# NUMERICAL MODELLING OF DILATANT ROCK JOINTS

by

M F SNYMAN



A thesis submitted for the degree of Doctor of Philosophy in the Faculty of Engineering.

Department of Civil Engineering

University of Cape Town

October 1991

## ACKNOWLEDGEMENTS

I would like to thank my supervisor, Prof. G.E. Jackson, for his patience, guidance and advice with every part of this thesis.

Thanks is also due to Mark Kelly, Zayed Brown, and Pete Roberts for their assistance.

The emotional support provided by my family, particularly during the past year, is greatly appreciated.

I acknowledge the financial assistance of the South African Council for Scientific and Industrial Research during the course of this study.

## ABSTRACT

Aluminium forms a highly neurotoxic complex with maltol (3-hydroxy-2-methyl-4H-pyran-one).

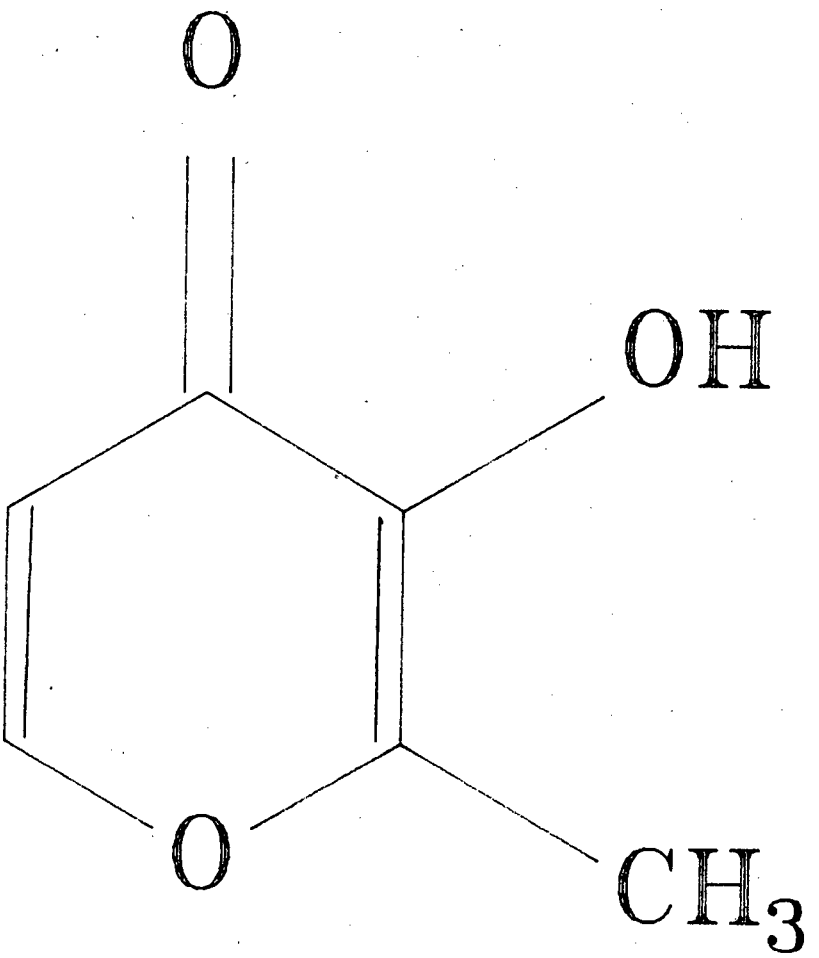
The stability of this complex has been determined using glass-electrode potentiometry.

Owing to the effect on nuclear relaxation behaviour, paramagnetic contrast agents have immense diagnostic potential and have recently received a great deal of attention in the literature. The gadolinium-maltol complex was studied with the view to developing a potential tissue-specific magnetic resonance imaging contrast agent.

Because of the interest in medical applications of radioactive isotopes of group IIIB elements, the indium-maltol complex was studied in order to assess its radiopharmaceutical usefulness.

The major analytical techniques used in this study are potentiometry and high resolution nuclear magnetic resonance spectroscopy.

STRUCTURAL FORMULA OF MALTOL




MALTOL

# CONTENTS

# CHAPTER 1

# INTRODUCTION

## 1.1 ALUMINIUM IN BIOLOGICAL SYSTEMS

Aluminium is the third most abundant element in the earth's crust and is the most abundant metal. Combined with oxygen, aluminium is found in the most common rock and minerals, namely, alumina silicates.

The ionic form of the element is ubiquitous in plant and animal tissues of all kinds and it is found in natural waters everywhere.

Although aluminium is abundant in our environment, the amount of aluminium in the food we eat is actually quite small. It should, however, be noted that certain plants like the *Theaciae* (tea family) are aluminium accumulators. Aluminium is also added as a filler in certain foods, such as cheese, is the active ingredient in various orally administered pharmaceuticals such as antacids, and is an ingredient in baking powder. It is also possible that aluminium enters our food through leaching from aluminium cooking utensils. The actual amount of aluminium ingested varies considerably depending on the individual eating habits.

Aluminium appears to be excreted from the body largely by the kidney. Under ordinary conditions in healthy people, the kidney appears to be able to excrete all the absorbed aluminium. As recently as 1974, the idea that aluminium can act as a human biological toxin was dismissed. It was not until 1976 that researchers began to appreciate the potential toxicity that the trace element might have for humans [1]. A major advance in the study of aluminium has been the development of more

precise methodologies for the measurement of aluminium in biological systems. The various analytical methods for measuring aluminium levels has recently been reviewed [2]. This new technology together with the documentation of the toxicity of aluminium in humans since 1976, provide the impetus for further studies which characterised the pathology, metabolism, and toxicity of this element. At present, the factors that modulate the absorption of aluminium are still poorly understood. Since aluminium would have to be in a soluble form for absorption this suggests that it may be dependent on gastric pH and that absorption occurs primarily in the stomach or proximal duodenum. Aluminium forms stable complexes with fluoride and so fluoride may be an important modulator of aluminium absorption. It has been suggested [3] that vitamin D may increase the absorption of a number of elements such as calcium, lead, magnesium, strontium, beryllium, barium, zinc, cadmium, cesium, and cobalt [4].

Considerable data now exists, which indicate that aluminium, when given to experimental animals, can induce various pathological processes. Among the clinical abnormalities encountered with aluminium loading are anemia, encephalopathy and bone disease. Elevated levels of aluminium have also been found in persons who have died of Alzheimer's disease [5].

Alzheimer's disease is the most common cause of the global deterioration of intellectual functioning in clear consciousness. The syndrome is generally termed dementia. Alzheimer's disease is characterised by the insidious onset and gradual worsening of cognitive defects over a period

of years [6,7]. The neuropathological lesions which characterise this disease, consists of a large number of neurofibrillary tangles in the frontal, temporal and parietal lobes, and prominent cell loss in the nucleus basalis of Meynert. Aluminium is present in the neurofibrillary tangles of patients suffering from the disease and it has been shown that neurofibrillary tangles can be produced in animal brains by the injection of aluminium, but the etiological significance is not clear [5]. A widespread loss of cholinergic enzyme is a consistent finding in patients with Alzheimer's disease. It has been postulated that aluminium may contribute to cholinergic neuronal dysfunction by inhibiting choline transportation. Both the number of neurofibrillary tangles as well as the quantity of cholinergic enzyme loss are related to the severity of the dementia [8,9].

Despite the general agreement on the elevated levels of aluminium involved in the pathogenisis of Alzheimer's disease, there is no agreement on how it is delivered to the brain.

Finnegan, Rettig and Orvig [10] suggest that the coordination chemistry of aluminium in aqueous solutions may be directly linked to the problem of aluminium in Alzheimer's disease through the use of ligands which would produce neutral six-coordinate complexes of some stability at physiological pH. The properties of solubility and stability in $H_2O$ are necessary if the complex is to be delivered to the blood brain barrier.

The blood brain barrier (BBB) is a highly selective barrier that prevents the passage of many substances from the blood into

extracellular fluid of the brain or into the cells themselves, and *vice versa*. The BBB allows the transfer of two kinds of substances: highly lipid soluble molecules and water soluble molecules that are transported by a carrier. The barrier is based on some unique characteristics of the capillaries of the brain. The cells of the capillary wall in the brain are tightly joined together, having no "split pores" between the cells. A basement membrane and a fatty, footlike sheath, called a glial foot, of adjacent nerve cells further separate any diffusing species from extracellular brain fluid. All substances must traverse the capillary cell membranes as well as the fatty tissue lining of the glial foot to cross the BBB. For a complex to pass the BBB, the requirements of electroneutrality and a molecular weight of about 400 daltons, or less, must be satisfied [11]. Finnegan *et al.* [10] examined the low molecular weight neutral compounds of maltol with aluminium.

Maltol (3-hydroxy-2-methyl-4H-pyran-one) is a natural product which may be obtained by the alkaline hydrolysis of streptomycin [12] and is used as a flavouring agent in certain foodstuffs. The investigation of Finnegan *et al.* [10] of the Al-maltol complex yielded spectroscopic data which is consistent with bidentate, monoanionic ligation of the metal ion. The analytical and mass-spectral results show a tris-ligand complex. The complex was shown to have a remarkable window of stability to hydrolysis (pH 4-9). Preliminary studies show the Al-maltol complex to be unusually neurotoxic suggesting that the complex remains intact delivering $Al^{3+}$ successfully into the brain cells.

Based on the findings of this study, it was decided to examine the indium and gadolinium analogs of this complex with the view of developing potential imaging agents. Since gadolinium, indium and aluminium are similar in various chemical aspects, such as ionization potentials and the fact that $M^{3+}$ is the prevalent aquo species of all three metals, it is likely that the gadolinium and indium maltol complexes will be capable of crossing the BBB. This will make them potential imaging agents specifically for the brain.

## 1.2 MAGNETIC RESONANCE IMAGING

In magnetic resonance imaging, (MRI), the images are generated by differences in proton density and relaxation times $T_1$ and $T_2$ between neighbouring tissue.

Depending on the particular radiofrequency pulse sequence used, image intensity may be weighted towards proton density, spin-lattice relaxation, $T_1$, or spin-spin relaxation, $T_2$. Although the images produced look similar to those produced by computed tomographic scanning, they are based on variables that have little to do with electron density.

MRI provides unique information regarding tissue characterization and status of organ function or blood flow.

The inherent difference in the observed nuclear resonance signals from various tissues and fluids may be enhanced by the use of contrast agents. These pharmaceuticals produce MRI images which provide more diagnostic information by altering $T_1$ and $T_2$. MRI contrast may be administered by intravascular injection, ventilation, oral ingestion or subarachnoid instillation.

A class of compounds which have received the most attention as potential contrast agents are the complexes of paramagnetic transition and lanthanide ions. Such contrast agents are not observed directly on the images; rather, their magnetic effects on neighbouring nuclei are the means of contrast enhancement. Alternatively, enhancement can be achieved by manipulation of tissue viscosity or temperature but since these are not clinically feasible, the use of paramagnetic pharmaceuticals appears to be more advantageous. We shall thus limit our discussion to enhancement by the use of paramagnetic contrast agents.

A paramagnetic substance may be defined as one that is attracted toward and aligns with the stronger portion of a magnetic field. A paramagnetic substance has its own magnetic moment and aligns in an external magnetic field but loses this alignment and becomes randomly oriented when the field is removed. Paramagnetic substances possessing such magnetic fields act as "relaxation centres" for other nuclei in

their microenvironment by altering relaxation times of neighbouring nuclei via dipolar interactions. There are numerous paramagnetic substances but only a few have promise as clinical contrast agents.

Ions from the first row transition series which include, iron, manganese, cobalt, nickel and copper contain unpaired electrons. There are five 3d orbitals capable of accommodating a total of 10 electrons. For the transition series, in high spin complexes, the first five electrons fill the 3d orbitals without pairing and it is not until the sixth electron is reached (for $Fe^{2+}$ and $Co^{3+}$) that electrons of opposite spin begin pairing. Maximum spin is achieved in the series with $Fe^{3+}$ and $Mn^{2+}$.

The first human MRI study using a paramagnetic agent was conducted by Young *et al* [13] using ferric chloride to enhance the gastrointestinal tracts. Besides $Fe^{3+}$ from the first row transition series, $Mn^{2+}$ has been used extensively as a paramagnetic probe for *in vitro* biochemical studies [14]. The lanthanide series also provides paramagnetic ions, which contain unpaired electrons in 4f orbitals. There are seven 4f orbitals and a maximum spin is achieved when ions in this group contain seven electrons with unpaired spin. This is the configuration of $Gd^{3+}$ which has a spin quantum number of 7/2. Studies of gadolinium suggest that the ion does significantly enhance the signal of certain tissue [15]. The potential disadvantages of using inorganic cations as MRI contrast agents are their relative lack of chemical versatility, delayed biological excretion times, and toxicity [16]. It is thus unlikely that these ions will achieve wide use as *in vivo* contrast agents but that

nontoxic complexes of these ions be developed which are rapidly excreted and have strong relaxation effects. The only contrast agent currently undergoing clinical trials is $[Gd(DTPA)(H_2O)]^{2-}$.

### 1.2.1 General requirements for metal complexes as MRI contrast agents

Besides the standard pharmaceutical features, the requirements relevant to metal complex based agents may be classified into the following categories;

i) relaxivity

ii) *in vivo* stability

iii) specific *in vivo* distribution

iv) toxicity

### 1.2.1.1 Relaxivity

The efficiency with which the complex enhances proton relaxation rates is described by the term relaxivity, which must be sufficient to significantly increase relaxation rates of target tissues at a nontoxic dose.

