it is not possible to model a half space, a large enough foundation has been taken in order to be sure to create an equivalent model. The dimensions of the block is 1000 by 1000 mm while the radius of the cylinder is 50 mm. The material properties of both the cylinder and the foundation are 10.8 MPa for the elastic modulus and 0.35 for the Poisson's ratio. The concentrated load applied on top of the cylinder is 3916N. Plane strain is assumed. The model used is shown in Figure 3.11, and Figure 3.12 shows a closer look at the contact zone between the cylinder and the half space.

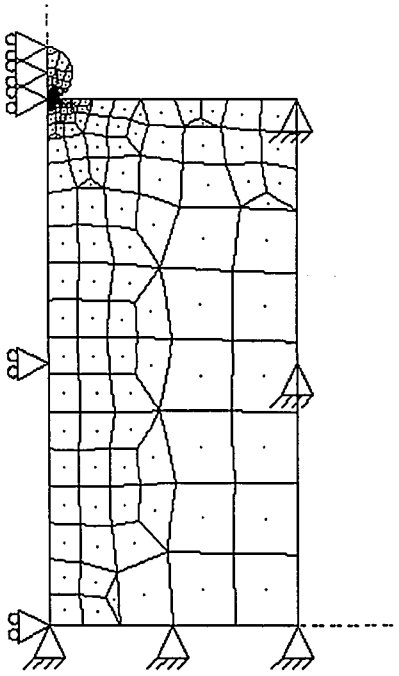

**Figure 3.11:** Cylindrical punch on elastic foundation model

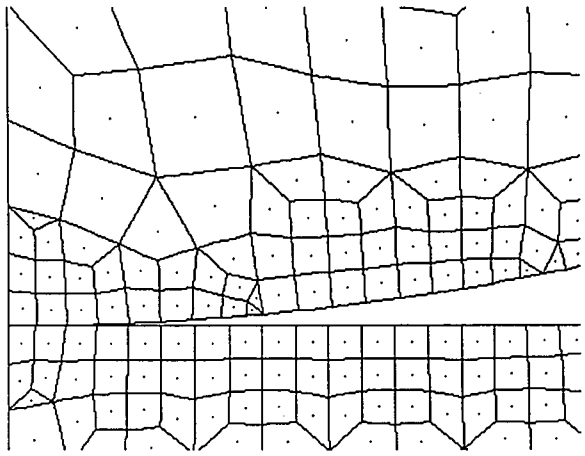

**Figure 3.12:** Close look to the contact zone in cylindrical punch problem

In Figure 3.13, the stresses on the contactor surface, target surface and the averaged stresses are shown. The averaged stresses are found by passing a $10^{th}$ degree regression curve using the stresses on both the target and the contactor surfaces. It can be seen that the stresses at the corner nodes, on both surfaces, are a little bit offset from other values. This is believed to be caused by the stress smoothing algorithm already being used in the code that the contact algorithm has been implemented, namely Frac2D. The stress smoothing algorithm is presented in detail in Appendix B.

The comparaison between the averaged stresses and the well-known Hertzian solution is in perfect agreement and is shown in Figure 3.14 . Despite the erroneous behavior of the stresses coming from either surfaces, the averaged stresses are very accurate.

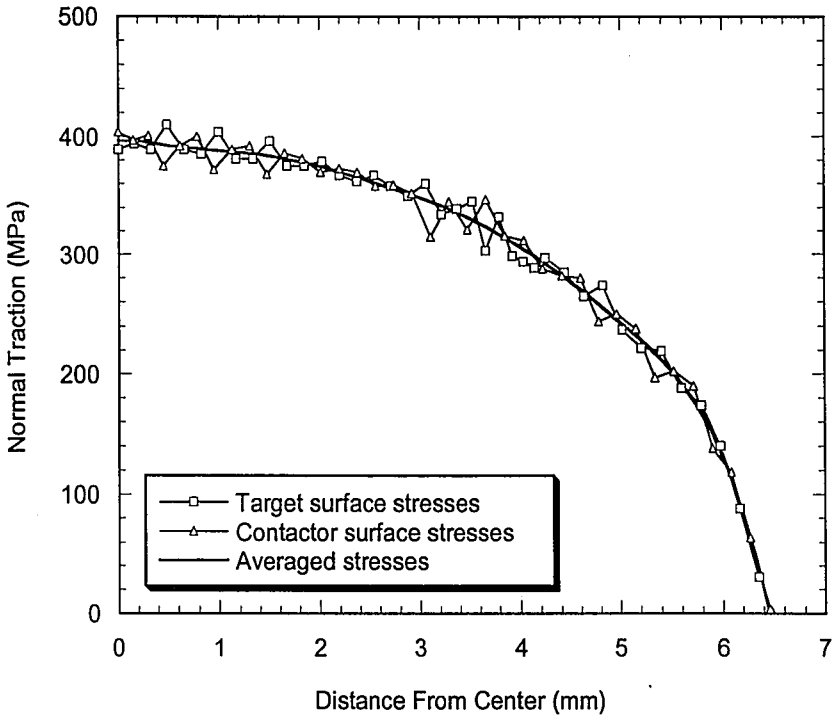


**Figure 3.13:** Stresses as a function of the distance from the center

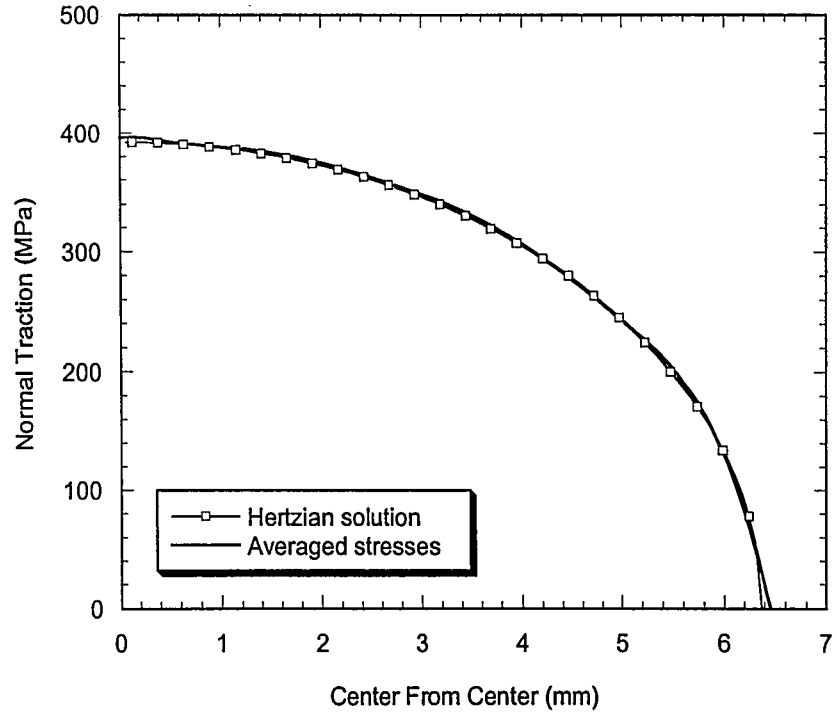


**Figure 3.14:** Comparison of the stresses found with the Hertzian solution

## 3.4 SEMICONDUCTOR PACKAGE APPLICATIONS

In this section, some interface crack contact problems in semiconductor packages will be modeled and the results will be presented. Edge cracked and center cracked flip chip packages with various boundary conditions will be examined. The boundary conditions applied are unconstrained substrate, substrate constrained in the vertical direction only, substrate constrained in both directions and substrate constrained at the end. The resulting mode I and mode II stress intensity factors, total strain energy release rates and phase angles will be plotted as a function of crack length. Also, for comparative purposes, the results with no contact constraints will be shown and discussed. Plane strain is assumed for every case.

Material properties for the constituent materials are as shown in Table 3.1 :

| material | $E_{xx}$, $E_{yy}$, $E_{zz}$ (psi) | $\nu_{12}$, $\nu_{23}$, $\nu_{31}$ | $G_{12}$ (psi) |
|---|---|---|---|
| silicon die | 3.407e6, 3.407e6, 3.407e6 | 0.33, 0.33, 0.33 | 1.281e6 |
| underfill | 1.056e6 | 0.301 | - |
| polymeric substrate | 1,884e7 | 0.279 | - |

| material | $\alpha_{xx}$, $\alpha_{yy}$, $\alpha_{zz}$ |
|---|---|
| silicon die | 1.5e-6, 5.7e-6, 1.5e-6 |
| underfill | 2.6e-5 |
| polymeric substrate | 3.3e-6 |

**Table 3.1:** Material properties of the constituent materials in a flip chip package

The dimensions of the flip chip package are shown in Figure 3.15

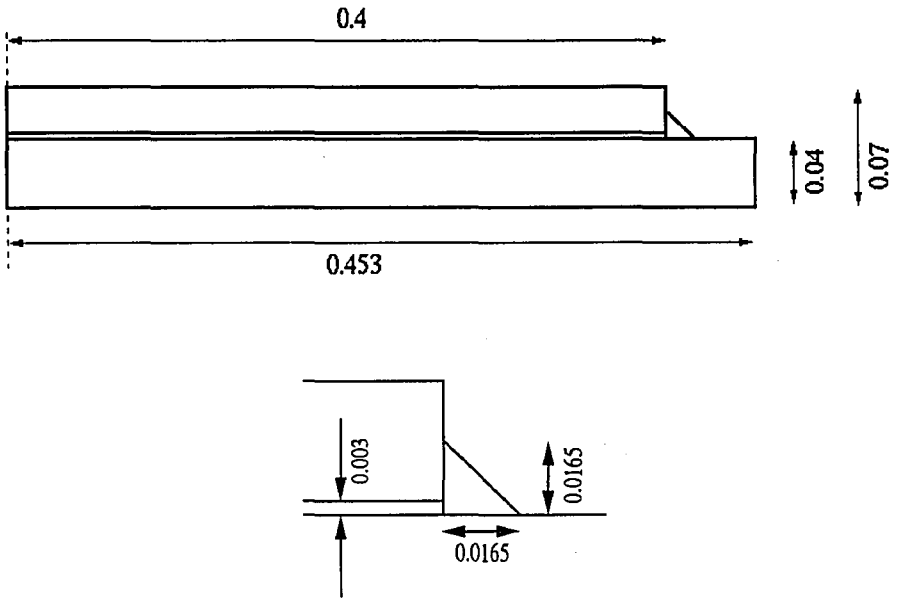


**Figure 3.15:** Dimensions of the flip chip package

## 3.4.1 EDGE-CRACKED FLIP-CHIP PACKAGE

This is the most common and probably the most severely cracked flip-chip package. It is well known that the corner between the underfill and the substrate has a singularity. Even though it is not as strong as a crack tip singularity, the strength may be sufficient to initiate a crack at that location. Four different boundary conditions will be applied to the substrate and relevant fracture mechanics quantities such as mode I and mode II stress intensity factors, total strain energy release rates and phase angles are found and plotted as a function of crack lengths ranging from 0.0045 in to 0.0315 in..

### 3.4.1.1 UNCONSTRAINED PACKAGE

Let's say the package is tested by the manufacturer for reliability by putting them on a surface and applying heat and/or moisture, like in an environmental

chamber. This test can be modeled by not constraining the bottom surface of the substrate, so the ends of the substrate are free to bend upwards of downwards depending on the material combinations and temperature loading.

The model used is shown in Fig 3.16.

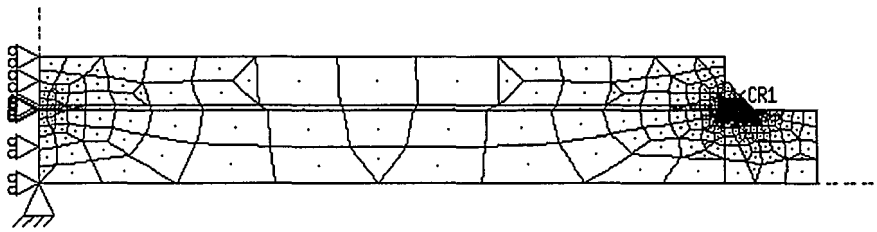


**Figure 3.16:** Unconstrained edge cracked flip-chip model

Two different temperature loads have been applied, $-10^0$C and $+10^0$C for the reason that in both of the temperature differences contact occurs between the crack surfaces for certain crack lengths. There's a rule of thumb stating that if the crack surfaces opens with a temperature difference, it should eventually close when the negative temperature difference is applied. This case is obviously an exception to that rule since in both cases contact occurs, not for every crack lengths tough but just for certain ones. For this reason the results for each case will be shown.

For $\Delta T=10^0$C, for certain crack lengths, a phenomenon that can be called "double contact zone" is observed. For this problem, the contact zone shape, number and position changes with crack length. Figure 3.17 shows the shape changes in the contact zone as the crack length is increased.

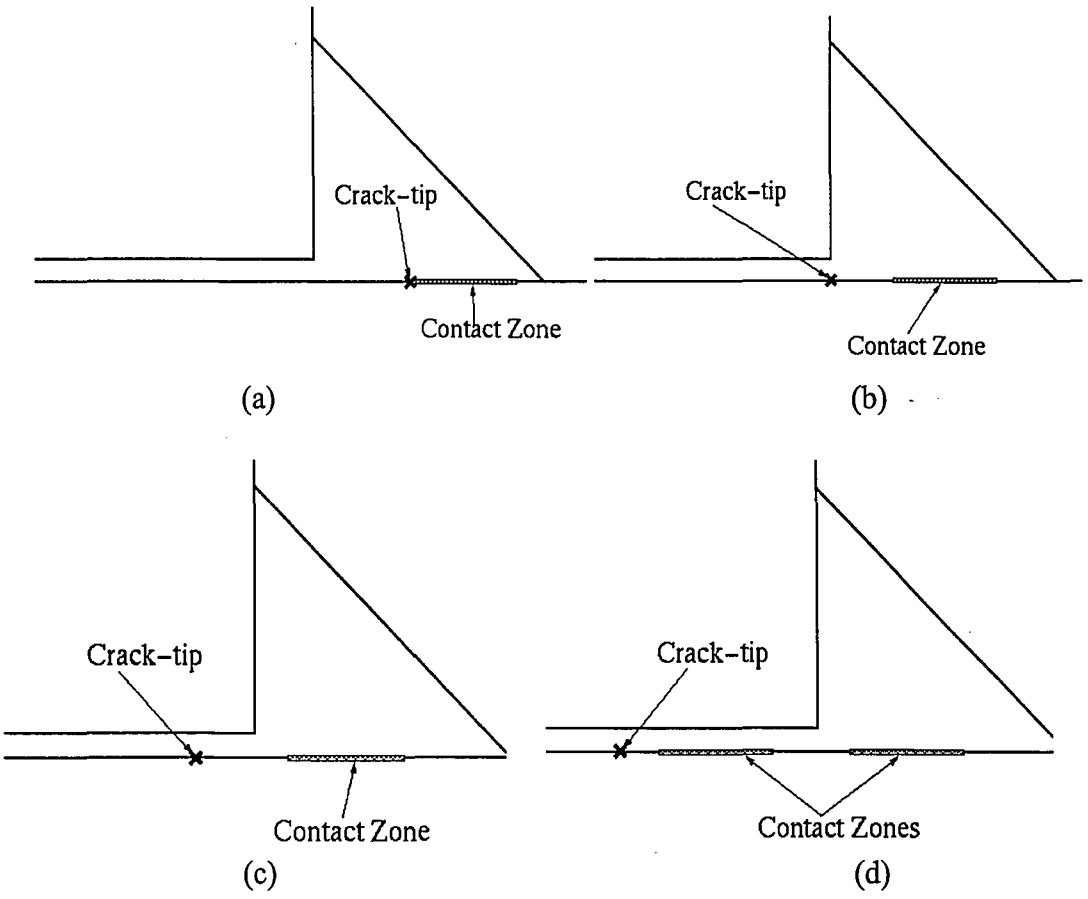


**Figure 3.17:** Shape changes in the contact zone as the crack length is increased in a flip chip package with unconstrained substrate, $\Delta T=10^0C$, a) Crack is between the silicon die corner and edge of the underfill. b) Crack is just underneath the silicon die corner c) Crack has just passed the corner d) Crack is relatively far from corner

When the crack is between the silicon die corner and the edge of the underfill layer (figure 3a) the contact zone starts right at the crack tip and ends just before the edge. Then crack is just underneath the silicon die corner (figure 3b), the contact zone does not start at the crack-tip anymore but just after it and extends some distance without reaching the edge. After having passed the corner, at about a crack length of 0.0195 in, the contact zone is still just ahead of the crack. When the crack is relatively far away from the corner, now a double contact zone is seen. There for a crack length between 0.0195 and 0.0225 in, a transition from one contact zone to double contact

zones occurs. One of double zones is after the corner, the other is before the corner. The size of the zone before the crack tip, i.e. lying in the triangular region, is pretty much constant while the size of the zone after the crack tip almost doubles in size.