Various theoretical approaches exist for the understanding of NMR relaxation properties on nuclei in the presence of unpaired electron spin [17-20]. This knowledge base is the starting point for discussing aspects of relaxivity.

The addition of paramagnetic solute causes an increase in the longitudinal and transverse relaxation rates ($1/T_1$ and $1/T_2$) of solvent nuclei.

The diamagnetic and paramagnetic contributions to the relaxation rates of such solutions are additive and are given by the following equation:

$$(1/T_i)_{obsd} = (1/T_i)_d + (1/T_i)_p \qquad i = 1,2$$

where $(1/T_i)_{obsd}$ is the observed solvent relaxation rate in the presence of a paramagnetic species, and $(1/T_i)_d$ represents the (diamagnetic) solvent relaxation rate in the absence of a paramagnetic species, and $(1/T_i)_p$ is the additional paramagnetic contribution.

The large and fluctuating local magnetic field in the vicinity of a paramagnetic centre provides the additional relaxation pathway for solvent nuclei. Since these fields fall off rapidly with distance, random translational diffusion of solvent molecules and the complex as well as specific chemical interactions which bring solvent molecules near the metal are important in transmitting the paramagnetic effect.

The contributions for water proton relaxivity can be classified in three distinct types of interactions, schematically indicated below;

In case **A** the water molecules bind to the metal ion and exchange with the bulk solvent. The relaxation mechanism which results from this type of exchange is loosely referred to as "inner sphere relaxation".

In case **B** the water molecules are H-bonded in the second coordination sphere. Due to a lack of understanding of this second coordination sphere interaction, investigators do not distinguish between this relaxation mechanism and that due to translational diffusion past the chelate (case **C**), referring simply to "outer sphere relaxation". Examination of complexes which have no coordinated water molecules

(q = 0) indicate that outer sphere relaxivity is significant. The total relaxivity of a paramagnetic agent is generally given by;

$$(1/T_i)_p = (1/T_i)_{\text{inner sphere}} + (1/T_i)_{\text{outer sphere}}$$

where i = 1,2.

There are various physical and chemical parameters important in relaxivity which may be optimized to increase the efficiency of a paramagnetic agent. The discussion of these parameters are based on the argument outlined by Lauffer [21].

i) **Number of coordinated water molecules, q**

The presence of at least one coordinated water molecule (inner sphere relaxation) is important in attaining high relaxivities. While outer sphere relaxivities may be enhanced to some degree upon immobilization, it is limited by the very rapid translational diffusion of water or transient hydrogen bond lifetimes.

Table 1.1 below indicates the relation between longitudinal relaxivities, $R_1$, and the number of coordinated water molecules, q, for low molecular weight complexes [21].

Table 1.1 Longitudinal relaxivities of various metal complexes

| Complex | q | $R_1$ $mM^{-1}s^{-1}$ | freq MHz | temp $^0C$ |
|---|---|---|---|---|
| $Gd^{3+}$ | | | | |
| aquo ion | 8,9 | 34.3 | 0.02 | 5 |
| DTPA | 1 | 7.7 | 0.02 | 25 |
| EDTA | 2,3 | 25 | 0.02 | 5 |
| $Mn^{2+}$ | | | | |
| aquo ion | 6 | 44 | 0.02 | 35 |
| DTPA | 0 | 3.4 | 0.02 | 5 |
| EDTA | 1 | 5.6 | 0.02 | 25 |
| $Fe^{3+}$ | | | | |
| aquo ion | 6 | 17 | 0.02 | 35 |
| DTPA | 0 | 0.92 | 0.02 | 37 |
| EDTA | 1 | 25 | 0.02 | 37 |

ii) **Distance between water protons and the unpaired electron spin, r**

The $r^{-6}$ dependence in dipolar interactions presents the opportunity to increase relaxivity by (a) chemically inducing an orientation of bound water molecules such that the protons are closer to the metal chelate centre or unpaired spin density; or (b) delocalizing the unpaired spin density towards the water through atomic or molecular orbitals of the metal ion, the chelate ligand, or the bound water itself.

### iii) Rotational correlation time, $\tau_R$

For metal ions with long electron spin relaxation times ($T_{1e}$) alteration of the rotational tumbling $\tau_R$ is the single most important source of relaxivity enhancement.

Three basic strategies exist to reduce the rotational mobility of metal complexes *in vivo*:

a. Distribution into a tissue or tissue compartment with high microviscosity.

b. Covalent attachment of the compound to a larger molecule such as a protein or antibody.

c. Non covalent binding of the complex in tissue to macromolecules.

Table 1.2 gives some relaxivities of protein complexes. As can be seen, if free exchange of coordinated water is possible, there is an enormous increase in relaxivity.

**Table 1.2.** Selected longitudinal relaxivities ($R_1$) for protein - metal ion complexes and for bovine serum albumin (BSA) covalently labelled with metal chelates.

| Complex | $R_1$ $M^{-1}s^{-1}$ | Freq. MHz | Temp. $^0C$ |
|---|---|---|---|
| $Gd^{3+}$ | | | |
| glutamine synthetase | 148 | 22.5 | 25 |
| immunoglobulin | 112 | 20 | 19 |
| concanavalin A | 60 | 20 | 25 |
| BSA | 72 | 24.3 | 30 |
| $Mn^{2+}$ | | | |
| pyruvate kinase | 275 | 20 | 25 |
| concanavalin A | 96 | 20 | 25 |
| carboxypeptidase | 43 | 20 | 25 |

### iv. Electron spin relaxation time, $T_{1e}$

The choice of $Gd^{3+}$ as one of the optimal relaxation agents stems from the ions long $T_{1e}$ and large magnetic moment.

In general increasing $T_{1e}$ yields higher relaxivities. This effect is limited by the correlation time for internal motion or by $\tau_R$ and $\tau_M$.

### v. Residence lifetime of coordinated waters, $\tau_M$

Chemical exchange of coordinated water molecules between the metal ion and the bulk solvent is necessary for the relaxation effect of the metal ion to be carried across to the bulk water. The rate of chemical exchange is given by $\tau_M$, the mean lifetime of a water molecule in the

coordination sphere of the metal. It is for this reason that $Cr^{3+}$, which is substitutionally inert, is such a poor relaxation agent.

### 1.2.1.2 Stability and Toxicity of Contrast Agents

The required dose of metal-complex for MRI greatly exceeds that used in radioscintigraphy. There thus exists a need to develop safe chelates and the chemist can contribute tremendously by elucidating the dissociation mechanism of such chelates in biological systems.

Toxic effects of metal complexes can arise from the intact complex or by the free ligand or free metal ion released by dissociation. The dissociation of a complex generally leads to a higher degree of toxicity. An example of how toxicity and *in vivo* stability depend on the chelate ligand is the comparison between $[Gd(EDTA)(H_2O)_n]^-$ and $[Gd(DTPA)(H_2O)_n]^{2-}$. The latter is a very stable complex, excreted intact readily by the kidneys, having a low degree of toxicity. The former complex has a toxicity comparable to $GdCl_3$ due to the quantitative *in vivo* dissociation of $[Gd(EDTA)_n]^-$ yielding the free metal ion which can then coordinate to macromolecules and cell membranes, altering the dynamic equilibria necessary to sustain life.

Other possible mechanisms of chelate toxicity include enzyme inhibition, alteration of membrane potentials, and nonspecific protein conformational effect.

The requirement of metal complex stability is essentially a kinetic requirement - stability is only required for the residence time in the body. The spherical electronic distribution of $Gd^{3+}$, $Mn^{2+}$ and $Fe^{3+}$, which lead to high relaxivities, is detrimental to complex stability. The lack of ligand field stabilization energy in complexes of these metal ions lead to labile metal-ligand bonds. Kinetic stability must therefore be derived from structure of an appropriate multidentate ligand.

### 1.2.2 Potential Clinical Applications of Contrast Agents

Contrast enhancement may prove to be increasingly valuable by differentiating between magnetically similar but histologically dissimilar tissue - i.e. by identification of an isointense abnormal tissue foci surrounded by normal tissue.

Another application of contrast agents is the direct assessment of tissue function. For example, the renal excretion of $[Gd(DTPA)(H_2O)]^{2-}$ and $[Gd(EDTA)(H_2O)]^{-}$ yield the application of imaging the kidney for structural and functional information [22,23].

Disease of the central nervous system which disrupt the blood brain barrier may not be identifiable on nonenhanced magnetic resonance images

but the regions of functional abnormality should become more evident if a contrast agent passes into areas where the blood brain barrier has broken down. Hepatobilary agents are another class of contrast agents which are becoming increasingly important since they give an indication of the state of hepatocytes of the liver.

In comparing the development of MRI contrast agents with that of radiopharmaceuticals, it is apparent that the former will require a great deal more characterization due to the complexity inherent in NMR relaxation phenomena and the higher dose requirements. Besides the fact that MRI is a welcome aid to the clinical practice of medicine, the two important factors that will catalyze the development of safe contrast agents are (a) the technique uses nonionizing radiation and (b) there appear to be no biological hazardous effects.

## 1.3 RADIOPHARMACEUTICALS

The procedures of nuclear medicine are divided into two increasingly divergent types:

(i) *in vitro* assays including radioimmuno-assays and saturation analyses.

(ii) *in vivo* counting and imaging procedures requiring gamma-emitting agents.

It is the procedure of the latter type which are of interest to us in general and more specifically the development of suitable gamma-emitting complexes for use as radioimaging agents.

### 1.3.1. TECHNETIUM

The number of g-emitting radionuclides which have been incorporated into compounds or complexes for external scintillation detection is small considering that about 100 g-emitting radionuclides have been known for many years.Today most radiopharmaceuticals employed in the field of nuclear medicine to visualize tissue, anatomical structures, and metabolic disorders are labeled with the element technetium (Tc). The isotope of technetium utilized is the metastable $^{99m}Tc$. This isotope is short lived ($t_{\frac{1}{2}}$=6.02 h) with a predominant single-photon g emission having an energy of 140.6 KeV. After intravenous administration, $^{99m}Tc$ radiopharmaceuticals localize in specific target tissues, which can be imaged using sodium iodide crystal cameras.

The most stable form of technetium is the pertechnetate ion $TcO_4^-$. Compounds of technetium are known in all valency states from +7 to -1 but in water the most stable forms are $TcO_4^-$ (+7) and the insoluble $TcO_2$(+4). Most of Tc-based pharmaceuticals contain the radionuclide in a reduced state.

The following examples are some of the uses of Tc-radiopharmaceuticals in nuclear medicine:

**BRAIN IMAGING.**

$^{99m}Tc$ as sodium pertechnetate is most commonly used. Sodium or potassium perchlorate is given to inhibit its uptake by choroid pexis, thyroid and salivary glands.

**THYROID IMAGING**

The trapping of $^{99m}Tc$ as sodium pertechnetate by the thyroid is useful for assessing both the function and anatomy of the gland.

**LIVER IMAGING.**

Sulphur or tin colloids are the agents of choice.

**HEPATOBILIARY IMAGING.**

Lidocaine derivations of iminodiacetic acid form stable chelates with stannous chloride. Adding $^{99m}TcO_4^-$ results in a radiopharmaceutical of rapid liver uptake and excretion into the bile at high concentration which is ideal for hepatobilary scanning.

**KIDNEY IMAGING.**

$^{99m}Tc$-stannous-DTPA is used for the morphological and functional investigation of the kidney.

**LUNG IMAGING.**

Either albumin microspheres or macroaggregates formed by the controlled denaturation of human serum albumin and labeled by the addition of $^{99m}TcO_4^-$ is used.

In addition to the examples discussed, $^{99m}Tc$ radiopharmaceuticals are extensively used for cardiovascular imaging as well as bone imaging.

Because of the practicality of this radionuclide, $^{99m}Tc$ is likely to remain the workhorse of diagnostic nuclear medicine. Even though there are many radionuclides with a wide range of energies available for use in radiopharmaceuticals, it is unlikely that a radionuclide will emerge which is superior to $^{99m}Tc$ in the way in which this element predominates diagnostic imaging. Instead, if a particular application demands certain decay properties, the radionuclide which will be used will be the one for which appropriate chemistry can be developed and which can be produced and distributed most economically.

## 1.3.2. GALLIUM AND INDIUM RADIOPHARMACEUTICALS

With the current interest in the designing of new radiopharmaceuticals for imaging, considerable attention is being directed to the radioactive isotopes of indium and its sister IIIB element gallium. These elements offer both cyclotron-produced and generator available radionuclides which have decay characteristics and half-lives attractive for nuclear medicine application.

Although indium and gallium are not transition elements, there are numerous aspects of chemistry which resemble those of iron. The ionization potentials, ionic radii, and coordination number of the aquo +3 cation are similar for all 3 metals. This complicates the design of radiopharmaceuticals labeled with gallium and indium since these metals form strong complexes with the plasma protein, transferrin. A gallium or indium labeled radiopharmaceutical will exchange its metal to transferrin unless the metal is bound with a stability constant greater than $10^{20}$.

Despite this lack of thermodynamic stability with respect to exchange with transferrin, the rate at which equilibrium is reached is often slow compared to the biological event. An example is the fast glomerular filtration rate of indium-DTPA which is much more rapid than the rate of exchange with transferrin.