In the following figures, between the region between the points A and B corresponds to the case in Figure 3.17a, between B and C corresponds to Figures 3.17b and 3.17c (the vertical line indicating the position of the silicon corner), between C and D corresponds to Figure 3.17d, i.e. the double contact zone case.

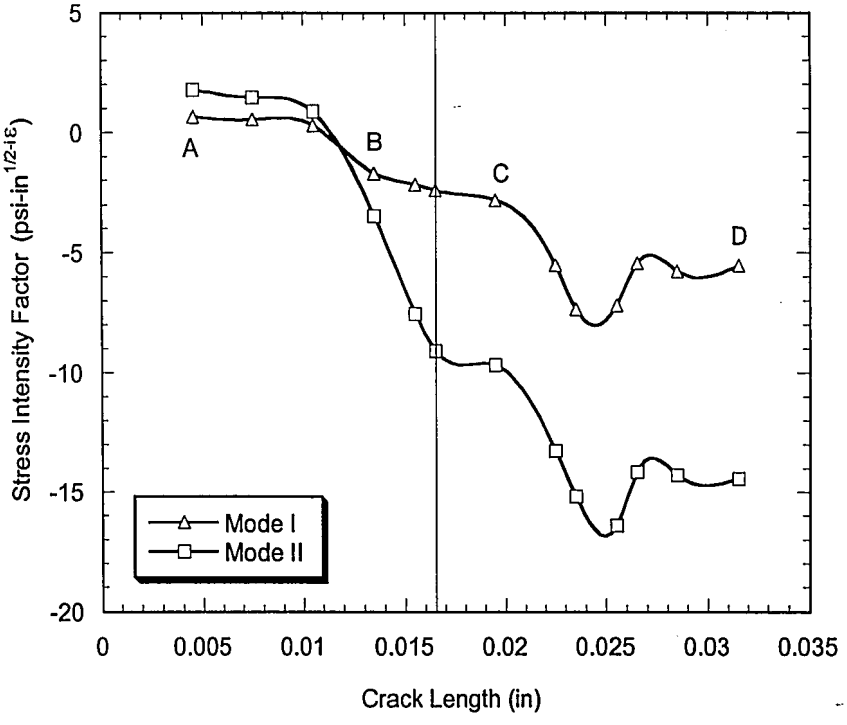


**Figure 3.18:** Stress intensity factors as a function of crack length for the edge cracked package under unconstrained substrate boundary condition, $\Delta T=10^0C$.

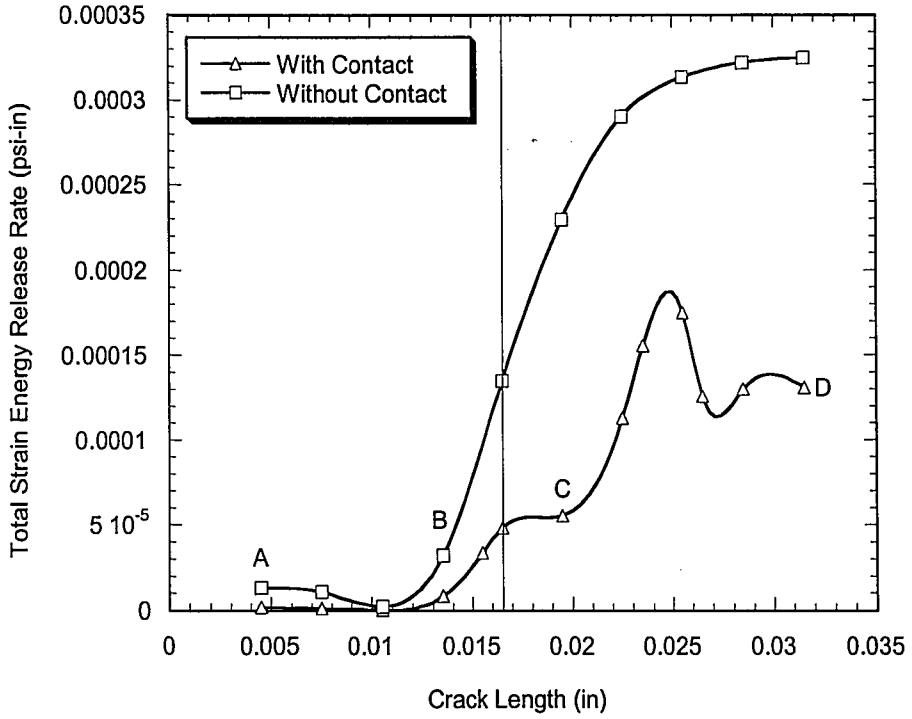


**Figure 3.19:** Total strain energy release rates as a function of crack length for the edge cracked package under unconstrained substrate boundary condition, $\Delta T=10^0C$.

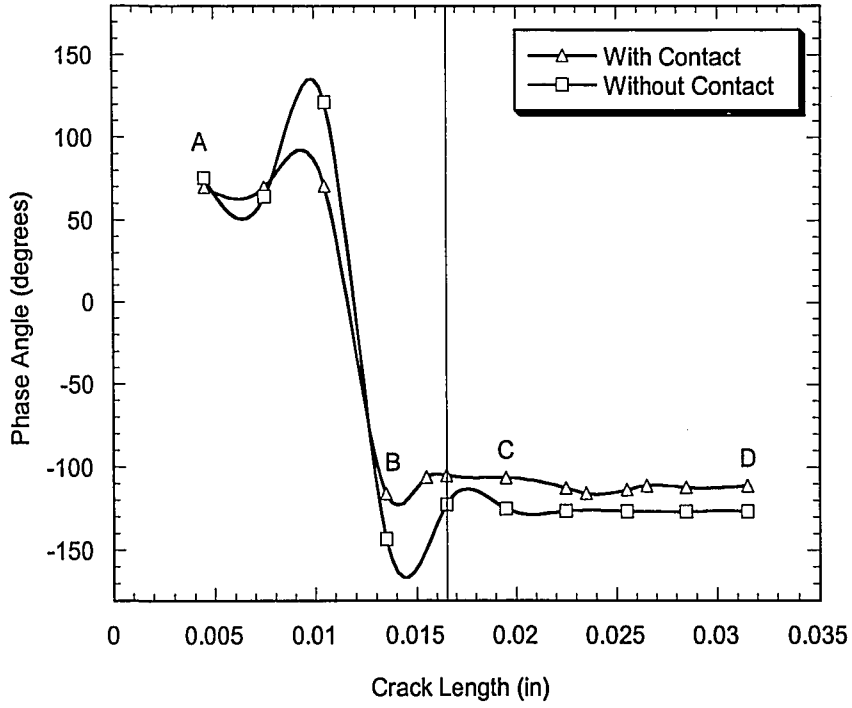


**Figure 3.20:** Phase angles as a function of crack length for the edge cracked package under unconstrained substrate boundary condition, $\Delta T=10^0C$.

Figure 3.18 shows the mode I and mode II stress intensity factors as a function of crack length for $\Delta T=10^0C$. As it can be seen, the behavior is extremely complex, due to the change in the shape of the contact zone(s), unconstrained boundary condition and the effect of the silicon corner chip. The only observation to be made is the increase in magnitude of both mode I and mode II stress intensity factors as the crack length increases.

Figure 3.19 shows the total strain energy release rates, obtained using contact analysis and without contact analysis, as a function of crack length, for $\Delta T=10^0C$. The values obtained without using contact analysis are quite smooth and increasing with increasing crack length. The values from contact analysis, on-the-other-hand, shows some ups and downs. This is mainly caused by the contact zone shape change previously presented. Between A and B (for crack lengths ranging from 0.0045 to 0.0135) the crack surfaces are closed starting right from the crack tip which makes the strain energy release rates quite low. The first jump occurs on the graph between B and C due to the fact that the contact zone doesn't begin right from the crack-tip. An even bigger jump occurs when the single contact zone becomes a double contact zone, for the region from C to D. One can observe that the total strain energy release rates are overestimated by a factor of about two when contact analysis is not performed.

Figure 3.20 shows the phase angles, obtained using contact analysis and without contact analysis, as a function of crack length, for $\Delta T=10^0C$. Both are quite oscillatory when the crack is before the silicon corner (in the triangular section), then gets stabilized after the corner. The oscillation is due to the sign changes in the mode

I and mode II stress intensity factors, as it can be seen in Figure 3.18. Both graphs show the same overall behavior.

For $\Delta T=-10^0C$, no double contact zones are observed, in fact for most of the crack lengths even contact does not occur.

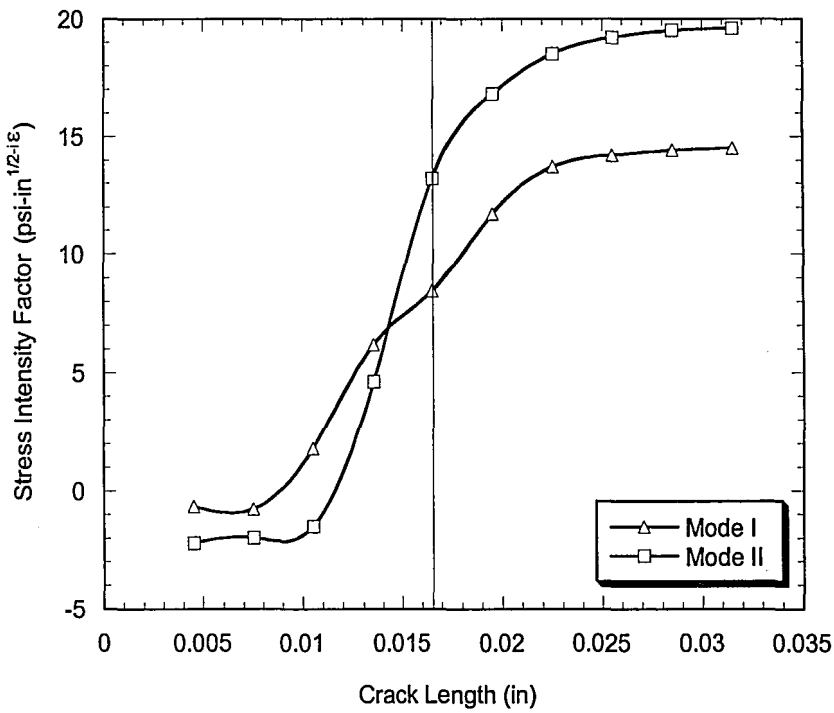


**Figure 3.21:** Stress intensity factors as a function of crack length for the edge cracked package under unconstrained substrate boundary condition, $\Delta T=-10^0C$.

Figure 3.21 shows the mode I and mode II stress intensity factors as a function of crack length, for $\Delta T=-10^0C$. Note that there is a small range of crack lengths between 0.0105 and 0.014 in where the mode I effect is dominant. For all other crack lengths the mode II effect is dominant.

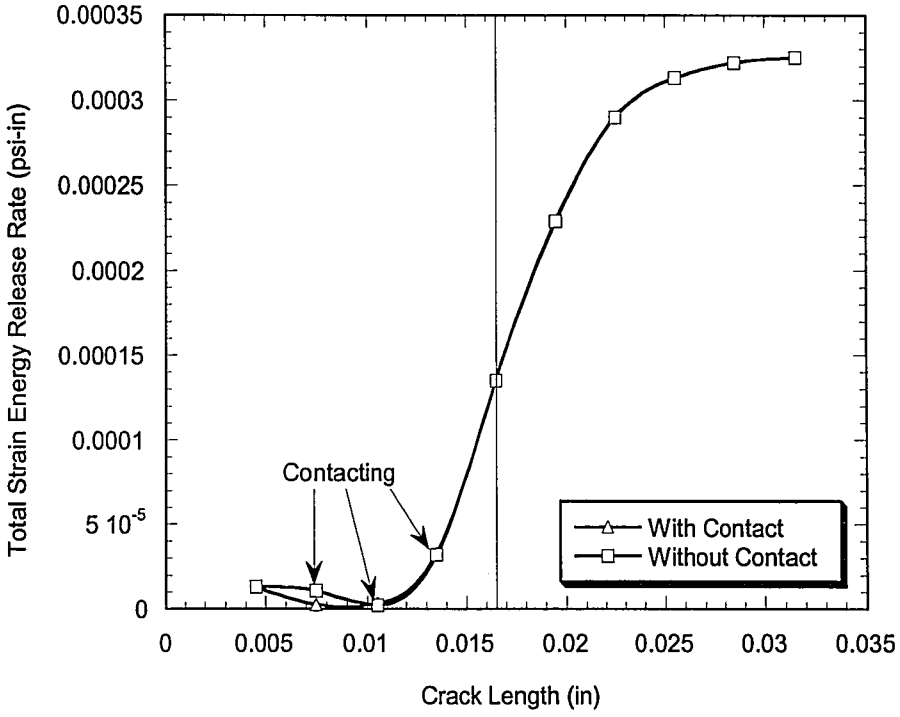


**Figure 3.22:** Total strain energy release rates as a function of crack length for the edge cracked package under unconstrained substrate boundary condition, $\Delta T=-10^0C$

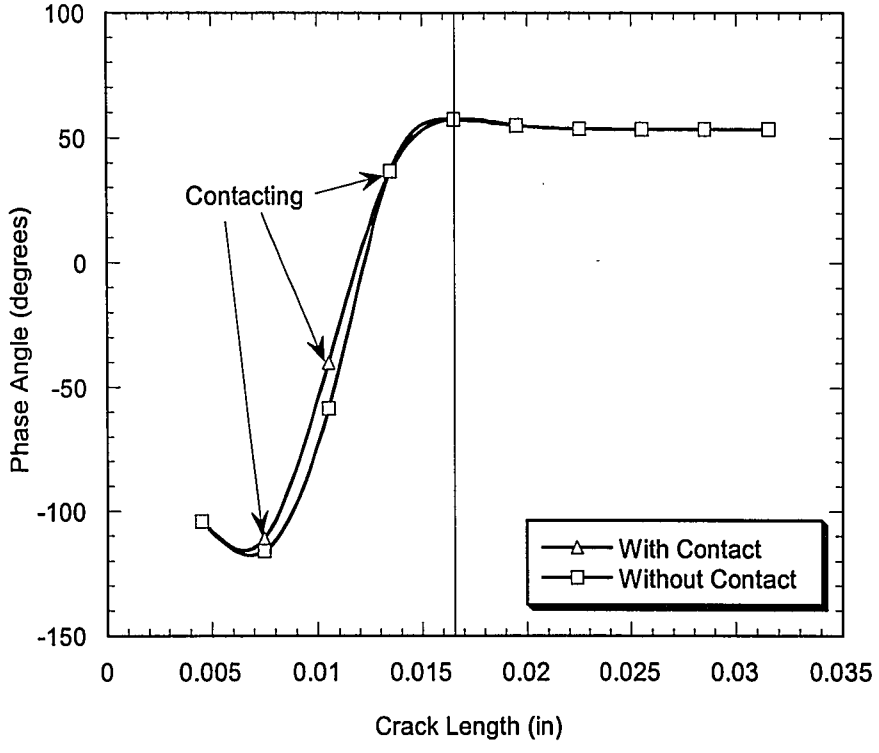


**Figure 3.23:** Phase angles as a function of crack length for the edge cracked package under unconstrained substrate boundary condition, $\Delta T=-10^0C$

Figure 3.22 shows the total strain energy release rates, obtained using contact analysis and without contact analysis, as a function of crack length, for $\Delta T=-10^0C$. The curves are coincident except for some crack lengths, shown on the graph, for which contact occurs. In the overall, total strain energy release rates increase as the crack length is increased.

Figure 3.23 shows the phase angles, obtained using contact analysis and without contact analysis, as a function of crack length, for $\Delta T=-10^0C$. Again the curves are coincident except for the crack lengths at which there is contact. The abrupt change in the sign of the phase angles is due to the change in the signs of the stress intensity factors. Having passed the silicon corner, phase angles get constant.

### 3.4.1.2 SUBSTRATE CONSTRAINED IN VERTICAL DIRECTION

If the substrate is attached or bonded to a board, then the boundary conditions on the bottom surface of the substrate are quite complex. It can be modeled either by constraining the substrate to move in the vertical direction only or by constraining it to move both in vertical and horizontal directions. The real case is neither of them but inevitably falls between the two cases.

This case is run only for a temperature difference of $\Delta T=-10^0C$ since no contact is occurring for any crack length when $\Delta T=10^0C$ is applied. The model used for modeling this problem is shown in Figure 3.24

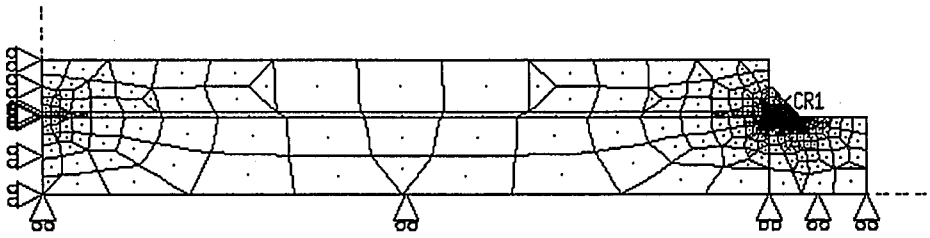


**Figure 3.24** Edge Cracked Flip Chip Model Constrained in the Vertical Direction

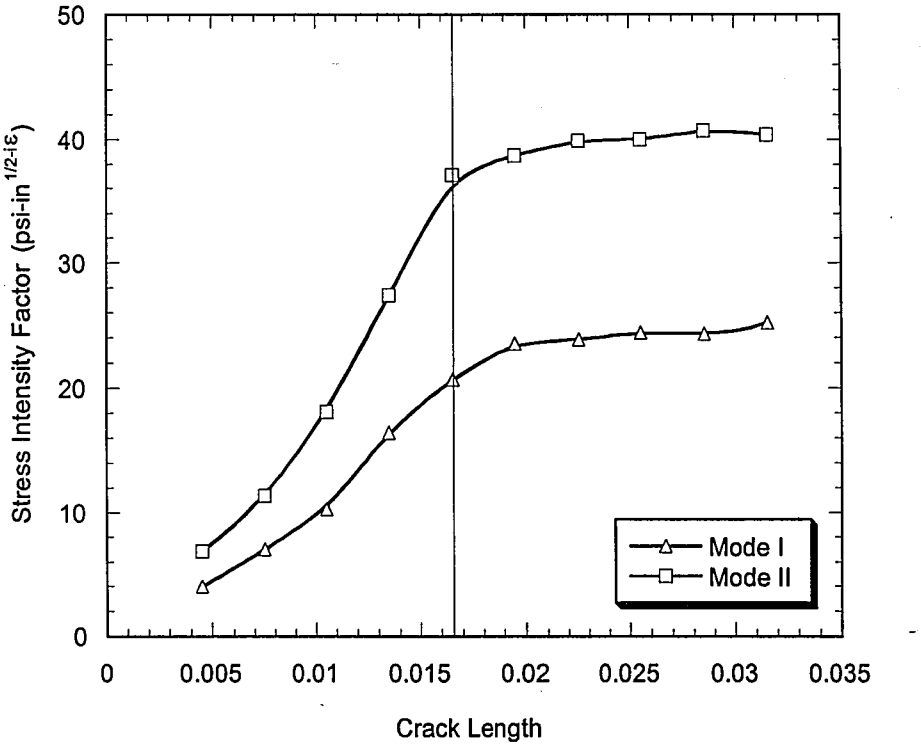


**Figure 3.25:** Stress intensity factors as a function of crack length for the edge cracked package constrained in the vertical direction, $\Delta T$=-10$^0$C.

Figure 3.25 shows the mode I and mode II stress intensity factors as a function of crack length. Both stress intensity factors increase as the crack length increase, until the crack length just passed below the corner of the silicon chip, which is shown on the graphs as a vertical line at the crack length of 0.0165 in. After passing the corner, the stress intensity factors seem to stabilize, they don't increase much. As it was expected, $K_{II}$ is higher than $K_I$ for every crack length since shear is predominant.

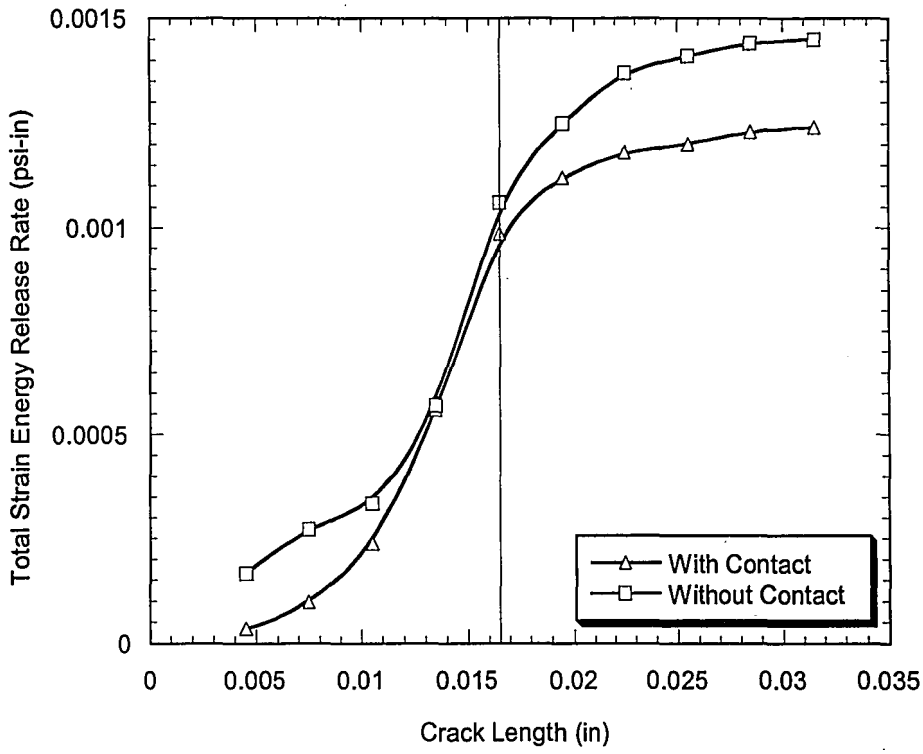


**Figure 3.26:** Total strain energy release rates as a function of crack length for the edge cracked package constrained in the vertical direction, $\Delta T=-10^0C$

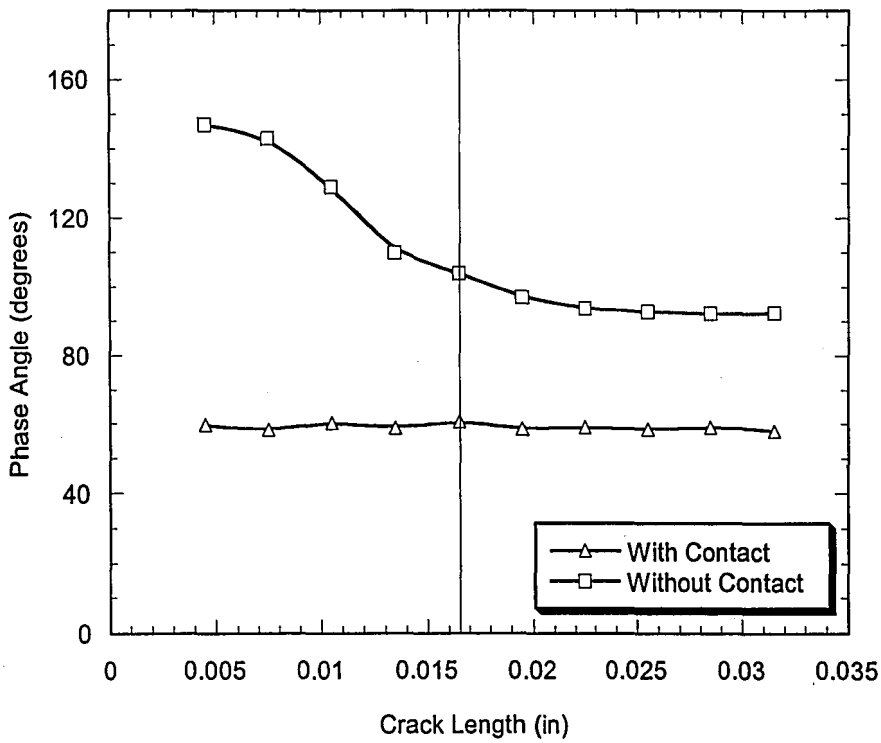


**Figure 3.27:** Phase angles as a function of crack length for the edge cracked package constrained in the vertical direction, $\Delta T=-10^0C$

The figure 3.26 shows the total strain energy release rates, obtained using contact analysis and without using contact analysis, as a function of crack length. The two curves show the same kind of curve, although the G values for the case without contact are larger than the ones with contact. The difference between the two curves are larger for small and large crack lengths, i.e away from the corner of the silicon chip and they are almost equal when the crack has just passed the corner.

Figure 3.27 shows the phase angles with and without contact analysis as a function of crack length. For the contact case, the phase angles are almost constant for every crack length, but for no contact case the angle decreases until the corner of the die after which it gets constant also. The difference is quite high, generally around $30^0$.

### 3.4.1.3 SUBSTRATE CONSTRAINED IN BOTH DIRECTIONS

As explained in section 3.4.1.2, neither boundary conditions represents the real problem with the substrate bonded to a board, but the real case falls between the two. Obviously constraining the substrate to move in both directions is a more severe condition than the real case, therefore in reality, the total strain energy release rates, phase angles, mode I and mode II stress intensity factors will be somewhere between the results of the two models.

For this case only $\Delta T=10^0C$ is considered since no contact is encountered for $\Delta T=-10^0C$. The model used for modeling this problem is shown in Fig. 3.28

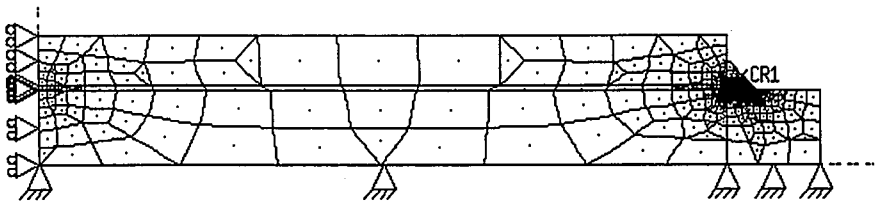


**Figure 3.28** Edge cracked flip chip model constrained in both directions

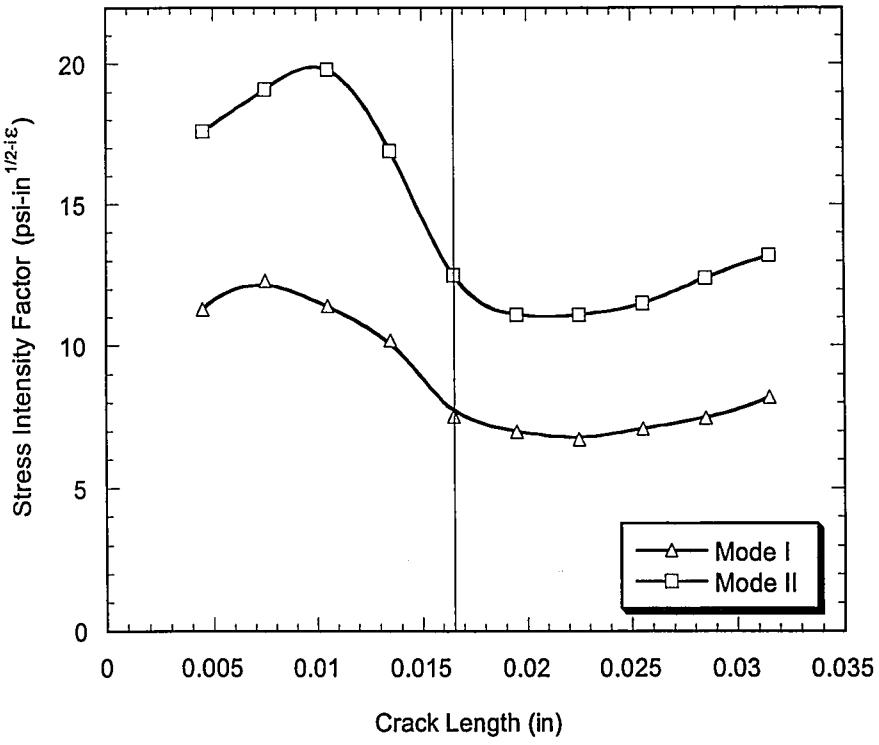


**Figure 3.29:** Stress intensity factors as a function of crack length for an edge cracked flip-chip package constrained in both directions, $\Delta T=10^0C$

Figure 3.29 shows the mode I and mode II stress intensity factors as a function of crack length. As expected, $K_{II}$ is greater than $K_I$ for every crack length, since shear is dominant. Both curves seem to have an inflection point somewhere near 0.0165 where the corner of the silicon die is. Both decreases as the crack approaches the silicon die corner and they increase again after having passed the corner.