An obvious requirement that gallium and indium radiopharmaceuticals must meet is the ability to resist hydrolysis of the metal in aqueous solutions at physiological pH (hydrolysis results in colloidal precipitation). This can be achieved if the metal complexes are thermodynamically stable to hydrolysis or if the kinetics of the ligand-to-hydroxide exchange process is slow relative to the nuclear medicine procedure Examples of the latter circumstances are complexes of EDTA and DTPA which have found clinical use in the detection of brain tumours and the study of renal function despite their relatively low thermodynamic stability.

In addition to these chemical requirements, certain criteria can be defined which may be used as a guide in designing radiopharmaceuticals for diagnostic imaging.

1. The intensity of the signal generated by the radionuclide should be reproducible from one examination to the next and must be concentration dependent.

2. The metal chelate should be chemically versatile so it can be bound to other compounds as a biological probe and thus permit selective tissue targeting. There has, for example, been major efforts in the development of bifunctional chelates which can form a bridge between the metal complex and a protein or peptide. The labelling of blood components with indium-8-hydroxy-quinolates has also been used to great advantage since transferrin has no access to such labelled species.

3. The radionuclide should be easily obtained and have suitable half-lives and decay characteristics to be practical for nuclear medicine applications.

4. The radiopharmaceutical should not be reactive in vivo and should be nontoxic in diagnostic doses.

## REFERENCES


1. Alfrey A.C., Le Gendre G.R., Kaehny W.D., *N. Engl. J. Med.*, **294** (1976) 184.
2. Supplement to Kidney International **29** (1986) s1.
3. Worker N.A., Migicovsky B.B., *J. Nutr.*, **75** (1961) 222.
4. Long J.F., Nagode L.A., Kindig O., Liss L., *Neurotoxicol.*, **1** (1980) 111.
5. Perl D.P., Brody A.R., *Science*, **208** (1980) 297.
6. Larsson T., Sjogren T., Jacobson G., *Act Psychiatry Scand.*, **167** (1963) 1.
7. Sjogren T., Sjogren E., Lindgren A., *Acta Psychiatry Neurol. Scand.*, **82** (1952) 68.
8. Blessed G., Tomlinson B.E., Roth M., *Br. J. Psychiatry*, **114** (1968) 797.
9. Perry E.K., Tomlinson B.E., *et al.*, *Br. Med. J.*, **2** (1978) 1457.
10. Finnegan M.M., Rettig S.J., Orvig C., *J. Am. Chem. Soc.*, **108** (1986) 5033.
11. Levin V.A., *J. Med. Chem.*, **23** (1980) 682.
12. Schenck J.R., Spielman M.A., *J Am. Chem. Soc.*, **67**, (1945) 2276.
13. Young J.R., Clarke G.J., Gailes D.R., *et al.*, *Comput. Tomogr.*, **5** (1981) 534.
14. Eisinger J., Shulman R.G., Szymanski B.M., *J. Chem. Phys.*, **36** (1962) 1721.
15. Carr D.H., Graif M., Niendorf H.F., *et al.*, *Clin. Rad.*, **37** (1986) 347.

16. Christiensen H.E. (ed.), "*Registry of Toxic Effects of Chemical Substances*", U.S.A. Department Health and Welfare, NIOSH, Rockville, 1975.

17. Koenig S.H., *J.Magn. Reson.*, **31** (1978) 1.

18. Koenig S.H., Brown R.D., *Magn. Reson. Med.*, **1** (1984) 478.

19. Koenig S.H., Brown R.D., *J. Magn. Reson.*, **61** (1985) 426.

20. Burton D.R., Forsen S., Karlstrom G., Dwek R.A., *Prog. Nucl. Magn. Reson. Spectrosc.*, **13** (1979) 1.

21. Lauffer R.B., *Chem. Rev.*, (1987) 901.

22. Wolf G.L., Fobben E.S., *Invest. Radiol.*, **19** (1984) 324.

23. Brasch R.C., Weinmann H., Wesbey G.E., *Am. J. Roentgenol.*, **142** (1984) 625.

# CHAPTER 2

# POTENTIOMETRY

## 2.1 INTRODUCTION.

Wherever metal ions and ligands are present, equilibria between them will be established. A knowledge of the chemical nature and concentration of each of the species present at equilibrium has many applications. Accordingly, the applications of stability constants that refer to the interactions of metal ions and ligands in solution are ubiquitous. Stability constants have found application in such diverse fields as electrochemistry, analytical chemistry, geochemistry, oceanography, photography and medicine.

In the medical field stability constants can be exploited to study the pathology or treatment of diseases such as Wilson's disease. This disease results from the accumulation of copper in the brain, liver and kidneys and is generally treated with $Na_2CaEDTA$.

Stability constants can also be used to evaluate the effectiveness of metal complexes which are administered as therapeutic pharmaceuticals. It is important to be able to predict the possibility of acute and chronic toxic effects. Toxic effects from metal complexes can arise from free ligand or free metal upon dissociation of the complex or from the metal complex itself. The development of computer models such as the blood plasma model [1] has greatly assisted this field of research. The blood plasma model performs speciation calculations for specific

metal-ligand systems taking into account the formation constants of all other naturally occurring ligands present in blood plasma.

There are various experimental methods available for the determination of stability constants. These include potentiometric, spectroscopic, distribution, electrochemical, and calorimetric methods as well as techniques such as viscometry, and reaction kinetics [2]. Because of its high accuracy and precision, potentiometry is still the preferred method of determination.

## 2.2 THEORY.

For the purpose of this thesis potentiometry is defined as the measurement of electromotive force (e.m.f.) of a reversible cell.

For the reaction

$$M + L \rightleftharpoons ML \qquad \ldots\ldots\ldots \quad (1)$$

the thermodynamic stability constant $K_{ML}$ is given by the equation

$$K_{ML} = \{ML\}/\{M\}\{L\} \quad \ldots\ldots\ldots \quad (2)$$

where the values within the brackets {} refer to the activities of the species concerned.

In many cases more than one ligand can coordinate to the metal and hence the stepwise formation of $ML_2$, $ML_3$ and so on occurs. In general it is not possible to first form $ML_n$ without first forming $ML_{n-1}$.

In the case where more than two metal ligand complexes are formed, several stepwise stability constants may be defined of the general form

$$K_n = \{ML_n\}/\{ML_{n-1}\}\{L\} \quad \ldots\ldots\ldots \quad (3)$$

The overall stability constant is denoted by $\beta$

$$\beta_n = \{ML_n\}/\{M\}\{L\}^n \quad \ldots\ldots\ldots \quad (4)$$

Since activity coefficients are difficult to measure and theoretical calculations based upon the Debye-Hückel approach are of limited accuracy, it becomes necessary to ensure that the term $\gamma_{ML}/\gamma_M\gamma_L$ remains constant. This is effectively done by (a) having a large excess of background electrolyte and (b) by using low concentrations of metal and ligand so that any change in their concentrations as a result of complex formation does not change the overall ionic strength of the medium significantly.

If $\gamma_{ML}/\gamma_M\gamma_L$ remains constant a stoicheiometric stability constant, K, can be defined as:-

$$K = [ML]/[M][L] \quad \ldots\ldots\ldots \quad (6)$$

The possibility of the formation of protonated, hydroxo-, and oligonuclear species has to be taken into account when calculating overall stability constants. This is done by the introduction of subscripts e.g. $M_pL_qH_r$ where

$$\beta_{pqr} = [M_pL_qH_r]/[M]^p[L]^q[H]^r \quad \ldots\ldots\ldots \quad (7)$$

In order to evaluate stability constants it is often necessary to define secondary functions which aid in the interpretation of data. Two such functions are the formation function, $\bar{Z}$, [3] and the deprotonation function, $\bar{Q}$.

The proton formation function, $\bar{Z}_H$, is defined as follows:-

$$\bar{Z}_H = (T_H - H + OH)/T_L$$

where $T_H$ = Total hydrogen ion concentration

$T_L$ = Total ligand concentration

$OH = K_w/H$

H = free hydrogen ion concentration

The metal formation function, $\bar{Z}_M$, is defined as follows :-

$$\bar{Z}_M = (T_L - A(1+\Sigma_n \beta_{LHn} H^n))/T_M$$

where $T_M$ = Total metal concentration

$$A = (T_H - H + OH)/(\Sigma_n \beta_{LHn} H)$$

The above $\bar{Z}_H$ and $\bar{Z}_M$ functions are only defined for simple stepwise mononuclear complex formation. While the functions are not valid for systems involving polynuclear or hydroxy species, the shape of the function can aid in the interpretation of the data. Figure 2.1 shows a theoretical $\bar{Z}$ curve for the system ML, MLH, and MLOH. The hump at high pA (low pH) is indicative of the presence of protonated species while the curl in the low pA region is indicative of hydroxy species formation.

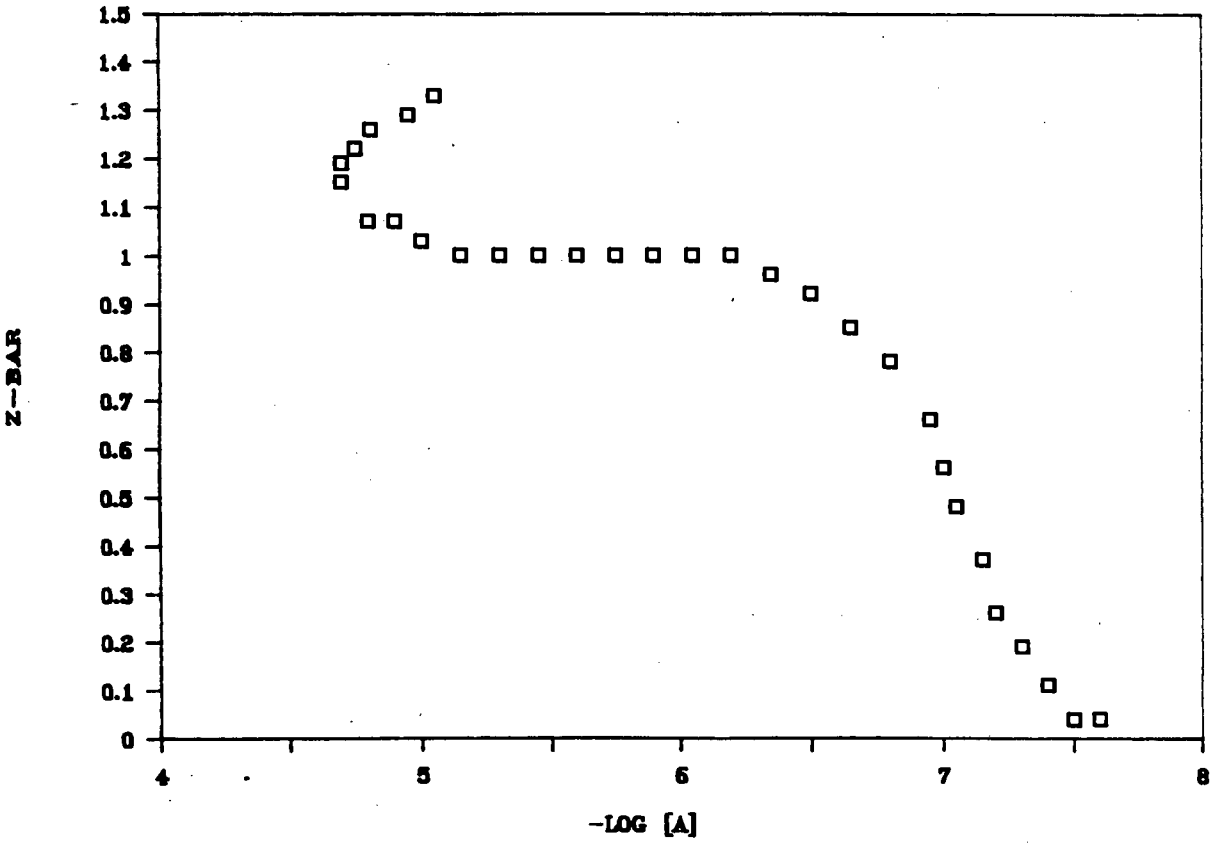



Figure 2.1 Theoretical $\bar{Z}$-curve for the system ML,MLH,MLOH.

Unlike the formation function the deprotonation function $\bar{Q}$ is valid for any complex stoichiometry. It can be regarded as the average number of protons released as a result of complexation, per metal ion and is defined as

$$\bar{Q} = (T_H^* - T_H)/T_M$$

where $T_H^*$ is the calculated total concentration of protons in the system at the observed pH ignoring the presence of all metal complexes.

For binary systems, a formation function is defined for the ligand subsystem

$$n = (T_H^* - H + OH)/T_L^r$$

An average proton stoichiometric coefficient can thus be defined for any given M and L stoichiometry (p and q).

$$r = qxn - \bar{Q}xp$$

For ternary systems, the average proton stoichiometric coefficient can be calculated given any M, L and X stoichiometry ($M_pL_qX_sH_r$).

$$r = qxn_L + sxn_X - \bar{Q}xp$$

In the calculations for proton stoichiometric coefficients it is assumed that the complexes are completely formed.

# ABSTRACT

The thesis considers a plane joint or interface element suitable for implementation into a standard non-linear finite element code. The element is intended to model discontinuities with rough contact surfaces, such as rock joints, where dilatant behaviour is present. Of particular concern is the formulation of a constitutive model which fully caters for all possible histories of opening, closing and sliding accompanied by dilation, in any direction.