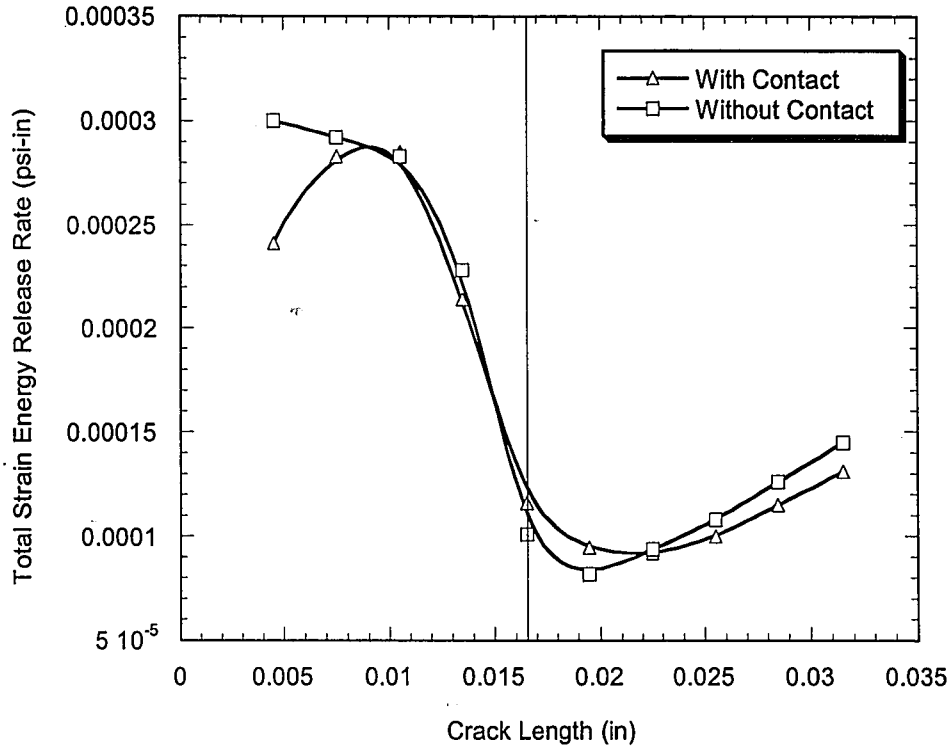


**Figure 3.30:** Total strain energy release rates as a function of crack length for an edge cracked package constrained in both directions, $\Delta T=10^0C$

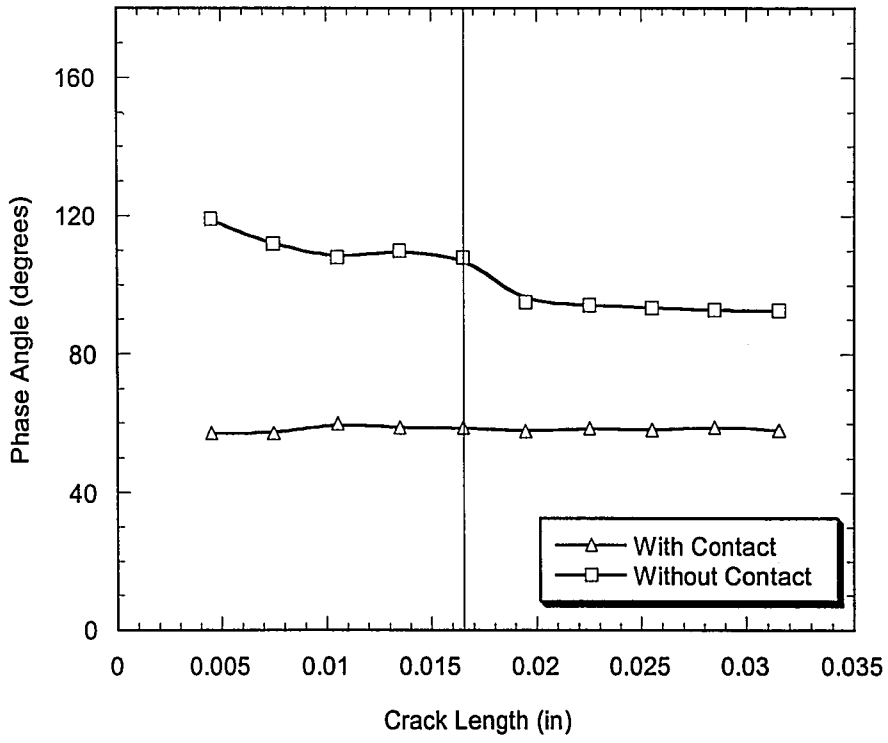


**Figure 3.31:** Phase angles as a function of crack length for an edge cracked package constrained in both directions, $\Delta T=10^0C$

Figure 3.30 shows the total strain energy release rates for the two cases considered, one with contact and the other without contact. The difference between two is larger for small crack lengths, up to a crack length of ~0.01 after which there is not much difference between the two cases. Both shows a minimum somewhere near the corner of the silicon die. Both decrease as the crack approaches the silicon die corner and they increase again after having passed the corner.

Figure 3.31 shows the phase angles for both cases as a function of crack length. The phase angles with contact analysis are again almost constant for every crack lengths. The phase angles obtained without contact analysis decrease for small crack length, then they get almost constant also. The difference between both results are quite high, more than $30^0$.

### 3.4.1.4 SUBSTRATE CONSTRAINED AT THE ENDS

This is another kind of boundary condition that is worth modeling. Today's computer chips are generally manufactured and assembled on the motherboard in such a way that the chip itself can be easily removed and replaced by a new one, for fixing or upgrading purposes. The substrate has some pins at the end which will lock into slots on the board. This problem can be modeled by assuming that the bottom surface of the substrate is free to move in both directions but the end of the substrate, where the pins are, is constrained to move in the vertical direction only. This is somewhat simplified model in the sense that the elasticity of the pins are neglected for simplicity purposes.

Only the temperature difference of $\Delta T=-10^0C$ is considered since no contact is observed to occur for the $\Delta T=10^0C$ case. The model used for this problem is shown in Figure 3.32

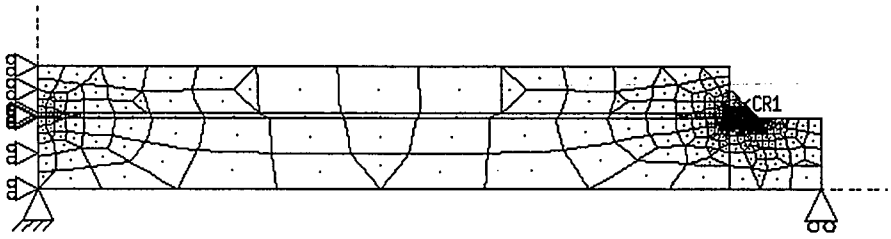


**Figure 3.32:** Edge cracked flip chip model constrained at the ends

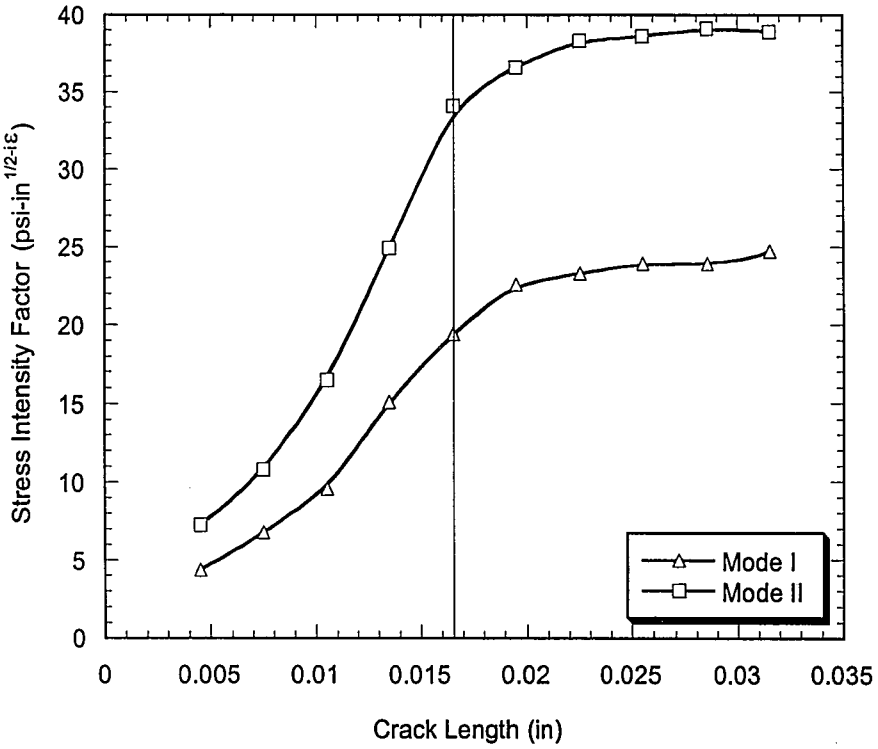


**Figure 3.33:** Stress intensity factors as a function of crack length for an edge cracked package constrained at the ends, $\Delta T=-10^0C$.

Figure 3.33 shows the mode I and mode II stress intensity factors as a function of crack length. The results are quite similar to the edge cracked package constrained in the vertical direction results in section 3.4.1.2. This is somewhat expected because the two boundary conditions are quite similar. Again the stress intensity factors increase rapidly for small crack lengths, then the slope tends to zero, after having passed the silicon die corner.

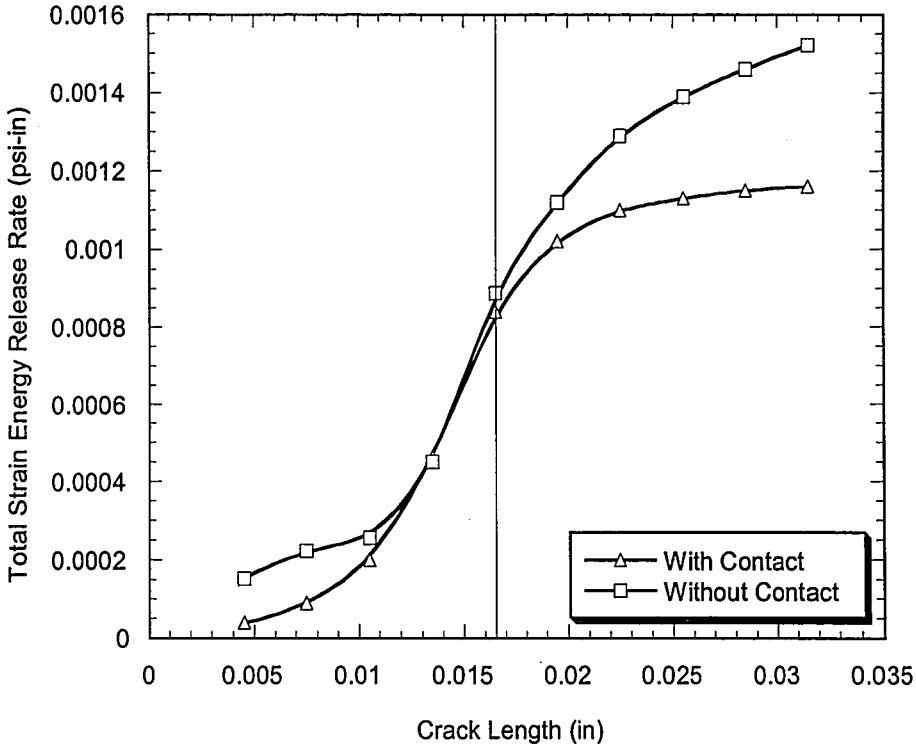


**Figure 3.34:** Total strain energy release rates as a function of crack length for an edge cracked package constrained at the ends, $\Delta T$=-10$^0$C.

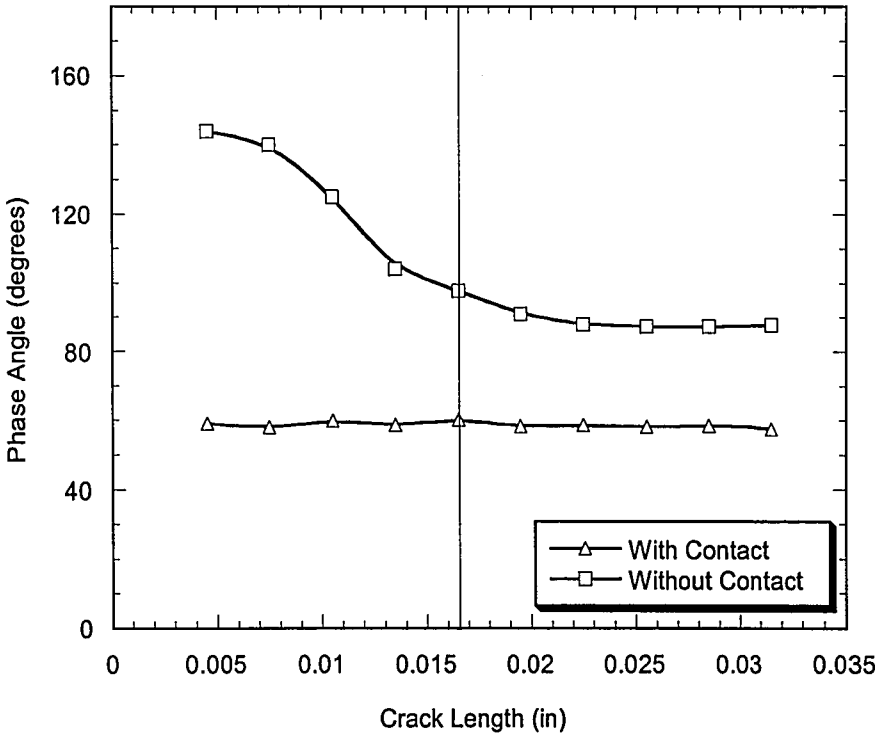


**Figure 3.35:** Phase angles as a function of crack length for an edge cracked package constrained at the ends, $\Delta T$=-10$^0$C.

Figure 3.34 shows the total strain energy release rates for the contact and no contact analysis cases as a function of crack length. Again the two curves are quite similar to the ones obtained in section 3.4.1.2. As the crack length increases, the results for both cases become almost equal to each other, after passing the silicon corner, which is showed on the graph as a vertical line at a crack length of 0.0165 in, the two curves separate again. The strain energy release rates for the case without contact analysis are almost always higher than the ones with contact analysis. In a way, without putting the contact constraints, one overestimates the total strain energy release rates.

Figure 3.35 shows the phase angles for the contact analysis and the analysis without contact as a function of crack length. Again the results are very close to the ones obtained in section 3.4.1.2. The phase angles obtained with contact analysis are almost constant for every crack length, while the ones obtained without contact analysis decreases for small crack lengths and then after having passed the silicon corner, gets constant also. The difference between them is quite high, almost $30^0$ everywhere.