The non-linear incremental constitutive equations are formulated in a manner appropriate for a backward difference discretisation in time along the path of loading. The advantage of such an approach is that no essential distinction need be drawn between opening, closing and sliding. Further, a convenient formulation of the constitutive equations is facilitated by representing the different contact conditions in relative displacement space. The state diagram in relative displacement space, however, changes from one time-step to the next, and evolution equations for the updating must be formulated.

These concepts are illustrated for three rock joint models; a dilatant Coulomb model, a sawtooth asperity model and a logarithmic spiral model. The models are based on a penalty formulation to enforce the contact constraints, and explicit equations for the tangent stiffness matrix and for the corrector step of the standard Newton-Raphson iterative algorithm are derived. These equations have been implemented as a user element into the finite element code ABAQUS. Six examples, mainly in the field of rock mechanics, are presented to illustrate the predictions of the formulation.

# DEDICATION

I would like to dedicate this thesis to my parents for their continuous interest in my work, and to my wife Antoinette, for her love, as well as endless patience and support during my postgraduate studies.

# ACKNOWLEDGMENTS

I acknowledge and greatly appreciate the help of the following people and organisations:

My supervisor, professor J B Martin, who was a great source of encouragement, who sacrificed valuable time, and who splendidly guided my research.

The Foundation for Research and Development (FRD) and The Chamber of Mines of South Africa Research Organisation (COMRO) for financial assistance.

My colleagues at the Centre for Research in Computational and Applied Mechanics (CERECAM), in particular Les Rencontre and Tony Pretorius, for helpful discussions.

Tony Pretorius for proofreading the final draft.

Finally, my family for their support and interest in my studies.

# TABLE OF CONTENTS

# NOMENCLATURE

This is a list of the symbols used in the main text of this thesis.

## Uppercase symbols

| | |
|---|---|
| $D_{ij}$ | Components of tangent stiffness matrix |
| $E$ | Young's modulus |
| $J$ | Jacobian |
| $K_t$ | Elastic constant normal to joint face |
| $K_s$ | Elastic constant tangent to joint face |
| $N$ | Normal force |
| $N_i$ | Shape function components |
| $R$ | Radial component of polar coordinates $(R, \theta)$ |
| $S$ | Shear force |

## Bold uppercase symbols

| | |
|---|---|
| $\mathbf{D}$ | Tangent stiffness matrix |
| $\mathbf{F}$ | Internal force vector |
| $\mathbf{K}$ | Stiffness matrix |
| ${}^{e}\mathbf{F}$ | Element internal force vector |
| ${}^{e}\mathbf{K}$ | Element stiffness matrix |
| $\mathbf{N}$ | Shape function matrix |
| $\mathbf{R}$ | External force vector |

## Lowercase symbols

| | |
|---|---|
| $l$ | Interface arc-length |

| | |
|---|---|
| $t$ | time |
| $u$ | Horizontal displacement |
| $v$ | Vertical displacement |
| $x$ | Global coordinate |
| $y$ | Global coordinate |

## Bold lowercase symbols

| | |
|---|---|
| $\mathbf{i}$ | Unit vector along $x$ coordinate axis |
| $\mathbf{j}$ | Unit vector along $y$ coordinate axis |
| $\mathbf{n}$ | Unit vector normal to joint plane |
| $\mathbf{s}$ | Unit vector tangential to joint plane |
| $\mathbf{u}$ | Displacement vector |

## Greek Symbols

| | |
|---|---|
| $\gamma$ | Relative shear deformation |
| $\gamma^e$ | Elastic component of shear deformation |
| $\gamma^p$ | Inelastic component of shear deformation |
| $\gamma^o$ | Shear component of apex position |
| $\delta$ | Relative normal deformation |
| $\delta^e$ | Elastic component of normal deformation |
| $\delta^p$ | Inelastic component of normal deformation |
| $\delta^o$ | Normal component of apex position |
| $\delta^m$ | Maximum dilation/height of asperities |
| $\zeta$ | Dimensionless parameter |
| $\eta$ | Isoparametric reference coordinate |
| $\theta$ | Angular component of polar coordinates $(R, \theta)$ |
| $\mu$ | Dilation angle |
| $\mu^o$ | Initial dilation angle |
| $\nu$ | Poisson's ratio |
| $\xi$ | Isoparametric reference coordinate |
| $\phi$ | Friction angle |
| $\psi$ | Angle in relative displacement space |

## Bold Greek symbols

| | |
|---|---|
| $\boldsymbol{\Theta}$ | Transformation matrix |
| $\boldsymbol{\sigma}$ | Stress vector |
| $\boldsymbol{\Phi}$ | Force residual |

## Superscripts

| | |
|---|---|
| $e$ | Elastic |
| $(i)$ | Iteration or increment |
| $p$ | Inelastic |
| $T$ | Transpose of a vector or matrix |

## Subscripts

| | |
|---|---|
| $b$ | Bottom |
| $n$ | Time step |
| $t$ | Top |

## Special Symbols

| | |
|---|---|
| $\partial$ | Partial differentiation with respect to |
| d | Differentiation with respect to |
| $\cdot$ | Differentiation with respect to time |
| $\Delta$ | Increment in |
| sgn( ) | Sign of ( ) |
| | if ( ) $\geq 0$ then sgn( ) = +1 |
| | ( ) $< 0$ sgn( ) = -1 |

# CHAPTER 1
# INTRODUCTION

Mining for gold in South Africa takes place at depths of up to 3500 m below the surface where the gravitational stress can exceed 100 MPa. The sedimentary gold reef occurs in strong quartzite rock which is separated at approximately 1 m intervals with thin almost horizontal shale layers called parting or bedding planes. Dykes and faults are common. The gold bearing reef is excavated by the longwall technique, in horizontal tabular stopes approximately 1 m high. The removal of this highly stressed rock causes a complicated pattern of fracturing in the rock mass surrounding the stope. The fracture zone is susceptible to rockfalls and rockbursts, increasing the hazard to miners. Of these fractures, the vertical shear fractures are the most important, and, together with the parting planes, divide the rock mass into blocks. The distribution of stress and deformation are influenced by deformation occurring along these discontinuities.

The work presented in this thesis forms part of an ongoing research programme, directed by the Chamber of Mines of South Africa Research Organisation (COMRO), which seeks an understanding into the behaviour of the zone of fractured rock surrounding these deep excavations. Such understanding is vital if safer and more productive mine layouts, extraction sequences, as well as support methods are to be developed.

The thesis is specifically aimed at establishing expressions for a plane joint or interface element suitable for implementation into an existing finite element code. The element is directed in particular at applications in rock mechanics. Rock joints experience complex behaviour during their loading history. A typical discontinuity may for instance experience a sequence of open, stick and slip states as the excavation advances. Furthermore, shearing is normally accompanied by dilation, which is an important contributing mechanism to the overall stability of the rock mass.

Few experimental results or analytical models have been reported in the literature which provide a detailed description of a rock joint undergoing any possible history of opening, closing and sliding, accompanied by dilation. The particular concern of the thesis is to formulate constitutive equations

for a dilatant joint element, essentially the relations governing the relative displacement between adjacent points on either side of the interface, which fully cater for all possible histories of opening, closing and sliding, in any direction. The constitutive equations are incorporated into the spatially discrete joint element by standard finite element techniques.

Special attention is paid to the establishment of a consistent formulation, retaining realistic joint response. This response should include both contact and no-contact conditions. When the interfaces are in contact, frictional sliding, accompanied by dilatant behaviour, is possible. It will be assumed that the sliding is time independent, but the constitutive equations will certainly be path dependent.

A classical formulation of a joint or interface element, which can be readily incorporated into a standard non-linear finite element code, is that proposed by Taylor, Goodman and Brekke[18] and Goodman and Dubois[17]. This element is based on a penalty formulation, accomplished by the introduction of an elastic compliance of the interface. The elastic compliance in the direction normal to the interface, when the joint is closed, is made as large as possible to enforce the contact constraint. The element, together with extensions proposed by various authors,[9,35,38] is the most used method for contact problems, and is also adopted in this formulation.

Since the interface element is to be incorporated into a standard non-linear finite element code, the formulation of the constitutive equations must be fully consistent with the forward integration algorithms along the path of loading adopted for time independent analysis in the code. These algorithms most commonly make use of an incremental formulation, with a two-step Newton-Raphson iteration to solve the non-linear incremental equations. The first step is the predictor step which involves a consistent tangent stiffness matrix and the out-of-balance nodal forces at the beginning of the iteration, and computes an updated trial displacement increment. The second step is the corrector step, which, in a Gauss-point-by-Gauss-point procedure, calculates the updated internal forces associated with the updated trial displacement increments, and subsequently the out-of-balance nodal forces for the next iteration.

The formulation of the incremental equations is most commonly carried out on the basis of a backward difference discretisation in time. The key feature of this approach is that an incremental problem for a material with characteristics of plasticity is converted into what is essentially a non-linear elastic

problem[28,29]. There are distinct advantages in formulating the incremental constitutive equations in this way before implementation into the finite element code. Opening and closing have the characteristics of non-linear reversible and hence elastic behaviour. Frictional sliding, however, has the characteristics of plasticity. The incremental non-linear elastic formulation implies that the non-linear elastic and plastic behaviour can be treated together, i.e. no qualitative distinction need be made between opening and closing on the one hand, and sliding on the other. Furthermore, a convenient formulation of the constitutive equations is facilitated by representing the different contact conditions in relative displacement (or state) space. The state diagram in relative displacement space, however, changes from one time-step to the next, and evolution equations for the updating must be formulated.

These principles are illustrated for three joint models: a Coulomb model, a sawtooth asperity model and a logarithmic spiral model. An important feature of these models is the simple explicit nature of the constitutive relations. The dilatant Coulomb model, however, has an inherent problem associated with reversed shear deformation. An inelastic shear deformation increment of any sign is always accompanied by a positive increment of inelastic normal deformation, even when the joint surfaces are not in contact. Consequently, an initially open joint will close after a finite number of shear cycles. This problem is avoided in the sawtooth asperity model. The model idealises the rough contact surfaces as interlocking sawtooth asperities which match in the virgin state. Sliding causes asperities on the one side of the joint surface to ride over the asperities on the other surface. The result is surfaces which dilate or contract, depending on the direction of sliding. This model provided important insights into the dilatant Coulomb concept, i.e. the Coulomb model is only applicable in the virgin state, and must be modified when the joint is displaced. The continuing dilation accompanied by monotonic shearing predicted by the sawtooth model is, however, a limitation. The logarithmic spiral model overcomes this limitation by assuming a non-linear relation between inelastic normal and shear relative displacement rates, with inelastic normal deformation ceasing after some shearing. As a consequence, the friction laws change from one time-step to the next in a way which is characteristic of a softening material. The model largely retains the simple explicit nature of the Coulomb and sawtooth asperity models, but offers a far more realistic model for dilatant rock joints.

The organisation of the thesis is as follows. Chapter 2 is a literature review of the behaviour of rock joints, of some of the constitutive models describing frictional behaviour, and of finite element formulations that are used to

model contact problems. The interface element, which captures the most important aspects of rock joint behaviour, is then formulated in the following three chapters. The first of these chapters, Chapter 3, considers the requirements that the joint constitutive equations must meet for implementation into a non-linear finite element code. This is followed by the formulation of explicit expressions for the corrector step of the Newton-Raphson algorithm for the three joint models. The formulation is completed in Chapter 5 where the tangent stiffness matrix and an isoparametric interface element, which incorporates the corrector step equations, are derived. These equations have been implemented as a user element into the finite element code ABAQUS[23]. Chapter 6 contains a number of examples, mainly in the field of rock mechanics, to illustrate the predictions of the formulation, to highlight the differences between the three constitutive models, and to show the effectiveness of the model.

# CHAPTER 2
# A REVIEW OF ROCK JOINTS

The material covered in this chapter is intended to provide background to concepts which follow in later chapters. The review concentrates on the mechanical properties which apply to all rock joints, but finite element formulations describing the contact problem, as well as some of the constitutive models which are used to model rock discontinuities, are also covered briefly. Attention will be restricted to two-dimensional time-independent problems. Further, we will assume that the rock mass is relatively intact, that the response is governed by a small number of discontinuities, and that each discontinuity can be represented individually by the finite element method.

The chapter is organised as follows: we first consider the mechanical behaviour of rock joints. This is followed by a discussion of the constitutive models describing frictional response, and finally, the different finite element formulations describing the contact problem are addressed.

## 2.1 ROCK JOINT BEHAVIOUR

The behaviour of a rock mass is generally governed by the properties of the intact rock as well as the properties of rock joints. Joints, for instance, provide most of the weakness, deformability and conductivity (fluid flow properties) of a rock mass. The term rock joint will be used as a collective term to refer to discontinuities of geological origin (such as faults, dykes and bedding planes) as well as fractures induced by high stress concentrations (excavations and foundations).

We will concentrate on two-dimensional discontinuities only. This is reasonable since rock discontinuities in three-dimensional space are usually planar, and can therefore be viewed as two-dimensional features having a preferred orientation. It is thus convenient to describe stress and relative deformation in directions perpendicular $(\sigma, \delta)$ and tangential $(\tau, \gamma)$ to the plane of the joint. The stress and deformation are usually related by means of compliance or stiffness terms in rock joint models. These terms can physically be

attributed to poor contact, resistance of the joints to closure, friction on the joint walls and compressibility of infilling material. The stiffness terms, together with peak and residual shear strength and displacement, maximum joint closure and maximum dilation, provide a complete description of the strength and deformability of rock joints. Strength and stiffness are strongly stress dependent and may vanish under small normal tensile stress. Under compression, however, they vary between relatively well understood limits.

A large number of experimental studies have been conducted to increase the understanding of the behaviour of natural as well as artificial joints. Artificial joints (normally made from plaster of Paris) have been studied mainly because the interface characteristics can be reproduced easily. It is especially the behaviour of rough joints undergoing shear deformation which has received considerable attention.

The parameter which is affected most of all by varying normal and shear stress is the joint aperture or normal deformation $\delta$. The aperture is firstly affected by the magnitude of normal stress and secondly by shear deformation $\gamma$. During shearing of rough joints the displacement vector deviates from the mean direction of the applied shear force due to overriding of surface irregularities. The component of the displacement vector normal to the joint plane causes an increase in aperture. This phenomenon is called dilation.

The mechanical effect of a dilational rock joint confined by surrounding rock is important; each increment of shear deformation, accompanied by dilation, results in an increase in the force normal to the joint plane, which in turn, arrests further shearing. This dilation-normal force interaction is primarily responsible for the self-stabilisation of rock masses.

Dilational properties of discontinuities are governed by the roughness of the wall surfaces (the size and shape of asperities), and the character of the intervening space. Bedding planes can derive significant roughness from sedimentary structures inherited from the deposition process, such as ripple marks, worm trails, mud cracks, etc. Fractures tend to be rough as they originate primarily from extension fractures. In cases where the joint is filled with deformable material, or where the two joint surfaces are mismatched, dilation may be preceded by a small amount of contraction.

### 2.1.1 Mechanical Properties

It is apparent that a number of parameters are directly linked to the mechanical properties of rock joints, and are therefore fundamental to a proper constitutive joint model. These parameters include the roughness of surfaces, initial aperture, wall strength, filling material and moisture.

The wall roughness determines the dilational properties of a joint.

The initial aperture is defined as the aperture under small or zero normal stress levels. Related to this is the initial contact area and the distribution of apertures and contact points between surfaces. The initial contact area changes when the joint is sheared and the asperities no longer match.

The strength and deformability of the asperities, and wall in general, determine which failure mechanism dominates at a given normal stress level.

The most obvious effect of filling material is that it separates the joint walls and reduces the rock-rock contact. The filling material is rarely of a better quality than the host rock. Sometimes, however, the joint surfaces are cemented together with infill material, such as quartz or calcite, which has properties as good, or even better than the properties of the host rock. It is known that shear strength decreases with increasing infill thickness, and that dilation is normally preceded by contraction if the joint has thick infill. It is further important to note that a gouge, which is considered as a weak infilling material, is formed when a joint is sheared and the asperities degrade.

Moisture plays an important role. Clay minerals, such as those of the montmorillonite group, can cause extremely slippery conditions with negligible shear strength when wet. Pore water pressure can be introduced in the stiffness matrix of the joint or applied as an external force to both joint walls. The effect of pore water pressure on the deformability and strength of the rock mass depends mainly on the permeability of joints.

Most of these parameters are affected by *scale* effects.[3,12] This term is used to describe the variation of test results with specimen sizes. Only small scale waviness, the first component of roughness, is represented by laboratory tests. The changing size of significant asperities along joints as sample or block size increases, and the different effective normal stress concentrations on the effective joint wall contacts during shear, are primarily responsible for scale effects. There are still different views regarding the effects of scale.

Most researchers agree, though, that the magnitude of the scale effect is directly related to the roughness pattern; the higher the roughness of the joints, the greater the scale effects. On the other hand, if the joint is filled with low strength and highly deformable material, and the width of the filling is sufficient to prevent wall contact, no significant scale effect can be observed. It is further known that increases in specimen size increase the peak shear deformation and reduce the asperity strength and roughness, and that different failure mechanisms may dominate at different sample lengths.

### 2.1.2 Normal Stress-Closure Behaviour

Experiments indicate that the normal stress $\sigma$ versus closure $\delta$ relation for a wide range of natural, unfilled joint types are highly non-linear. All joints display an initial amount of joint aperture at low normal stress. As the normal stress increases, the joint closes until a point of maximum closure is reached at high normal stress. Figure 2.1 shows the behaviour of a joint subjected to one complete loading/unloading cycle. The behaviour can be explained as follows. Upon initiation of loading the joint closes rapidly as asperities readjust to their initial seating condition. As the normal stress increases and the initial seated position is taken up, further closure depends almost exclusively on the deformability of asperities. The tight mechanical interlock between asperities at high normal stress, as well as the increased actual contact area, creates a very effective confined environment, thus stiffening the deformational response of the joint. Upon subsequent unloading the joint responds in a hysterectic and inelastic manner, with some permanent, or irrecoverable, closure. The joint surfaces separate under zero or low tensile stress conditions.

Barton *et al.*[4,7] proposed a hyperbolic function of the form

$$\sigma = \frac{\delta}{a - b\delta} \tag{2.1}$$

to describe the normal stress-closure behaviour. The maximum joint closure $\delta^m$ is equal to $\frac{a}{b}$, and the initial normal stiffness $K_t^0$ is equal to $\frac{1}{a}$. Estimates of $K_t^0$ and $\delta^m$ for successive loading/unloading cycles are given by Barton,[7] and are based on the joint roughness coefficient, the joint wall compressive strength and the initial joint aperture.

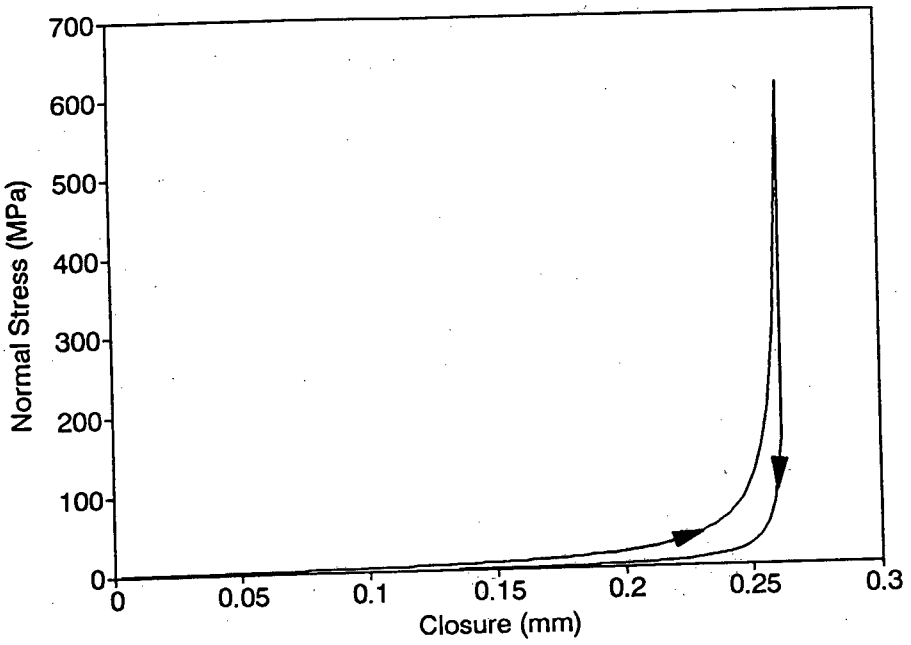



Figure 2.1: Normal Stress *vs* Closure Behaviour

Goodman *et al.*[18] suggested the following relationship

$$\frac{\sigma - \sigma_0}{\sigma_0} = C \left( \frac{\delta}{\delta^m - \delta} \right)^k , \tag{2.2}$$

where $\sigma_0$ is the initial stress, and $C$ and $k$ are constants. Beer[9] uses a relationship of the form

$$\sigma = \frac{1}{\delta^m - \delta} + \frac{C}{\delta^m K_t^p} , \tag{2.3}$$

where $C$ is a constant determining the rate of closure, and $K_t^p$ is the peak normal stiffness (stiffness at maximum closure). The models of both Goodman and Beer are non-linear elastic. Barton, on the other hand, provided different constants to describe the different loading and unloading paths. These constants, however, represent complete loading/unloading cycles; the specimen is loaded until maximum closure is reached and then completely unloaded before the next loading cycle is performed. No results of partially loaded/unloaded specimens could be found.

The normal stress-closure behaviour of joints sheared from the virgin position reveals similar behaviour to joints loaded in a fully seated position. Compar-

isons of several interlocked and dislocated normal stress-closure records of the same joint specimens indicated that the interlocked stiffness was several times higher. The much lower stiffness of mismatched joints can be attributed to a stress concentration over a lower contact area and the lack of asperity confinement. However, the differences seem to reduce considerably under repeated loading. Results further indicate that only an extremely small amount of dislocation from an original interlocked position is required to cause a large reduction of the original stiffness. Small post-peak displacements indicate a further reduction in stiffness, but at a much lower rate. Barton[7] recommended a logarithmic relationship to describe the normal stress-closure behaviour of unmatched joints. He also proposed an empirical relation to modify the normal stiffness to allow for changes induced by shearing.

### 2.1.3 Shear Stress-Shear Deformation Behaviour

The main problem in finding a suitable frictional model is that widely different behaviour can be expected at different normal stress levels. The shear stress $\tau$ versus shear displacement $\gamma$ curves of Figure 2.2 indicate that the behaviour can range from *brittle* to almost *ductile*. Brittle behaviour is most typical of rough joints, small block sizes, low normal stress and high host rock strengths. Ductile behaviour results from exactly the opposite characteristics.

The reason for the wide range of behaviour is that the asperity surfaces, which are responsible for dilatency, have finite strength, and, depending upon the severity of the stress and the amount of sliding, will degrade and affect subsequent behaviour. Figures 2.3 and 2.4 show how the wall strength, joint roughness and residual friction angle influence the shear strength and dilational properties of a rock mass.

One of the first researchers to distinguish between different types of shear response was Patton.[34] He proposed a bi-linear model where the behaviour is characterised by overriding of asperities at low normal stress, and shearing through asperities at high normal stress levels.

Barton[4,5] subsequently extended Patton's ideas, and proposed that the two failure mechanisms occur simultaneously. The amount of overriding/shearing is determined by $\mu$, which is a geometrical parameter and referred to as the dilation angle. The dilation angle depends on the magnitude of normal stress,

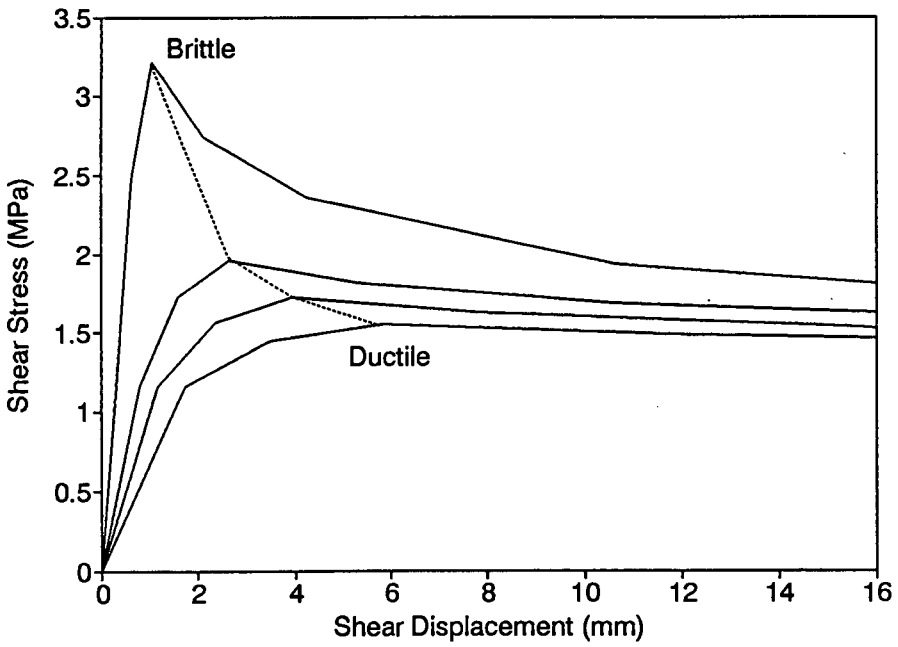



Figure 2.2: Shear Stress *vs* Shear Deformation Behaviour

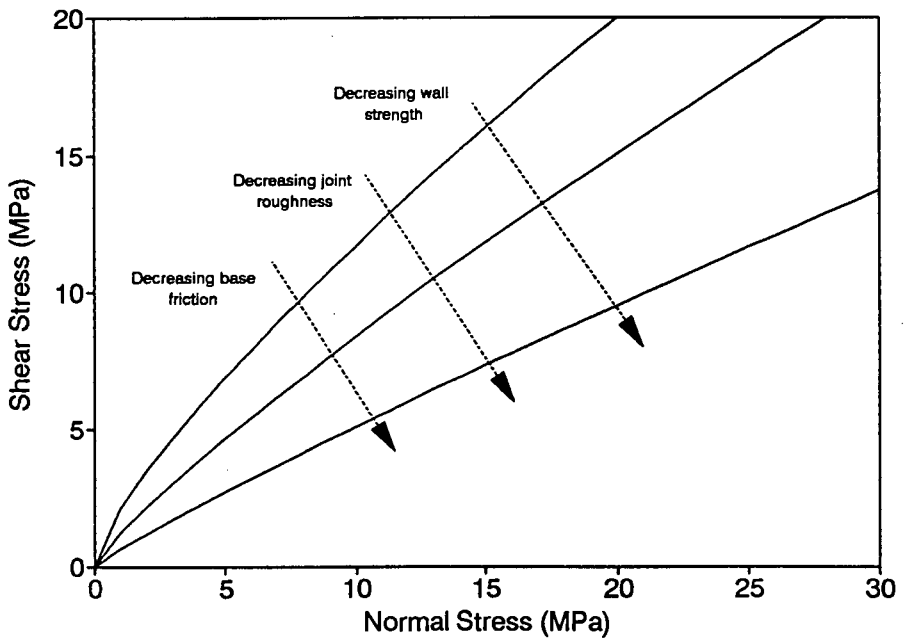



Figure 2.3: Influence of joint parameters on shear strength

increasing shear deformation by reducing the dilation angle as plastic work is accumulated. In his model the degree of surface degradation is not sensitive to normal stress variations.