## 3.4.2 CENTER-CRACKED FLIP-CHIP PACKAGE

This is the other common crack problem in flip chip packages. As it will be shown in the following sections, it is a less critical case than the edge cracked situation in a design point of view since the total strain energy release rates, which are the primary comparative quantity in interface crack problems, are much lower in center cracked packages. Nevertheless, it is still an interesting problem to be handled. Three boundary conditions are modeled which are substrate constrained in vertical

direction, substrate constrained in both directions and substrate constrained at the ends. The unconstrained substrate case is not investigated because the stress intensity and total strain energy release rates were far too small to be of any interest. Results are obtained for crack lengths from 0.02 in to 0.20 in, which is about the quarter of the length of the underfill.

### 3.4.2.1 SUBSTRATE CONSTRAINED IN VERTICAL DIRECTION

As explained earlier in section 3.4.1.2, the real fracture mechanics quantities such as total strain energy release rates, phase angles, mode I and mode II stress intensity factors of a center cracked flip chip package bonded to a board will be somewhere between the results two boundary conditions which will be modeled in this section and in section 3.4.2.3. In this section the bottom surface of the substrate is allowed move in the horizontal direction only. The model used for this case is shown in Fig 3.36

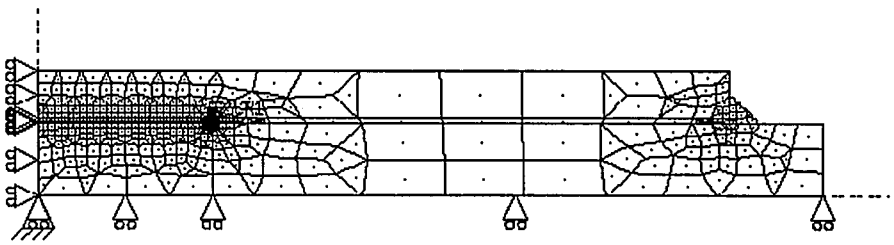

**Figure 3.36:** Center cracked flip chip model constrained in the vertical direction only

Figure 3.37 shows the stress intensity factors as a function of crack length. As expected, the magnitude of the mode II stress intensity factors are higher than the ones for mode I, since shear is dominant at the interface. Due to the complex nature of the problem, it's very difficult to explain the bumps on the plots

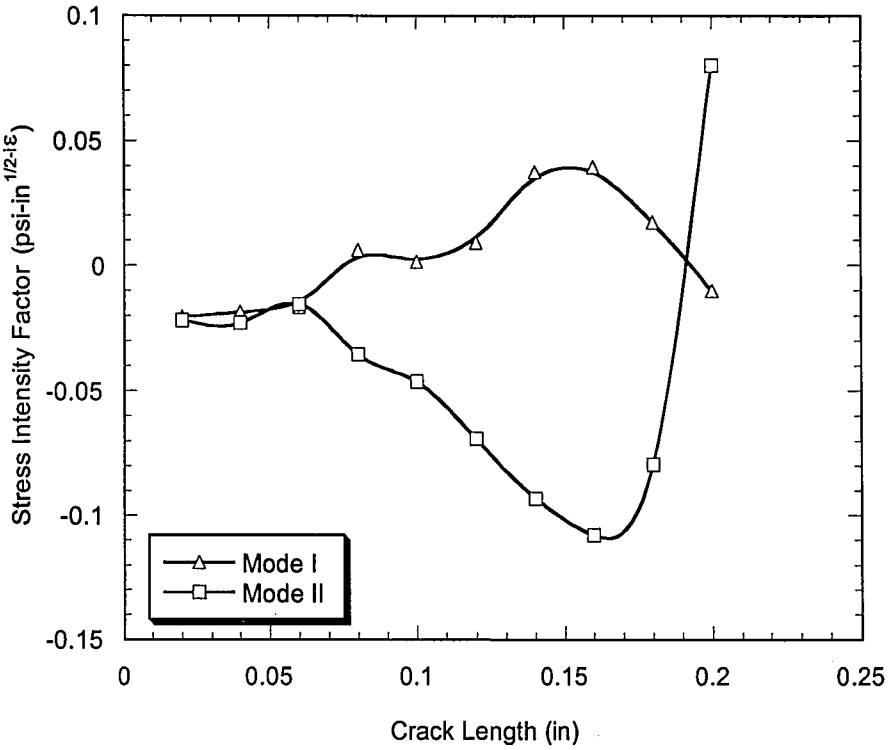


**Figure 3.37:** Stress intensity factors as a function of crack length for a center cracked flip chip package constrained in the vertical direction, $\Delta T$=-10$^0$C.

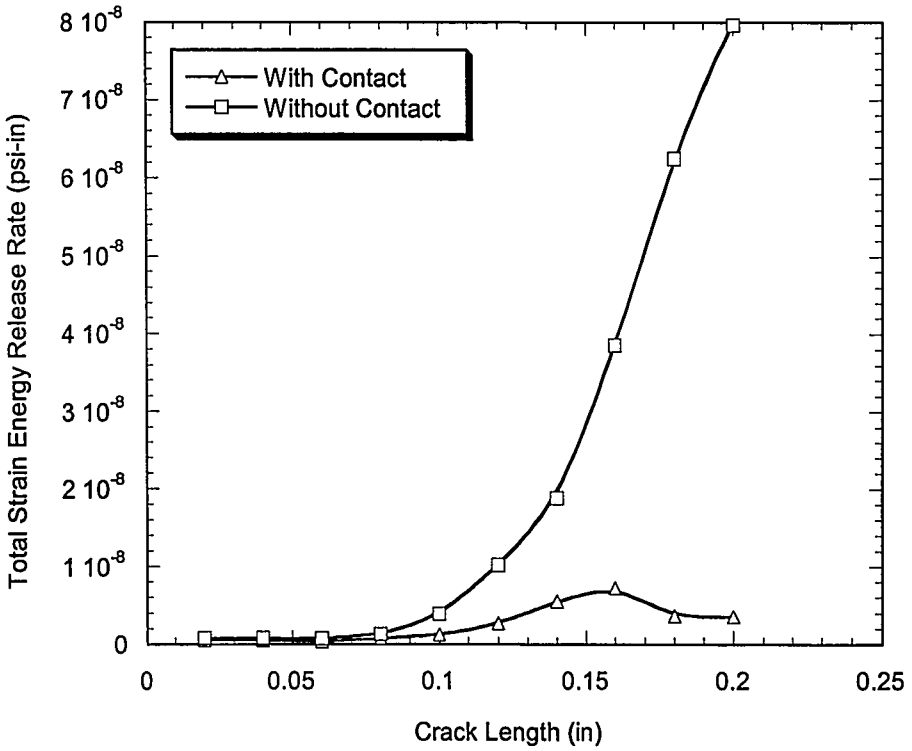


**Figure 3.38:** Total strain energy release rates as a function of crack length for a center cracked flip chip package constrained in the vertical direction, $\Delta T$=-10$^0$C.

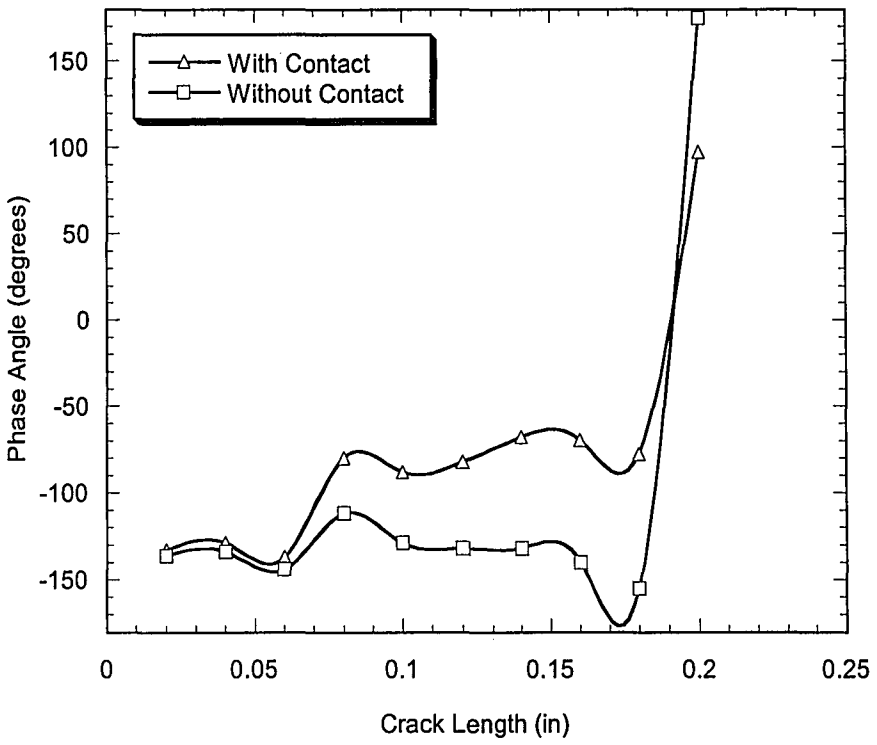


**Figure 3.39:** Phase angles as a function of crack length for a center cracked flip chip package constrained in the vertical direction, $\Delta T=-10^0C$.

Figure 3.38 shows the total strain energy release rates as a function of crack length, obtained from contact analysis and without contact analysis. It is clear that without using contact analysis, the G values are overestimated. For small crack lengths they are almost equal, after 0.1 in of crack length the difference becomes quite large, at a crack length of 0.2 in, the one obtained without contact analysis is eight times the one obtained with contact analysis.

Figure 3.39 shows the phase angles as a function of crack length, obtained from contact analysis and without using contact analysis. Again, the complex nature of the problem makes it very difficult to explain the bumps on this plot. Note that

the difference in the phase angles is quite small for small crack lengths and it increases continuously with increasing crack length.

### 3.4.2.2 SUBSTRATE CONSTRAINED IN BOTH DIRECTIONS

As a continuation of the section 3.4.2.2, now the other boundary condition left, which is the substrate constrained in both directions, will be modeled. The model used in this problem is shown in Fig. 3.40

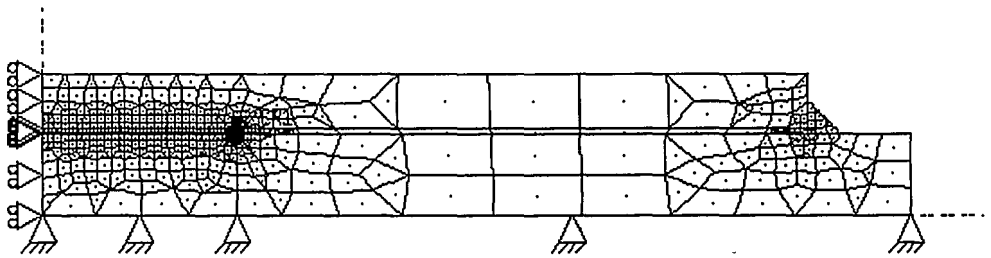

**Figure 3.40:** Center cracked flip chip model constrained in both directions

Figure 3.41 shows the mode I and mode II stress intensity factors as a function of crack length. They both show the same behavior. Both increase in magnitude as the crack length is increased. As it was expected in other cases, the mode II stress intensity factor dominates in magnitude over mode I, since shear is dominant at the interface.

Figure 3.42 shows the total strain energy release rates, obtained with contact analysis and without using contact analysis, as a function of crack length. They both increase as the crack length is increased and show the same kind of behavior with respect to crack length. The difference between them is quite small, but nevertheless the results obtained without contact analysis are higher then the ones obtained with contact analysis.

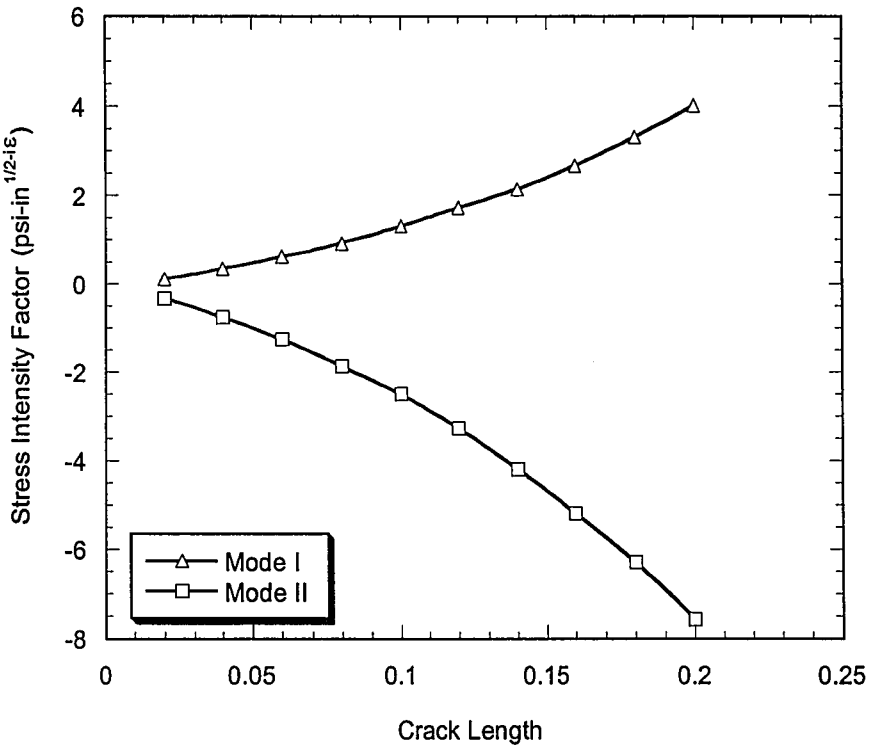


**Figure 3.41:** Stress intensity factors as a function of crack length for a center cracked flip chip package constrained in the vertical direction, $\Delta T=-10^0C$.

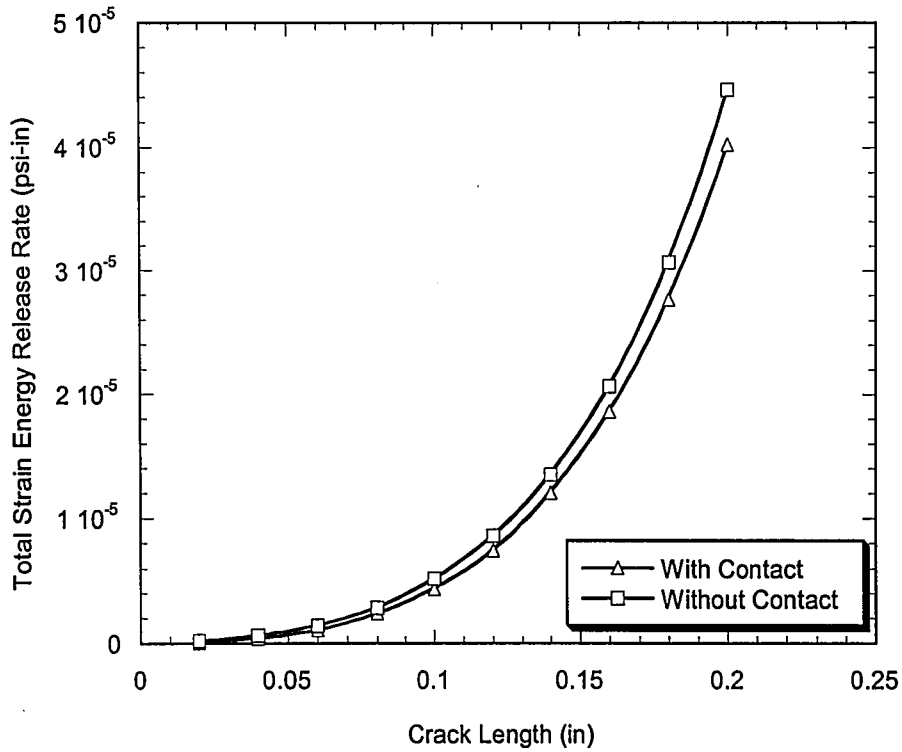


**Figure 3.42:** Total strain energy release rates as a function of crack length for a center cracked package constrained in the vertical direction, $\Delta T=-10^0C$.