Roberts and Einstein[37] distinguished between four different shear responses. Firstly, at very low normal stress the two sides of the discontinuity slide over each other without shearing off any of the asperities. As the normal stress increases dominant asperities are sheared through, while others are overridden. As the normal stresses increases still further, all asperities are sheared off at relatively small shear deformation and softening behaviour occurs. Finally, at very high normal stresses there is a transition from sliding along the discontinuity to fracturing through the intact material, and a simultaneous brittle-ductile transition followed by hardening behaviour of intact material.

Handanyan *et al.*[20] investigated the exact failure of asperities. They realised that asperities may fail in one of three ways: shearing of the asperities, elastic and plastic deformation and eventually sliding over the asperities, or tensile splitting of asperities. Their experiments indicate that tensile splitting is the most likely to occur.

The behaviour shown in Figure 2.2 is typical of a direct shear test carried out at constant normal stress. These testing procedures are, however, inadequate for cases where dilation is partially or totally restricted by the stiffness of the surrounding rock mass since the normal and tangential behaviour is largely uncoupled. Constant normal stiffness boundary conditions are more likely to exist across *in situ* joint surfaces,[39] but little test information is documented. Amadei and Saeb[1] proposed a graphical as well as a numerical model that is based on constant normal stress results to predict the behaviour under different boundary conditions. Archambault *et al.*[2] reported that increases in normal stiffness increases the peak shear strength, the residual shear strength as well as the normal stress, but decreases the peak dilation. Ohnishi and Dharmaratne[31] showed that the shear stress-shear displacement curves under constant normal stress and constant normal stiffness become similar, with identical peak shear displacements at high normal stress conditions.

Few experimental results exist for cyclic shear tests. Results from Kutter and Weisbach[27], as well as the more resent work by Hutson and Dowding,[24] indicate that the specimen contracts to almost the virgin or interlocked state when the direction of shear is reversed. This behaviour seems reasonable when the normal stress across the joint is low and asperity overriding is predominant. No experimental results could be found were the joint was

subjected to different histories of opening, closing and sliding states. Pande *et al.*[10] offer the following explanation

> "... our knowledge of joints subjected to a combined action of tension and shear is very limited since it is quite difficult to carry out experimental investigations in this stress regime. Conceptually, it appears that a joint will be able to withstand shear while it is opening under tensile normal stress as the asperities will be interlocked. However, when the joint opening is equal to or greater than the average height of asperities, interlocking should cease, and the joint will be incapable of withstanding neither shear nor tensile stress ..."

Such behaviour is anticipated to occur along discontinuities which deform due to excavations in highly stressed rock. A typical example is the behaviour of vertical fractures at the proximity of a deep excavation in rock. We will consider this problem in more detail later.

## 2.2 CONSTITUTIVE MODELS

The most widely accepted constitutive models to describe rock joint behaviour are the classical Coulomb friction law and the empirical Barton-Bandis model. These models, as well as some other analytical and empirical models which highlight important developments and contributed to the understanding of rock behaviour, are considered in this section.

The most popular method of modelling frictional behaviour is by means of the theory of elasto-plasticity. Elastic behaviour is represented by the initial shear stiffness $K_s$. The peak strength and dilatency, on the other hand, are represented by a failure criterion and flow rule respectively, and post-peak behaviour by a hardening/softening rule.

The French military engineer, C.A. Coulomb, introduced the concept that shearing resistance $\tau$ is composed of two components, namely cohesion $c$ and friction $\phi$. The relation in terms of a failure criterion is written as

$$F = |\tau| - \sigma \tan \phi - c \,. \tag{2.4}$$

Although his formula embodying these concepts is still universally used in many engineering fields, it represents only special limiting cases which are

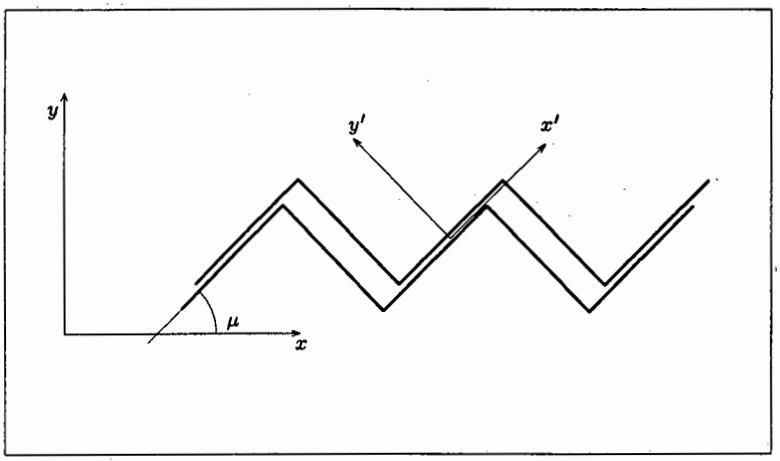



Figure 2.5: Asperity surface model

seldom encountered. In this relation the strength parameters $\phi$ and $c$ are assumed to be constant. In the case of rock joints, however, they depend on the magnitude of the normal stress which can vary by several orders of magnitude. Rates of inelastic normal deformation $\dot{\delta}^p$ and shear deformation $\dot{\gamma}^p$ can be obtained by assuming an associated flow rule. This implies that

$$\dot{\delta}^p = \dot{\gamma}^p \tan\phi \,. \tag{2.5}$$

An increment of shear deformation is therefore accompanied by normal deformation or dilation. The rate of dilation is, however, constant. Various studies have confirmed that the flow rule of rock joints should be non-associated and that dilation must cease when the normal displacement is equal to the average height of the asperities.

Patton[34] conducted experiments on plaster of Paris specimens in order to investigate the mechanism of shear failure in rock. His specimens were cast with irregular surfaces, and the test variables included the inclination, number and strength of the teeth, and the normal loads applied. At low compressive stresses ($\sigma < \sigma_{cr}$) the behaviour is characterised by overriding of asperities resulting in an increase of normal deformation. For high compressive stresses ($\sigma > \sigma_{cr}$) the behaviour is characterised by shearing through asperities with no normal deformation. The bi-linear failure envelope can be expressed as

$$\tau = \sigma \tan(\phi + \mu) \qquad \sigma < \sigma_{cr} \,, \qquad \text{and}$$

$$\tau = c + \sigma \tan \phi \qquad \sigma \geq \sigma_{cr} \,, \tag{2.6}$$

in which $c$ = shear strength of the asperities, $\sigma_{cr}$ = critical stress, $\mu$ = the orientation of asperities with respect to the plane of the joint, and $\phi$ = residual or basic angle of friction.

Several other authors adopted Patton's *angle* approach. Plesha[35,36] distinguished between macroscopic and microscopic features of the contact surface to formulate his analytical constitutive model. Through macroscopic considerations, an incremental constitutive law is derived. By idealising the microstructure to consist of interlocking asperity surfaces, the constitutive equations are specialised for the description of rock joints and include effects such as dilation and asperity surface degradation. He considered a sawtooth model as well as a sine asperity model. He assumed Coulomb friction on the asperity surface and wrote a yield function $F$ and flow potential $Q$ in the plane of the asperity surfaces. By suitable transformation to macroscopic coordinates of the interface, the following equations can be obtained

$$\begin{aligned} F &= |\sigma \sin \mu_k + \tau \cos \mu_k| + (\sigma \cos \mu_k - \tau \sin \mu_k) \tan \phi \,, \\ Q &= |\sigma \sin \mu_k + \tau \cos \mu_k| \,, \end{aligned} \tag{2.7}$$

where $\mu_k$ is taken from either the sawtooth model or the sinusoidal asperity model, and $k$ denotes which asperity surface (left or right) is active. He assumed further that asperity degradation is a function of plastic sliding work $dW_t^p$. The relationship is given by

$$\mu_k = (\mu_k)_0 \cdot e^{-cW_t^p} \,, \qquad \text{with } \mathrm{d}W_t^p = \tau \cdot \mathrm{d}\gamma^p \,, \tag{2.8}$$

where $(\mu_k)_0$ is the initial asperity surface angle and $c$ is a rock joint degradation constant which has units of length/force and reflects how rapidly the asperity surface deteriorates.

The Barton-Bandis[4,5,6] model is considered the most sophisticated constitutive model available at present since it takes a wide range of behaviour into consideration. The model is based on a number of easily measured parameters: the joint roughness coefficient JRC which varies between 0 for smooth joints and 20 for very rough joints, the joint wall compressive strength JCS, and the residual friction angle $\phi$. The model is essentially an extension of Patton's $(\phi + \mu)$ concept. In this model the total frictional resistance is composed of three angular components: a basic frictional component, a geometrical component of asperity overriding, and an asperity failure component. The resistance against overriding of asperities (given by the dilation angle $\mu$), and the resistance against asperity failure, are almost identical for

a wide range of conditions. The shear strength is therefore given by

$$\tau = \sigma \tan(\phi + 2\mu) \,. \tag{2.9}$$

The dilation angle depends on the effective normal stress and varies with shear displacement. It is therefore controlled by surface roughness, asperity strength and rock type. The peak dilation angle can be obtained from

$$\mu = \frac{1}{2} \,\mathrm{JRC}\, \log\left(\frac{\mathrm{JCS}}{\sigma}\right) \,. \tag{2.10}$$

At stress levels approaching the level of the joint wall compressive strength, JCS is substituted by the confined compressive strength $(\sigma_1 - \sigma_3)$. The higher asperity strength is due to the increased confinement with greater areas of contact. The peak shear stiffness can be estimated as

$$K_s = \frac{100}{L}\sigma \tan\left[\mathrm{JRC} \cdot \log\left(\frac{\mathrm{JCS}}{\sigma}\right) + \phi\right] \,. \tag{2.11}$$

This equation is based on the assumption that the peak shear strength is reached after the joint has sheared 1% of the total length $L$ of the joint. Barton also provided curves to modify JRC and JCS so that post-peak behaviour can be modelled. Pande and Xiong[33] implemented this formulation into a finite element code.

Ghaboussi *et al.*[16] uses a capped plasticity model with a perfectly plastic yield surface to limit shear stress, and a hardening cap to control dilation. The yield surface is shown in Figure 2.6 in terms of $\sigma$ and $\tau$. Their equations are of the form

$$F_1(\sigma, \tau) = 0 \qquad \text{and} \qquad F_2(\sigma, \tau, \kappa) = 0 \,. \tag{2.12}$$

Stress states on $F_1$ are associated with expansion and dilation of the joint and $F_2$ are associated with contraction. Inelastic shear deformation may occur on either the yield surface or the cap. The hardening parameter $\kappa$ may be defined so that the cap moves in the compressive $\sigma$ direction during contraction and in the tensile $\sigma$ direction during dilation. The cap controls the amount of dilation.

Heuze and Barbour[22] considered the shear strength envelopes shown in Figure 2.7. The figure shows the residual strength envelope which is a straight line going through the origin, and with a slope equal to $\tan\phi_r$, in which $\phi_r$ is called the residual friction angle. It also shows the peak envelope being composed of two parts: below a critical normal stress $\sigma_{cr}$, the envelope is curved, and above $\sigma_{cr}$ it is a straight line parallel to the residual envelope. The curved part corresponds to the lower stress region where overriding is predominant. Above $\sigma_{cr}$ shearing through asperities dominates. The dila-

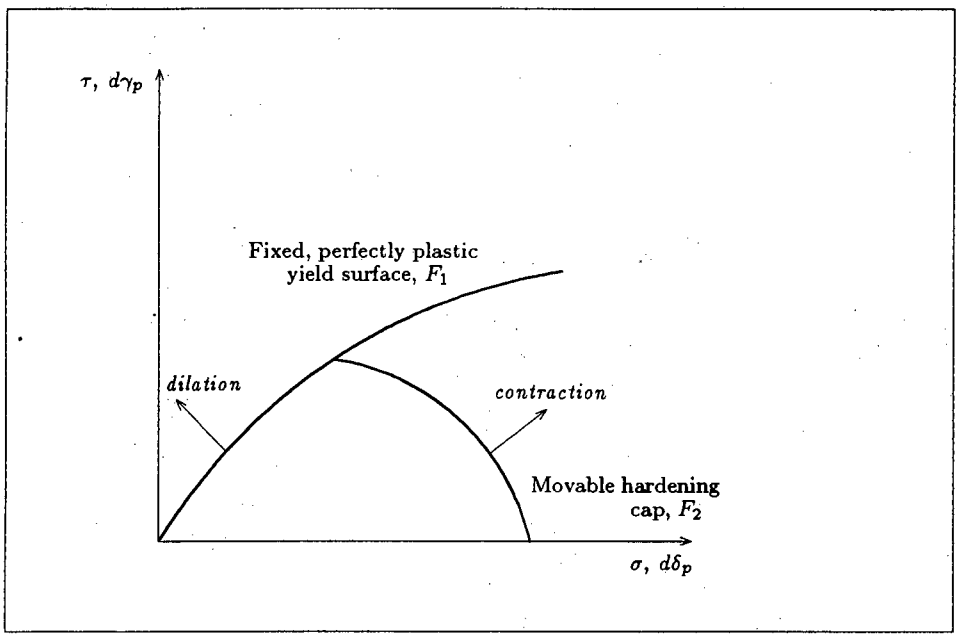



Figure 2.6: Yield Surface with Cap

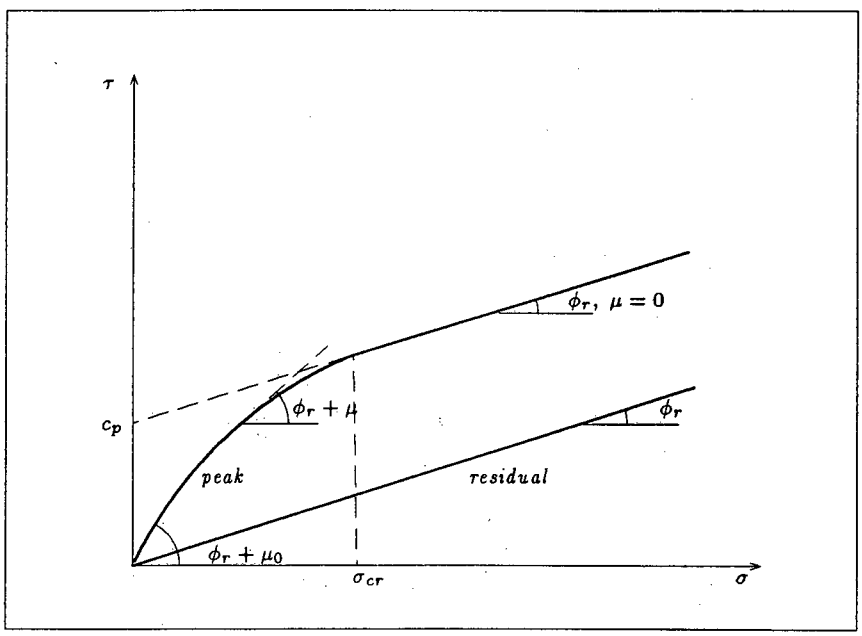



Figure 2.7: Shear Strength Envelopes

tant range is the range in which $\sigma \leq \sigma_{cr}$. In this region, the friction angle $\phi$ is the sum of the residual friction angle $\phi_r$, and the dilation angle $\mu$, and decreases from a peak value $\phi_p$ at the origin to $\phi_r$ for $\sigma > \sigma_{cr}$. This decrease corresponds to the decrease of dilation angle from its maximum value $\mu_0$ to a minimum value of zero for $\sigma > \sigma_{cr}$. The following relationships describe the various envelopes

$$\tau = A\sigma + B\sigma^2 + C\sigma^3 , \qquad \sigma \leq \sigma_{cr} \tag{2.13}$$

where $A, B$ and $C$ are functions of the residual friction angle, the peak friction angle and the critical normal stress across the joint surfaces. When $\sigma > \sigma_{cr}$ the peak strength is simply

$$\tau = c_p + \sigma \tan \phi_r , \tag{2.14}$$

and in the residual range

$$\tau = \sigma \tan \phi_r . \tag{2.15}$$

## 2.3 NUMERICAL MODELLING

One method of modelling the discontinuities in a rock mass is to consider each joint as a contact surface between two bodies. The problem is then essentially one of establishing compatibility conditions between the two contact surfaces. The basic surface compatibility conditions include no-contact, or contact, but with no penetration of one side of the body allowed into the other, and a pressure developing along the region in contact. When the surfaces are in contact, sticking or frictional sliding is possible. The contact problem is non-linear due to these varying conditions. Much of the difficulty of handling contact problems lies in deciding which contact conditions to apply, and an exhaustive set of criteria is often required to test the validity of an assumed state.[8,25] Katona[25] reported convergence problems when the response is a borderline case between two or more possible states.

The basic approaches employed to solve the contact problem are the Lagrange-multiplier method[8,25] and the penalty method.[18,40] A mixed formulation was suggested by Herrmann.[21] The finite element program ABAQUS[23] also uses a mixed formulation. The code imposes the normal conditions by Lagrange-multipliers, and the tangential conditions by the penalty method. Bird and Martin[11] proposed a method based on the internal variable method of plasticity.

In the Lagrange-multiplier method the contact forces are treated as independent variables which are added to the system of finite element equations. The disadvantage of this approach is that the introduction of additional variables increases computational effort. The method may further lead to singular stiffness matrices.[30,42] The advantage is that the no-penetration condition is enforced exactly, and that the uncertainty associated with choosing artificial constraints (penalty terms) is avoided. The contact conditions in this method are often enforced in a *point-wise* manner at the element nodes. Wriggers *et al.*[42] proposed a method where the contact pressures, and not the nodal forces, are taken as the independent Lagrange multipliers. The contact conditions in this formulation are thus enforced along the element boundary.

The penalty method is the most frequently used method in modelling discontinuities. In this method the contact conditions are enforced in an approximate manner, and the contact pressures do not explicitly appear in the formulation. The method avoids most of the problems associated with the Lagrange-multiplier method, but is susceptible to ill-conditioning due to large stiffness terms which may numerically overshadow the contributions from surrounding continuum elements. Wilson[41] demonstrated by a simple numerical example that the accuracy of the solution is greatly improved if relative displacements are introduced as independent variables. Pande and Sharma[32] later showed that the problem of ill-conditioning is not critical on accurate main frame machines. Oscillations in the stress along the surfaces in contact is another problem attributed to high penalty terms. Kikuchi *et al.*[26,30] advocate the use of reduced numerical integration to avoid this problem.

The penalty formulation produces an element stiffness matrix and a load vector analogous to the element stiffness matrix and load vector of a typical solid element (which can be assembled into the global finite element equations by standard techniques), and the term *joint* or *interface element* is therefore appropriate. There are different ways of formulating the interface element. Desai *et al.*[13] and Francavilla and Zienkiewicz[14] treated the joint as a solid or continuum element of small thickness, assuming constant strain in the thickness direction. These elements are commonly referred to as thin-layer elements. Other authors[19,32] have shown that even the familiar eight-node plane continuum element with reduced numerical integration performs well as an interface up to quite high aspect ratios. A more widely accepted approach is the zero thickness elements, which are especially suited for modelling rock discontinuities. Modern versions of this element, first proposed by

Taylor, Goodman and Brekke,[18] are based on isoparametric principles[9,38,35] and curved contact surfaces can therefore be represented by higher order approximations. This implies that the contact conditions are imposed in an *average* sense along the element boundary.

The method can be interpreted as inserting stiff springs between the contact surfaces to enforce the contact condition

$$N = K_t \delta \,, \tag{2.16}$$

where $N$ is the contact force, $\delta$ the relative displacement between points on either side of the contact surface, and $K_t$ is the normal stiffness or penalty term. Since the contact conditions do not hold exactly, material overlapping occurs. Most formulations consider this part of the behaviour as either linear elastic (constant stiffness) or non-linear elastic[9] (stiffness increasing with increasing penetration). When the surfaces separate the penalty term is set to zero or a very low value.

Separation is often monitored by means of a failure criterion, and the deformation when the joint opens is considered to be plastic. Such a formulation is restricted in the sense that the opening-closing response is unidirectional, i.e. subsequent reloading leads to either physically unreasonable or inconsistent behaviour. This problem is further complicated when closing is preceded by shearing (accompanied by dilation) with the joint surfaces in contact. Pande *et al.*[10] encountered this problem and stated that

> "... compressive stresses are transmitted across the open joint implying a mismatch of joint asperities ..."

and further that

> "... joints have no memory of their opening/closing ..."

They suggested an additional yield function which ensures that contact stresses are transmitted across joints only when they are closed.

The tangential contact conditions (sticking and frictional sliding) are modelled in a similar way. The frictional response is assumed elasto-plastic. Elastic behaviour, represented by an initial tangential stiffness $K_s$, represents the sticking contact state. Peak shear strength and inelastic deformation are represented by a failure criterion and flow rule respectively. Post-peak behaviour

is modelled by hardening/softening rules.

Time dependent behaviour of joints can similarly be represented by visco-plastic concepts.[37] Some authors,[9,10] however, use visco-plasticity purely as an alternative elasto-plastic algorithm for computational convenience.

## 2.4 SUMMARY

The penalty method, which imposes the contact compatibility requirements in an approximate sense, is the most popular numerical technique to model contact problems in rock masses. Elasto-plastic concepts are often used to describe deformation in both the normal and tangential directions.

The most important behaviour of a rough joint is that shearing at low normal stress levels is accompanied by a tendency for ride-up action over asperities leading to dilation. At high normal stress the asperities are sheared at the base, inhibiting the tendency to dilate. A number of constitutive models, ranging from the crude Coulomb model with an associated flow rule to the sophisticated model of Barton-Bandis, can be used to approximate this behaviour.

However, it appears from the literature that no analytical model currently exists which addresses the problem of a rough joint subjected to opening/closing during its loading history, in much detail. The remainder of this thesis is devoted to the establishment of dilatant joint models that cater for any possible history of opening, closing and sliding accompanied by dilation. We will not incorporate all the detailed mechanical behaviour discussed in this chapter into the constitutive models. However, special attention will be paid, firstly to establish a consistent formulation, and secondly to retain realistic rock joint behaviour as far as possible. The dilation/contraction accompanied by sliding is considered to be the most important behaviour.

characteristics of non-linear reversible and hence elastic behaviour. Frictional sliding, on the other hand, has the characteristics of plasticity. The incremental holonomic (or non-linear elastic) formulation, however, implies that non-linear elastic and plastic behaviour can be treated together, i.e. no qualitative distinction need be made between opening, closing and sliding.

We will mainly focus on the corrector step problem in this thesis.

## 3.2 THE CORRECTOR STEP PROBLEM

Consider, as shown in Figure 3.1, a plane joint of unit length. In the reference configuration the joint is in its virgin state, which we define as closed with no relative displacement between the two faces of the joint, and no forces transmitted across the interface. With reference to this virgin configuration $\delta$ and $\gamma$ denote opening and sliding components of the relative displacement of the positive side of the interface with respect to the negative. Positive senses of $\delta$ and $\gamma$ are shown in Figure 3.1. The unit vector $\mathbf{s}$ along the joint must be specified, and the unit vector $\mathbf{n}$ is in the direction 90° anticlockwise from $\mathbf{s}$; positive $\mathbf{n}$ denotes the positive side of the interface.

The pairs of forces $N, S$ conjugate to $\delta, \gamma$ are also shown in Figure 3.1. These pairs of forces are respectively the normal and shear forces acting across the joint. In order to execute the corrector step of the Newton-Raphson algorithm we must be able to solve the following problem: given at time $t_n$ relative displacements $\delta_n = \delta(t_n)$, $\gamma_n = \gamma(t_n)$ and the associated conjugate forces $N_n = N(t_n)$, $S_n = S(t_n)$, and given estimated relative displacement increments $\Delta\delta_{n+1}$, $\Delta\gamma_{n+1}$ in the $(n+1)^{th}$ time increment, leading to relative displacements $\delta_{n+1} = \delta_n + \Delta\delta_{n+1}$ and $\gamma_{n+1} = \gamma_n + \Delta\gamma_{n+1}$, find the associated conjugate forces $N_{n+1}, S_{n+1}$.