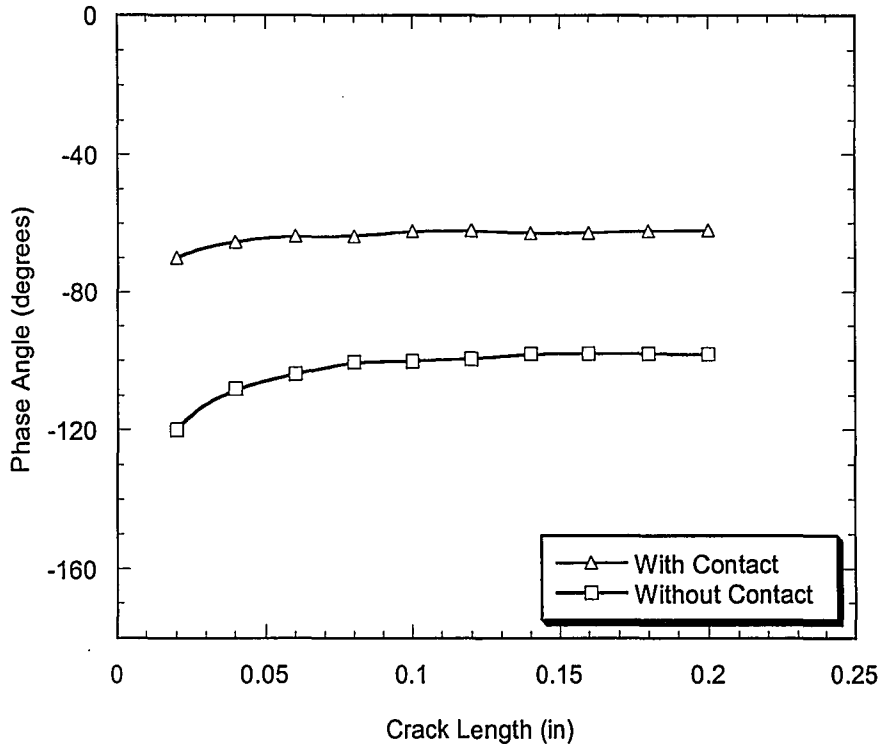


**Figure 3.43:** Phase angles as a function of crack length for a center cracked package constrained in the vertical direction, $\Delta T=-10^0C$.

Figure 3.43 shows the phase angles, obtained with contact analysis and without using contact analysis, as a function of crack length. Both increase for small crack lengths which, then they both get almost constant. The difference is quite high, about $30^0$.

### 3.4.2.3 SUBSTRATE CONSTRAINED AT THE ENDS

As the last of the semiconductor applications, in this section edge cracked flip chip package where the substrate is constrained at the end, as some PC chips are pined to the board for easy removal. Only the case for $\Delta T=-10^0C$ is considered because with an a loading with the opposite sign contact doesn't occur. The model used in this case is shown in Fig. 3.44

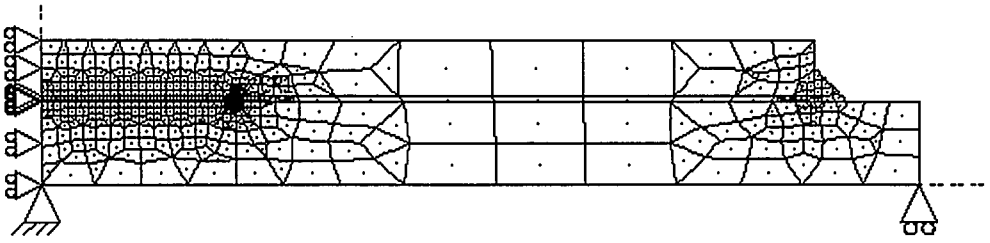

**Figure 3.44:** Center cracked flip chip model constrained at the ends

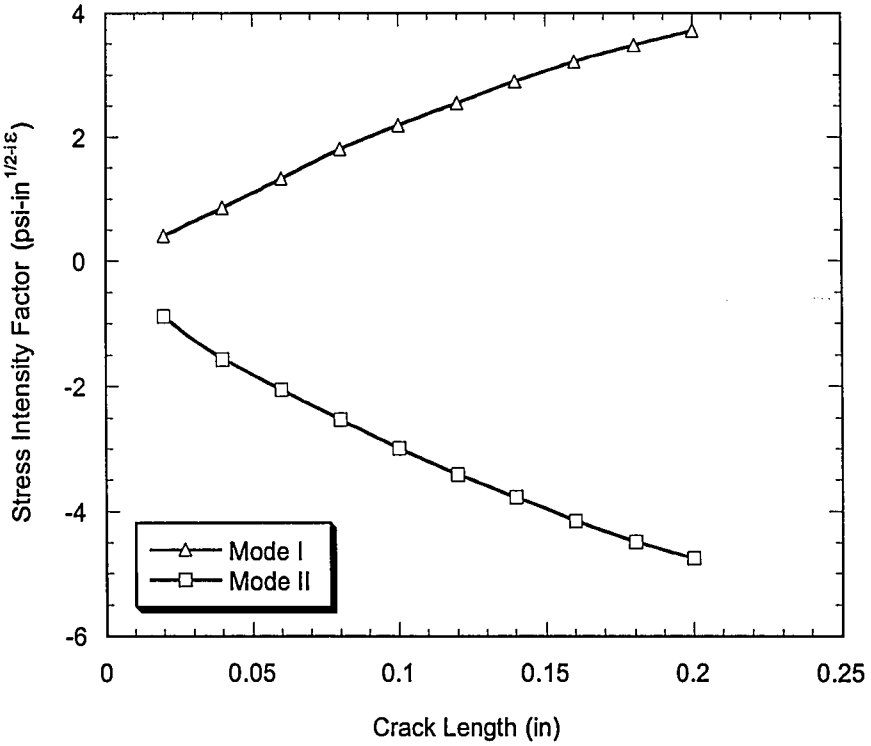


**Figure 3.45:** Stress intensity factors as a function of crack length for a center cracked package constrained at the ends, $\Delta T=-10^0C$.

Figure 3.45 shows the mode I and mode II stress intensity factors as a function of crack length. Both shows a linear-like increasing behavior with increasing crack length, the slop decreasing quite slowly. In magnitude, the mode II dominates over mode I, which is an expected observation since shear is dominant at the interface. Their behavior is quite similar to the one for which package was constrained in both directions.

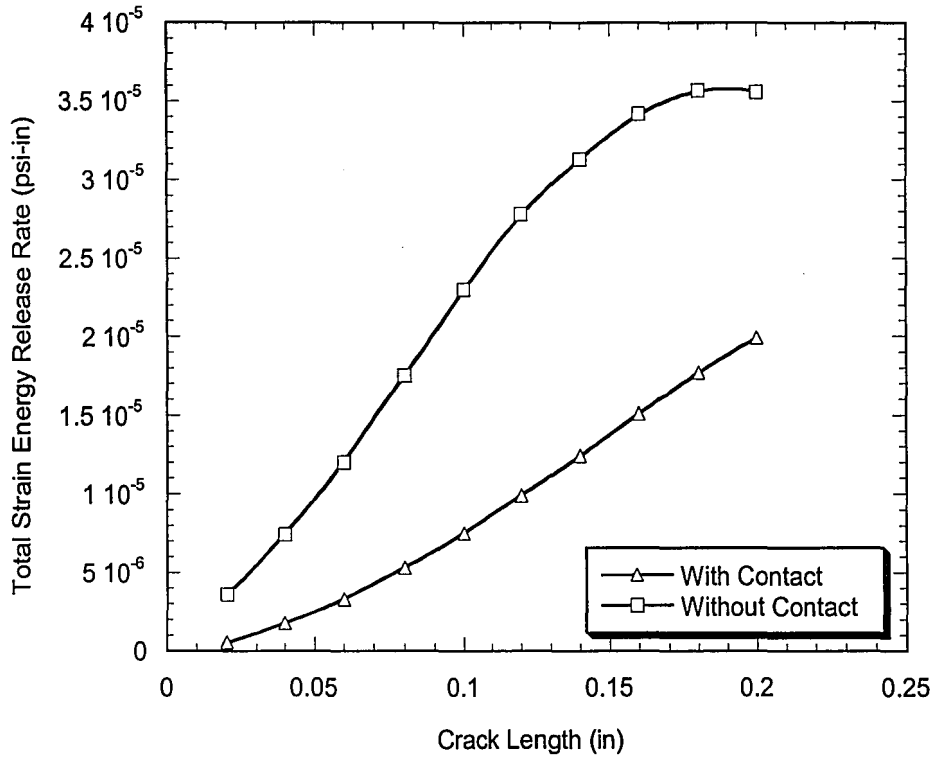


**Figure 3.46:** Total strain energy release rates as a function of crack length for a center cracked package constrained at the ends, $\Delta T=-10^0C$.

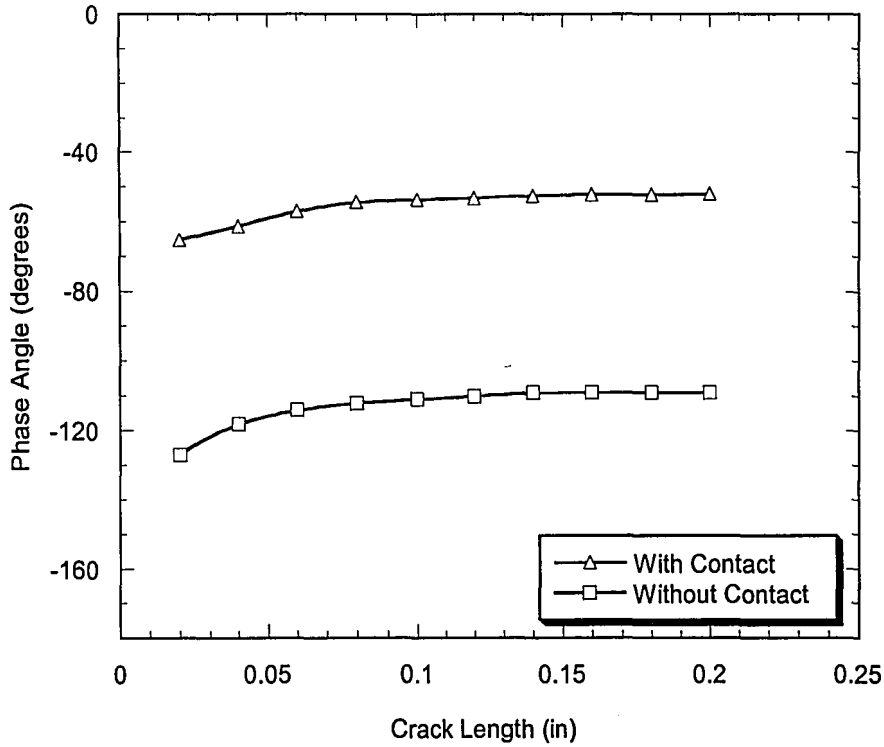


**Figure 3.47:** Phase angles as a function of crack length for a center cracked package constrained at the ends, $\Delta T=-10^0C$.

Figure 3.46 shows the total strain energy release rates, obtained from contact analysis and without contact analysis, as a function of crack length. The values obtained without contact analysis are again overestimated. The energy release rates for the contact case show a linear-like behavior in terms of crack length and it increases directly proportional to the crack length.

Figure 3.47 shows the phase angles, obtained from contact analysis and without contact analysis, as a function of crack length. They both increase for small crack lengths and then both gets almost constant. The difference is quite large, almost $50^0$ everywhere.

# CHAPTER 4

# CONCLUSION

In this study, interface crack surface contact problems are investigated and special attention is given to semiconductor package applications. The finite element method is briefly overviewed. Enriched element formulation, which is used to find the relevant fracture mechanics quantities such as stress intensity factors, strain energy release rates, phase angles, is presented. A general two dimensional contact analysis algorithm using the penalty function method is described in detail. This algorithm has been implemented into a finite element code, namely Frac2D. In order to prove the correctness of the solutions some examples have been solved: two blocks in sticking/sliding contact, a flat punch on an elastic foundation and a cylindrical punch on an elastic foundation. Numerical examples are mainly focused on flip chip packages. Both edge-cracked and center-cracked packages are considered with four different boundary conditions: unconstrained package, package constrained in the vertical direction only, package constrained in both directions and package constrained at the ends. Mode I and mode II stress intensity factors, total strain energy release rates and phase angles are plotted as a function of crack length. Total strain energy release rate and phase angle values are compared with the results obtained without using a contact algorithm, i.e. the mesh is allowed to overlap.

The following are the general conclusions. For interface crack problems, there is no magical rule to prevent contact. Depending on the boundary conditions, geometry of the package, material combinations, and loads applied contact will or will not occur. Detecting contact is quite easy with homogeneous cracks since $K_I$ will

become negative. There's no such rule for bimaterial interface cracks. The only possibility for catching a mesh overlap, other than using a contact algorithm, is to plot the deformed shape, but this may be misleading. Therefore, the best way would be to use a contact algorithm whenever contact is suspected to occur. Mesh overlap is inadmissible and thus it must be prevented.

For bimaterial interface cracks, the integrity of the package is decided upon comparing the actual total strain energy release rate with a critical strain energy release rate which is found experimentally. In almost all of the cases, the total strain energy release rate obtained without using a contact algorithm, i.e. allowing the mesh to overlap, is found to be higher than the one obtained with the contact algorithm. This means that in most of the cases even without considering possible contacts one can still be on the safe side, from a design point-of-view. This is not to be generalized, since for some cases the reverse may be true. Allowing the mesh to overlap may give results that are close to or quite different than the actual case (considering contact), depending on the geometry, boundary conditions, etc. For example, consider a center-cracked flip chip package whose substrate is constrained in both directions. Both G values are quite close to each other and ignoring contact gives very close results. Consider an edge-cracked flip chip package whose substrate is unconstrained. Two G values are quite different from each other; the one obtained without using a contact algorithm is generally two times higher than the other. This may result in an overly designed package: a package predicted to fail may still be on the safe side. One interesting observation can be made: for both edge and center cracked packages whose substrate is constrained in both directions, the G values from contact analysis and from usual finite element methods are quite close to each other for a large range of crack lengths. This may be due to the severe boundary condition

on the substrate: the crack surfaces does not penetrate into each other considerably, resulting in similar G values. It should be noted that there is no evidence that this can be generalized. As a general statement, it is recommended that a finite element code into which a contact algorithm is implemented should be used whenever mesh overlap is a possibility.

In the calculations, the coefficient of friction is taken to be zero for simplicity purposes. As a future work, coefficients of friction other than zero may be used and their effect on the total strain energy release rates, stress intensity factors and phase angles may be examined.

In semiconductor package applications, contact occurs mainly due to the mismatch in the coefficients of thermal expansion. Since plain strain is assumed for every case, it is indirectly assumed that the difference in the CTE's in the third dimension is not important, which is generally not the case. The correctness of the results depends on this assumption. Therefore, it should be noted that interface crack surface contact problems in semiconductor packages under thermal or moisture loads is truly three dimensional and consequently a three dimensional finite element analysis with a three dimensional contact algorithm is necessary to fully understand and solve the actual problem. Also, the axisymmetric case, which is a more realistic interpretation of the actual situation, can be investigated and the results may be compared with the plane strain and 3-D cases.

# APPENDIX A

## HERTZIAN CONTACT THEORY

In this appendix, Hertzian theory on the contact of elastic bodies will be shortly presented. The first studies on elastic bodies in contact dates back to work of Hertz [34], which is accepted as the birth of this field. Since then, a great number of papers have been published in this area but the results and conclusions of Hertz still remains valid, partly due to the simplicity of the solutions.

Hertz came up with this theory while he was investigating the possible effects of elastic deformations on Newton's optical interference fringes in a Christmas vacation at the age of twenty-three. The theory applies to smooth (i.e. zero coefficient of friction) and nonconformable surfaces in contact. He assumed that each body can be regarded as an elastic half-space loaded over a small elliptical region of its plane surface. By this assumption, the highly concentrated contact stresses are analyzed separately from the general distribution of stresses in the bodies due their shapes and due to the various boundary conditions present. Also, the methods of solving boundary value problems involving elastic half-spaces were already well developed. In order for this assumption to hold, the following should be satisfied:

a) The size of the contact zone has to be small compared to the dimensions of the bodies. This condition ensures that the various boundary conditions acting on the bodies does not affect the solution at the contact zone.

b) The size of the contact zone has to be small compared to the radii of curvature of the two bodies. This condition first ensures that the region just outside the contact zone can be regarded as half-space and secondly keeps the strains at the contact zone in the limits of linear elasticity.

Finally, he assumed that the contacting surfaces are smooth, therefore only normal tractions can be transmitted through the contact zone. With these assumptions and after a derivation that can be found in [35], the following formulas are obtained for two cylinders in contact:

Defining an "equivalent modulus" $E^*$ and "equivalent radius" $R^*$,

$$E^* = \left(\frac{1-\nu_1^2}{E_1} + \frac{1-\nu_2^2}{E_2}\right)^{-1}$$

$$R^* = \left(\frac{1}{R_1} + \frac{1}{R_2}\right)^{-1}$$

The semi-contact width $a$ is given by

$$a = \left(\frac{4PR^*}{\pi E^*}\right)^{\frac{1}{2}}$$

The maximum contact pressure, which occurs at the first point of touch on the axis of symmetry, is

$$p_0 = \frac{2P}{\pi a} = \left(\frac{PE^*}{\pi R^*}\right)^{\frac{1}{2}}$$

Maximum shear stress occurs at x=0, z=0.78a and is equal to

$$\tau_1 = 0.30 p_0$$

If one of the cylinders have an infinite radius, it will become a half space, and the problem will become a cylindrical punch on elastic foundation problem. Setting one of the radii as infinite and assuming that both bodies have the same materials properties $E$ and $\nu$, one would obtain

$$E^* = \frac{E}{2(1-\nu^2)}$$

$$R^* = \frac{R}{2}$$

$$a = \left(\frac{4PR(1-\nu^2)}{\pi E}\right)^{\frac{1}{2}}$$

In the numerical examples chapter, chapter 3, a cylindrical punch on an elastic foundation problem was solved. It was found that the contact zone is small compared to both the radius of the cylindrical punch and the dimensions of the block. Zero coefficient of friction was used to satisfy the smooth surface condition. The strains at the contact zone were also small such that linear theory of elasticity holds. Therefore each and every assumption of the Hertz theory have been satisfied. The results of the finite element analysis proved to be very close to the Hertzian solution.

# APPENDIX B

## STRESS SMOOTHING

It is known that when Gaussian integration is used, the stresses found are the most accurate when computed at the Gaussian points. Since 4x4 Gaussian integration has been used in this work for isoparametric cubic elements, we obtain the stresses at 16 different locations within an element. Rarely these stresses are being used because generally nodal stresses are desired. These nodal stresses are found by a process called *stress smoothing*. In a more general way, stress smoothing *smoothes* the discontinuous stress distribution within an element or within the whole model by minimizing a certain functional accepting that the stresses found at the Gaussian points are the most correct ones. The reason for the discontinuity of the stress field at the extremities of an element is that in conventional finite elements, the continuity of the displacements are imposed into the equations but no action is taken regarding the stresses. The nodal stresses could have been found by using the cubic shape functions but it is known that the interpolating functions tend to behave badly at the extremities of an interpolation region. Two methods are quite popular:

### B.1 GLOBAL STRESS SMOOTHING

Besides being the computationally most difficult, global stress smoothing is also the most accurate one. A least squares minimization technique is adopted. The proposed smoothing function has the form:

$$g(x,y) = a_1 + a_2 x + a_3 y + a_4 xy + \ldots \qquad \text{(B.1)}$$

$$= \sum a_{ij} x^i y^j \qquad \begin{cases} i = 1,3 \\ j = 1,3 \end{cases}$$

If the unsmoothed stresses are expressed by $\sigma(x, y)$, by minimizing the functional $\phi$ with respect to the coefficients $a_{ij}$ we have

$$\phi = \int\int (\sigma - g)^2 dx\, dy \tag{B.2}$$

$$\frac{\partial \phi}{\partial a_{ij}} = 0 \tag{B.3}$$

Equation B.3 gives a set of linear simultaneous equations which will be solved for the unknown coefficients. As the name implies, all of the stresses in the model will be smoothed at the same time. After the smoothing the stresses at the extremities of the elements becomes continuous but the method is quite time consuming. Also a finite element like formulation is proposed in [36].

## B.2 LOCAL STRESS SMOOTHING

In local stress smoothing, the stress within an element is assumed to vary like

$$\sigma = a_1 + a_2 s + a_3 t + a_4 st + a_5 s^2 + a_6 t^2 + a_7 s^2 t + ... + a_{16} s^3 t^3 \tag{B.4}$$

Since we have 16 Gaussian points, Eq B.4 can be written in matrix form as

$$\begin{Bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \cdots \\ \cdots \\ \sigma_{16} \end{Bmatrix} = \begin{bmatrix} 1 & s & t & \cdots & \cdots & s^3t^3 \\ 1 & s & t & \cdots & \cdots & s^3t^3 \\ 1 & s & t & \cdots & \cdots & s^3t^3 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & s & t & \cdots & \cdots & s^3t^3 \end{bmatrix} \begin{Bmatrix} a_1 \\ a_2 \\ a_3 \\ \cdots \\ \cdots \\ a_{16} \end{Bmatrix} \tag{B.5}$$

Here the unknowns are $a_i$'s. This equation can be inverted as

$$\begin{Bmatrix} a_1 \\ a_2 \\ a_3 \\ \cdots \\ \cdots \\ a_{16} \end{Bmatrix} = \begin{bmatrix} 1 & s & t & \ldots & \ldots & s^3t^3 \\ 1 & s & t & \ldots & \ldots & s^3t^3 \\ 1 & s & t & \ldots & \ldots & s^3t^3 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & s & t & \ldots & \ldots & s^3t^3 \end{bmatrix}^{-1} \begin{Bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \cdots \\ \cdots \\ \sigma_{16} \end{Bmatrix} \tag{B.6}$$

Since we known the isoparametric coordinates s and t of the Gaussian points, this 16 by 16 equation can easily be solved for the unknown $a_i$'s. Once they're known, we have the assumed distribution of the stresses within that element. To find the nodal stresses at a node, Eq. B.4 will be used with the isoparametric coordinates s and t of that node, which are already known.

Note that if a particular node is shared between $n$ elements, for that node $n$ different stresses will be found. These stresses will be then averaged arithmetically to find the nodal stress values.

Although this method is quite simple and useful, it cannot force the stress field within the model to be continuous at the extremities, like does the global stress smoothing. Also, in some cases it can give slightly wrong results. Such problems occurred in Chapter 3 in the flat punch on elastic foundation and cylindrical punch on elastic foundation problems. Previously it has been mentioned that the stresses of shared nodes will be averaged. Note that although contacting surfaces *touch* each other, the stress values of their nodes are not averaged because they are not physically attached to each other by means of connectivity matrices. Therefore, when calculating the nodal stresses on the contactor surface for example, the stress information that should come from the target surface is not taken into consideration. If one takes a look at Figure 3.10 and Figure 3.13, it can be seen that if the stresses on the contactor

side are somewhat above the known solution, then the stresses on the target side are found to be somewhat below the solution, therefore making the average fall almost exactly on the known solution of the problem. The difference between nodal results and known solution are bigger at the corner nodes of an element. This is an expected results since in stress smoothing, the stresses at nodes farther from Gaussian points, i.e corner nodes, will be obviously less accurate. A more detailed discussion on the various stress smoothing techniques is presented in [36].

As an advice, if local stress smoothing algorithm is used to find the nodal stresses, it is advised that the stresses coming from the contactor side and from the target side are averaged.

# APPENDIX C

## INTEGRATION ORDER FOR ENRICHED ELEMENTS

As presented in Section 2.1.1, there are some integrals to be calculated in the formulation of the finite element equations. In most of the cases, this kind of integrals are found using Gaussian integration, which is explained in almost every numerical methods books. The main parameter in the Gaussian integration scheme is the integration order. For every type of interpolating function used, there is a minimum integration order to be used in order to get accurate results. For example, for elements with cubic interpolation, 4 by 4 integration has to be used. Things get a little bit complicated when a crack tip element (whether enriched or transition) is to be used, since there is no such a minimum integration order. This is due to the extra terms in equations 2.9 and 2.10 that comes from the asymptotic displacements. Therefore the choice of the integration order is solely left to the initiative and experience of the user. Generally the use of 12 integration points gives good results for linear elastic fracture mechanics problems, but some problems arise when the crack surfaces are in contact with each other.

A certain "*displacement jump*" problem occurred when modeling center cracked flip-chip package with the substrate being constrained in the vertical direction only. Figure C.1 shows the y-displacements on the contactor surface. It can be seen that just at the enriched element, the displacements are not smooth and exhibits a *jump.* The same kind of jump has been observed for all of the crack lengths examined. This behavior disappears, as shown in Figure C.2, when the integration order for the enriched elements is taken as 60. It seems like 12-point-integration was not enough to

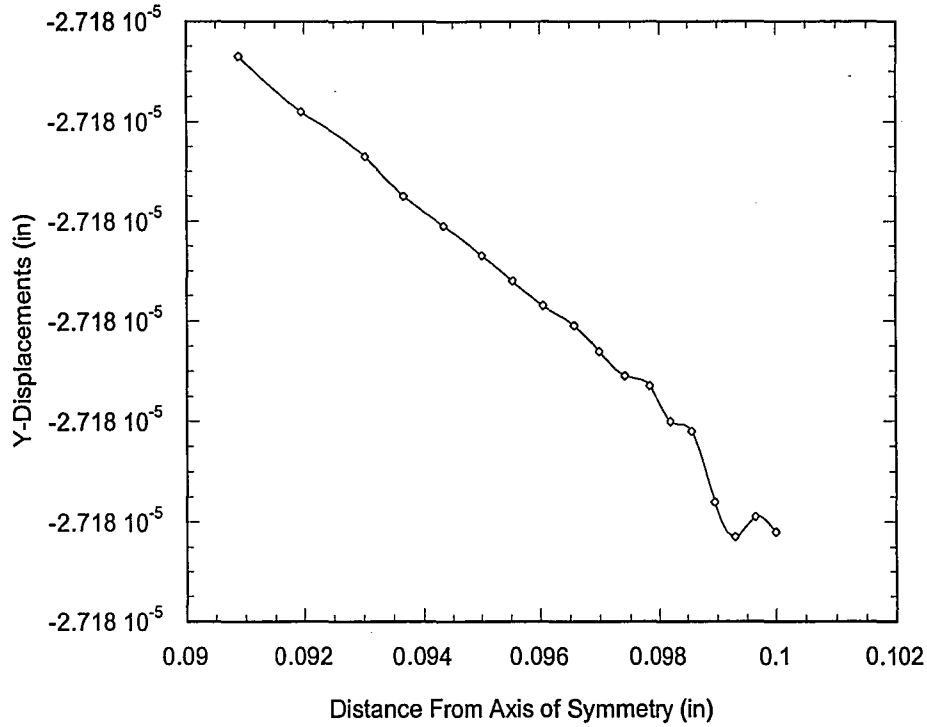


**Figure C.1:** Y displacements in a center cracked flip chip package constrained in the vertical direction only, with a crack length of 0.10 in, 12 integration points used.

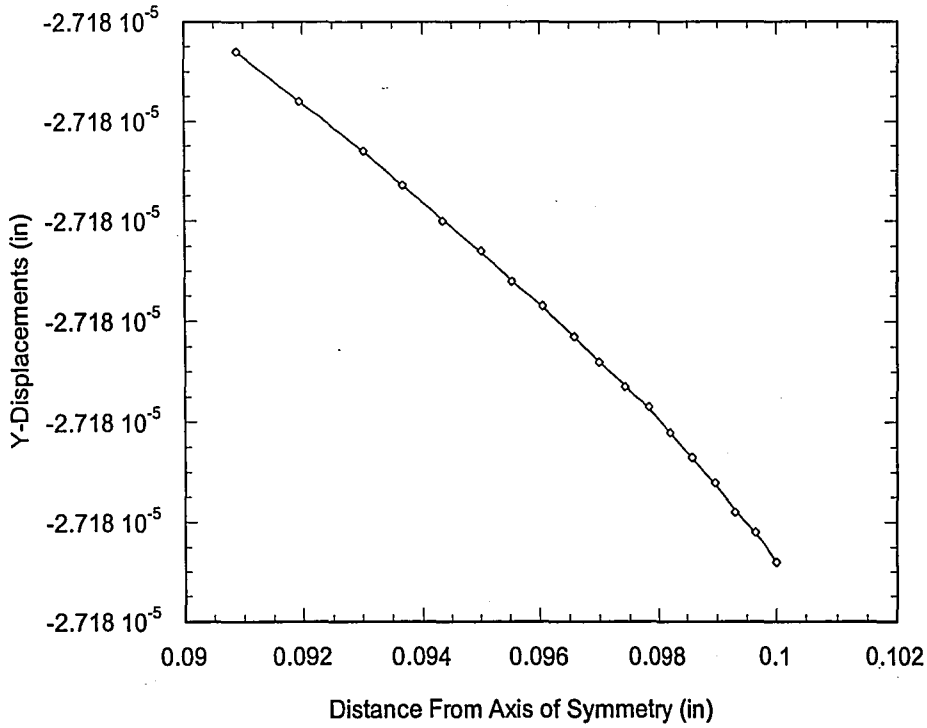


**Figure C.2:** Y displacements in a center cracked flip chip package constrained in the vertical direction only, with a crack length of 0.10 in, 60 integration points used.

correctly find the displacements in the neighborhood of the crack tip. Analytical solutions for interface cracks predicts some wild oscillations and contact at the crack tip. This may be the reason for the jump, since the displacements are quite oscillatory and complex in nature, apparently 12 was not enough as the integration order.

The same problem with the same crack length of 0.10 in has been run using various integration orders in order to catch the order at which the solutions converges. The results are shown in Figures C.3, C.4 and C.5 for the stress intensity factors, total strain energy release rates and phase angles. It can be seen that the results pretty much converges for the order of 60. Therefore, to not worry about this problem for other boundary conditions and crack geometries, 60 is taken as the integration order for the enriched elements whenever they are present in a model.

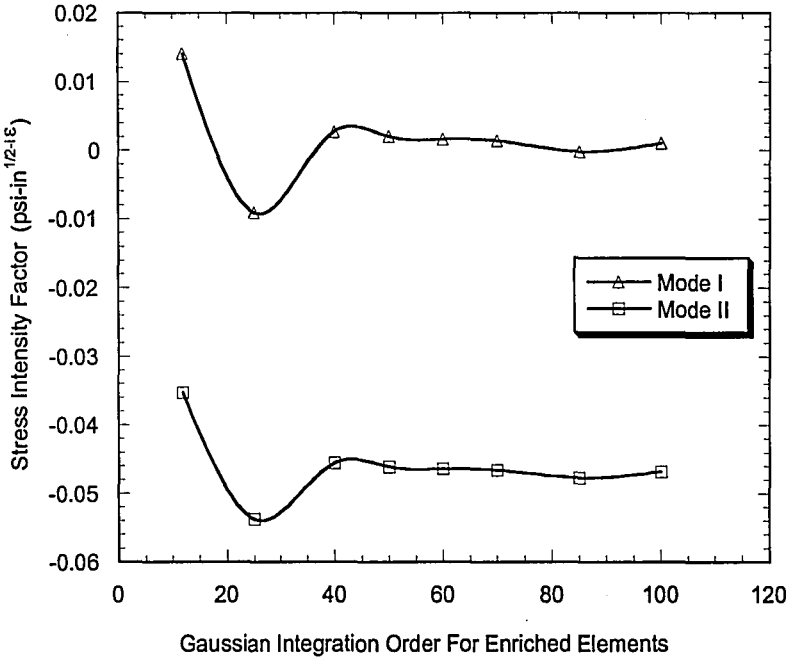


**Figure C.3:** Variation of the stress intensity factors with increasing integration order for enriched elements in a center-cracked flip-chip model with crack length 0.10 in.

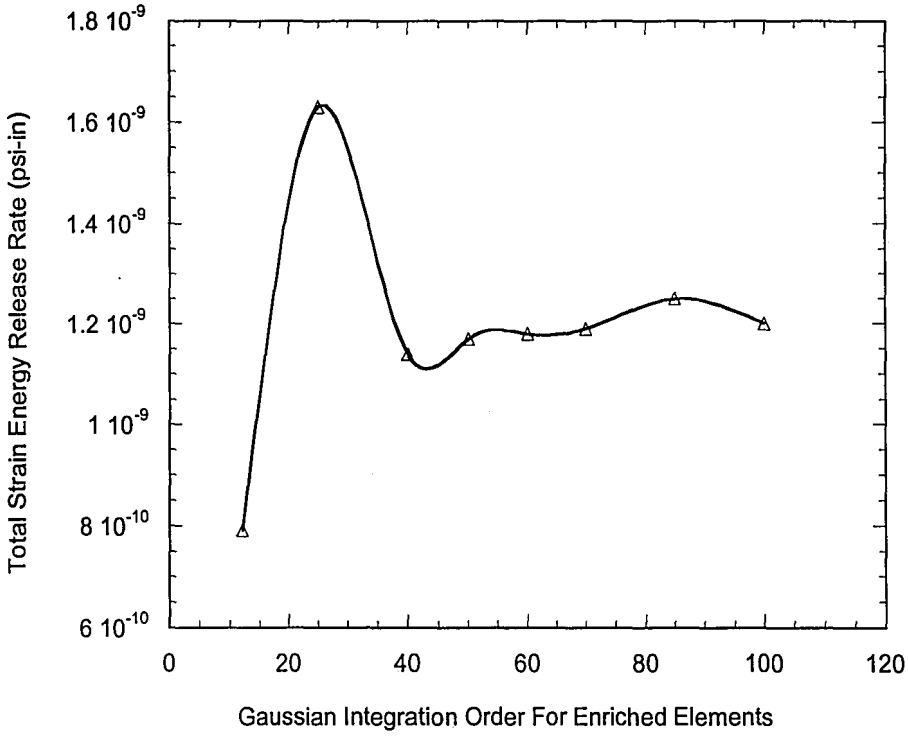


**Figure C.4:** Variation of the total strain energy release rate with increasing integration order for enriched elements in a center-cracked flip-chip model with a crack length of 0.10 in.

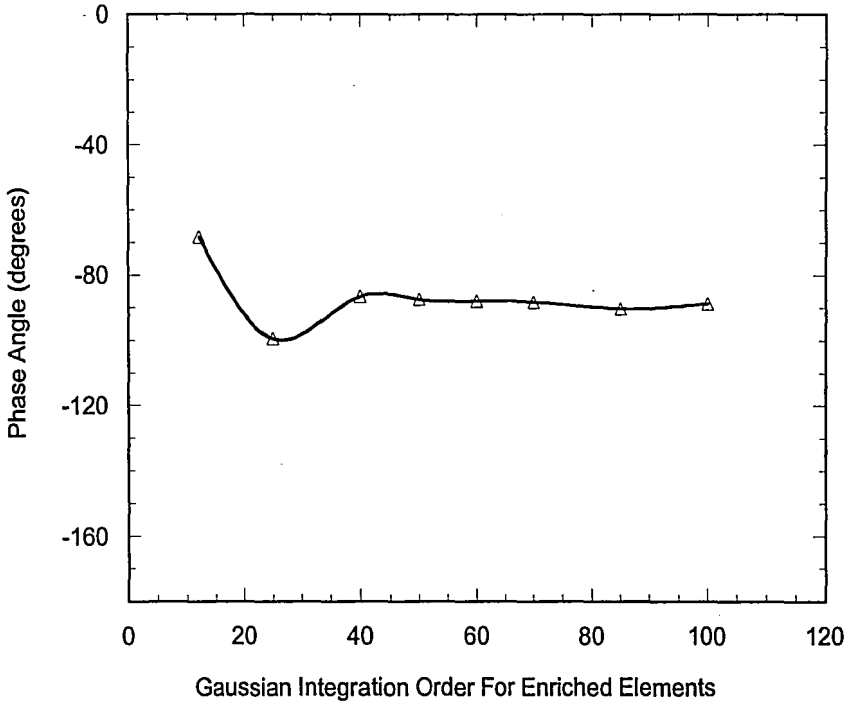


**Figure C.5:** Variation of the phase angle with increasing integration order for enriched elements in a center-cracked flip-chip model with a crack length of 0.10 in.

# REFERENCES

[1] Williams, M. L., The stresses around a fault or crack in dissimilar media, Bull. Seismological Soc. America, Vol.49, pp.199-204, 1959.

[2] Erdogan, F., Stress distribution in a nonhomogeneous elastic plane with cracks, J. Appl. Mech., Vol.30, pp. 232-237, 1963.

[3] England, A. H., A crack between dissimilar media, J. Appl. Mech., Vol.32, pp. 400-402, 1965.

[4] Rice, J. R., Sih, G. C., Plane problems of cracks in dissimilar media, J. Appl. Mech., Vol.32, pp.418-423, 1965.

[5] Atkinson, C., On stress singularities and interfaces in linear elastic fracture mechanics, International Journal of Fracture, Vol.13, pp. 807-820, 1977.

[6] Comninou, M., The interface crack, J. Appl. Mech., Vol.44, pp.631-636, 1977.

[7] Comninou, M., The interface crack in a shear field, J. Appl. Mech., Vol.54, pp. 287-290, 1978.

[8] Comninou, M., An overview of interface cracks, Engineering Fracture Mechanics, Vol.37, pp.197-208, 1990

[9] Van der Zande, H. D., Grootenboer, H. J., Jr, A finite element approach to interface cracks, J. Appl. Mech., Vol.53, pp. 573-578, 1986.

[10] Dattaguru, B., Buchholz, F. G., Venkatesha, K. S., Ramamurthy, T. S., Finite element estimates of strain energy release rate components at the tip of an interface crack, Symp. on Failure and Damage Mechanics: Analysis, Predictions and Innovative Methods at ICES-92, Hong Kong, 1992.

[11] Dattaguru, B., Venkatesha, F. G., Buchholz, F. G., Ramamurthy, T. S., Finite element estimates of strain energy release rate components at the tip of an interface crack under mode I loading, Engineering Fracture Mechanics, Vol.49, pp. 451-463, 1994

[12] Venkatesha, F., Dattaguru, B., Ramamurthy, T. S., Finite element analysis of interface crack with large crack-tip contact zones, Engineering Fracture Mechanics, Vol.54, pp. 847-860, 1996.

[13] Liu, S., Debonding and cracking of microlaminates due to mechanical and hydro-thermal loads, Structural Analysis in Microelectronics and Fiber Optics, ASME 1993, Vol.7, pp. 1-10, 1993.

[14] Irons, B. M., "A Frontal Solution Program", International Journal for Numerical Methods in Engineering, Vol.2, pp 5-32, 1970.

[15] Benzley, S. E., "Representation of Singularities with Isoparametric Finite Elements", International Journal for Numerical Methods in Engineering, Vol.8, pp.537-545, 1974.

[16] Simo, J. C., Taylor, R. L., A perturbed Lagrangian formulation for the finite element solution of contact problems, Com. Meth. in Appl. Mech. and Eng., Vol50, pp. 163-180, 1985.

[17] Simo, J. C., Laursen, T. A., An augmented Lagrangian treatment of contact problems involving friction, Computers & Structures, Vol.42, pp. 97-116, 1992.

[18] Simo, J. C., Wriggers, P., Schweizerhof, K. H., Taylor, R. L., Finite deformation post-buckling analysis involving inelasticity and contact constraints, International Journal for Numerical Methods in Engineering, Vol.23, pp. 779-800, 1986

[19] Yagawa, G., Yoshimura, S., Kitajima, Y., Ueda, H., Fully plastic solutions of semi-elliptical surface cracks (plates and cylindrical shells subjected to tensile loads), Trans. JSME, Part A, Vol.56, pp.1161-1167, 1990.

[20] Hughes, T. J. R., Taylor, R. L., Sackman, J. L., Curnier, A., Kanoknukulchai, W., A finite element method for a class of contact-impact problems, Comp. Methods Appl. Mech. Eng., Vol.8, pp. 249-276, 1976.

[21] Okamoto, N., Nakawaza, M., Finite element incremental contact analysis with various frictional conditions, International Journal for Numerical Methods in Engineering, Vol.14, pp. 337-357, 1979.

[22] Guerra, F. M., Browning, R. V., Comparaison of two slideline methods using ADINA, Computers & Structures, Vol.17, pp. 819-834, 1983.

[23] Bathe, K. J., Chaudhary, A., "A Solution Method For Planar and Axisymmetrci Contact Problems", International Journal for Numerical Methods in Engineering, Vol.21, pp. 65-88, 1985.

[24] White, D. J., Enderby, L. R., Finite element stress analysis of a non-linear problem: A connecting rod eye loaded by means of a pin, J. Strain Anal., Vol.5, pp. 387-395, 1970.

[25] Stadter, J. T., Weiss, R. O., Analysis of contact through finite element gaps, Computers & Structures, Vol.10, pp. 867-873, 1979.

[26] Mazurkiewicz, M., Ostachowicz, W., Theory of the finite element method for elastic contact problems of solid bodies, Computers & Structures, Vol.17, pp. 51-59, 1983.

[27] Zolti, E., A finite element procedure for time dependent contact analysis, Computers & Structures, Vol.17, pp. 555-561, 1984.

[28] Ostachowicz, W., Mixed finite element method for contact problems, Computers & Structures, Vol. 18, pp. 937-945, 1984.

[29] Cheng, J. H., Kikuchi, N., An incremental constitutive relation of unilateral contact friction for large deformation analysis, J. Appl. Mech., Vol.52, pp. 639-648, 1985.

[30] Padovan, J., Moscarello, R., Stafford, J., Tabaddor, F., Pantographing self adaptive gap elements, Computers & Structures, Vol.20, pp.745-758, 1985.

[31] Yagawa, G., Hirayama, H., "A finite element method for contact problems related to fracture mechanics", International Journal for Numerical Methods in Engineering, Vol.29, pp. 2175-2195, 1984.

[32] Mottershead, J. E., Pascoe, S. K., and English, R. G., "A General Finite Element Approach for Contact Stress Analysis", International Journal for Numerical Methods in Engineering, Vol.33, pp. 765-779, 1992.

[33] Man, K. W., Aliabadi, M. H., Rooke, D. P., Analysis of contact friction using the boundary element method, Computational Methods in Contact Mechanics, Elsevier Science Publishers, 1993.

[34] Hertz, H., Uber die Beruhrung festes elasticher Korper, J. reine und angewandte Mathematik, Vol.92, pp.156-171, 1882.

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

[36] Hinton, E., Campbell, J. S., Local and global smoothing of discontinuous finite element functions using a least squares method, International Journal for Numerical Methods in Engineering, Vol.8, pp. 461-480, 1974.

[37] Chan, S. K., Tuba, I. S., "A Finite Element Method for Contact Problems of Solid Bodies - Part I. Theory and Validation", International Journal of Mechanical Sciences, Vol.13, pp. 615-625, 1971.

[38] Sachdeva, T. D., Ramakrishnan, C. V., "A Finite Element Solution for the Two Dimensional Elastic Contact Problems with Friction", International Journal for Numerical Methods in Engineering, Vol.17, pp. 1257-1271, 1981.

[39] Lee, S. S., "A Computational Method for Frictional Contact Problem Using Finite Element Method", International Journal for Numerical Methods in Engineering, Vol.37, pp. 217-228, 1994.

[40] Galego, F. J., Anza, J. J., A mixed finite element model for the elastic contact problem, International Journal for Numerical Methods in Engineering, Vol/28, pp1249-1264, 1989.

# VITA

Yavuz Bertan Bayram was born in Ankara, Turkey on the 10th of March, 1973. After secondary school, he went to Belgium for three years. He entered Middle East Technical University Mining Engineering Department in 1991, from where he transfered to Mechanical Engineering. Graduated in 1995, he served as teaching assistant in the same department for a year after which he came to Lehigh University for his graduate study. He is the second son of Asuman and Mehmet Bayram.

# END OF TITLE