We shall assume that when the two faces of the joint are in contact reversible linear elastic relative displacements may occur. In consequence, we divide the relative displacements into elastic and inelastic parts, denoted by superscripts $e$ and $p$ respectively;

$$\begin{aligned} \delta &= \delta^e + \delta^p \,, \\ \gamma &= \gamma^e + \gamma^p \,. \end{aligned} \tag{3.1}$$

We shall further assume that a simple uncoupled relation exists between the

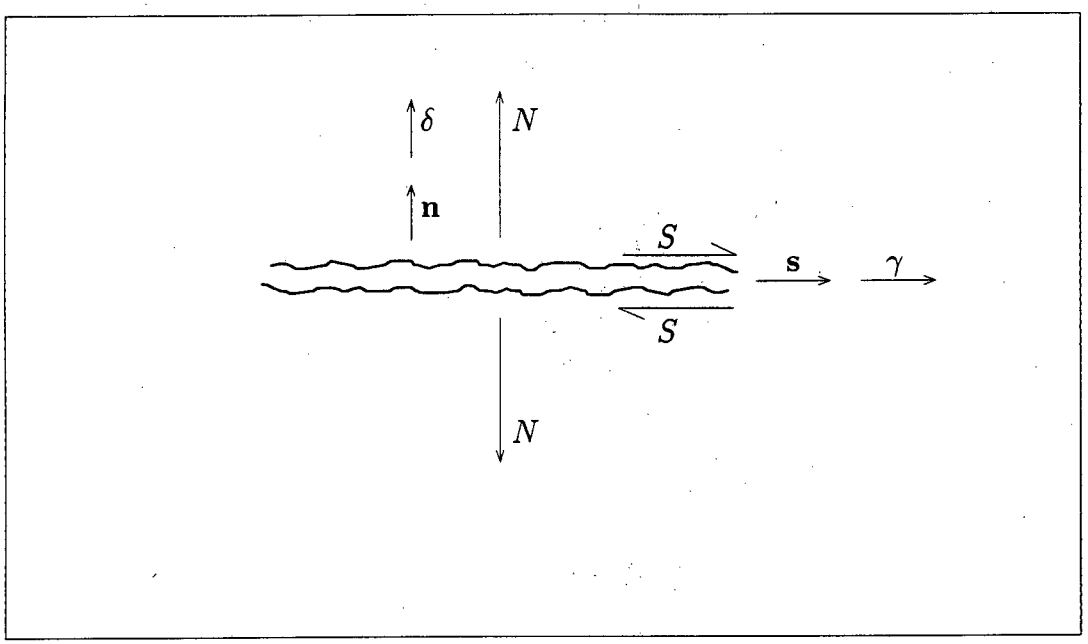



Figure 3.1: Plane rock joint of unit length

elastic components of the relative displacements and the conjugate forces, of the form

$$\left\{ \begin{array}{c} N \\ S \end{array} \right\} = \left[ \begin{array}{cc} K_t & 0 \\ 0 & K_s \end{array} \right] \left\{ \begin{array}{c} \delta^e \\ \gamma^e \end{array} \right\} , \tag{3.2}$$

where $K_t$ and $K_s$ are the elastic constants. It follows then that if $N_n, S_n$ are known, $\delta_n^e, \gamma_n^e$ are also known. The fundamental corrector step problem posed above thus reduces to the problem of determining the inelastic components $\Delta\delta_{n+1}^p, \Delta\gamma_{n+1}^p$ of the increments $\Delta\delta_{n+1}, \Delta\gamma_{n+1}$. We can then find

$$\begin{aligned} \delta_{n+1}^e &= \delta_n^e + \Delta\delta_{n+1}^e = \delta_n^e + (\Delta\delta_{n+1} - \Delta\delta_{n+1}^p) , \\ \gamma_{n+1}^e &= \gamma_n^e + \Delta\gamma_{n+1}^e = \gamma_n^e + (\Delta\gamma_{n+1} - \Delta\gamma_{n+1}^p) , \end{aligned} \tag{3.3}$$

and hence

$$\left\{ \begin{array}{c} N_{n+1} \\ S_{n+1} \end{array} \right\} = \left[ \begin{array}{cc} K_t & 0 \\ 0 & K_s \end{array} \right] \left\{ \begin{array}{c} \delta_{n+1}^e \\ \gamma_{n+1}^e \end{array} \right\} . \tag{3.4}$$

The backward difference assumption adopted here essentially means that the interface force change can be divided into two parts. First, the forces change from $N_n, S_n$ to $N_{n+1}, S_{n+1}$ along an elastic path (i.e. within the yield surface). No inelastic relative displacements occur in this first part. Then,

in the second part, with the interface forces fixed at $N_{n+1}, S_{n+1}$, inelastic relative displacement changes occur satisfying some flow rule. Since the interface forces remain constant, no changes in the elastic components occur in the second part.

# CHAPTER 4
# CONSTITUTIVE MODELS FOR DILATANT JOINTS

We proceed with the discussion in this chapter by first formulating expressions for a very simple dilatant model where sliding is governed by Coulomb friction, with no cohesion, and a non-associated flow rule. These concepts are then extended in later sections to formulate more realistic dilatant models.

## 4.1 SIMPLE COULOMB FRICTION MODEL

### 4.1.1 Displacements from the Virgin State

We demonstrate the underlying principles by considering the joint (Fig. 3.1) in the virgin position where $\delta_0 = \gamma_0 = 0$, and $N_0 = S_0 = 0$; the joint is in contact, with no forces across the interface, and has not been subjected to any previous relative movement. Displacements from an arbitrary position are considered in the following section.

As indicated, we assume that sliding of the joint is governed by Coulomb friction with a non-associated flow rule. The yield surface is shown in Figure 4.1, with an angle of friction $\phi$ and zero cohesion. The state point $N, S$ is constrained to lie in the shaded region in Figure 4.1, i.e. within or on the yield surface. If the state point lies within the yield surface, i.e. if

$$|S| < -N \tan \phi \,, \tag{4.1}$$

the inelastic relative displacement rates $\dot{\delta}^p, \dot{\gamma}^p$ are zero. If, on the other hand, the state point lies on the yield surface, i.e. if

$$|S| = -N \tan \phi \tag{4.2}$$

at a point denoted by P, the inelastic relative displacement rates may be non-zero, and must satisfy the relation

$$\dot{\delta}^p = |\dot{\gamma}^p| \tan \mu \,, \tag{4.3}$$

were $\mu$ is the dilation angle. If $\mu = \phi$ an associated flow rule is recovered. Physical considerations require that $0 \leq \mu \leq \phi$

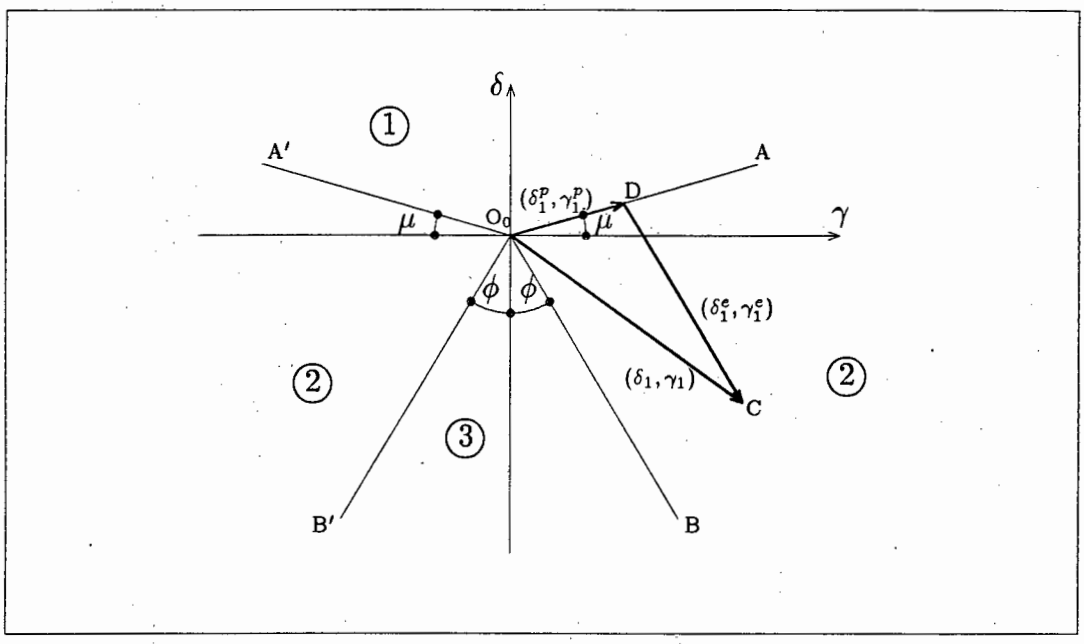



Figure 4.2: Displacements in relative displacement space

then

$$\begin{aligned} \delta_1^e &= \gamma_1^e = 0\,, \\ \delta_1^p &= \delta_1\,, \qquad \gamma_1^p = \gamma_1\,, \\ N_1 &= S_1 = 0. \end{aligned} \tag{4.4}$$

For geometric simplicity, let us now assume that the elastic constants $K_t, K_s$ are equal. The vector $(N, S)$ thus has the same direction as the vector $(\delta^e, \gamma^e)$. If this is the case, we can define region 3 in Figure 4.2 by drawing lines $O_0B$, $O_0B'$ at an angle $\phi$ through the origin on either side of the negative $\delta$-axis. If $\delta_1, \gamma_1$ lies in region 3, or on its boundaries, the response is entirely elastic, and the inelastic components of $\delta_1, \gamma_1$ are zero. Hence we can write:
if

$$\delta_1 \leq -|\gamma_1| \cot \phi\,,$$

then

$$\begin{aligned} \delta_1^p &= \gamma_1^p = 0\,, \\ \delta_1^e &= \delta_1\,, \qquad \gamma_1^e = \gamma_1\,, \\ N_1 &= K_t \delta_1\,, \qquad S_1 = K_t \gamma_1. \end{aligned} \tag{4.5}$$

In the remaining areas of Figure 4.2, which we denote by region 2, both the elastic and inelastic components of $\delta_1, \gamma_1$ will be non-zero, and the interface forces $N_1, S_1$ will be non-zero and will lie on the yield surface. If we consider the case where $\delta_1, \gamma_1$ is represented by the vector $O_0C$, we may locate a point D on line $O_0A$ by drawing DC parallel to $O_0B$. The inelastic component of $\delta_1, \gamma_1$ is then given by $O_0D$, and the elastic component by DC. Formally we write, after some straightforward manipulation:
if

$$|\gamma_1| \tan \mu \; > \; \delta_1 \; > \; -|\gamma_1| \cot \phi \, ,$$

put

$$\begin{aligned} a &= \frac{\delta_1 \sin\phi + |\gamma_1| \cos\phi}{\cos(\phi - \mu)} \, , \qquad \text{and} \\ b &= \frac{-\delta_1 \cos\mu + |\gamma_1| \sin\mu}{\cos(\phi - \mu)} \, , \end{aligned}$$

then

$$\begin{aligned} \delta_1^p &= a \sin\mu \, , & \gamma_1^p &= a \cos\mu \, \mathrm{sgn}\,(\gamma_1) \, , \\ \delta_1^e &= -b \cos\phi \, , & \gamma_1^e &= b \sin\phi \, \mathrm{sgn}\,(\gamma_1) \, , \\ N_1 &= K_t \delta_1^e \, , & S_1 &= K_t \gamma_1^e \, . \end{aligned} \tag{4.6}$$

We differentiate in the remainder of the thesis between the yield surfaces drawn in force space and relative displacement space by referring to them as *the yield surface* and *the state diagram* respectively.

### 4.1.2 Updating in Relative Displacement Space

At the commencement of the second time-step the yield surface in force space (Fig. 4.1) remains unchanged since no hardening or softening takes place. However, the state diagram in relative displacement space (Fig. 4.2) will change, and evolution equations for the updating must be formulated.

In this updating, the figure representing the lines joining $O_0$ to A,A',B and B' simply translates, without rotation or changes in the angles between the lines. The updated position is shown in Figure 4.3; the vertex moves to a position $O_1$ defined by coordinates $\delta_1^o, \gamma_1^o$. It then follows that for the virgin state, $\delta_0^o = \gamma_0^o = 0$.

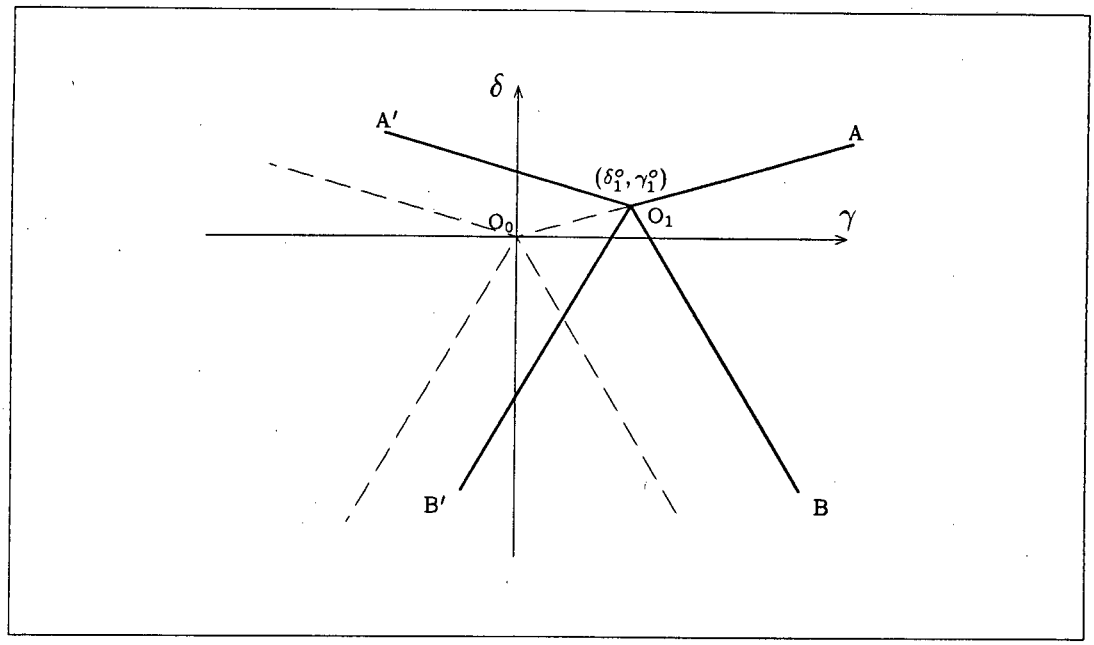



Figure 4.3: Updated position in relative displacement space

Consider first the case where the state point $\delta_1, \gamma_1$ lies in region 3 of Figure 4.2. In this case the response is elastic, and the state diagram is unchanged. Hence $\delta_1^o = \delta_0^o, \gamma_1^o = \gamma_0^o$.

Next, consider the case where $\delta_1, \gamma_1$ lies in region 2. For the example shown in Figure 4.2, the vertex $O_1$ will move to the point D; thus $\delta_1^o, \gamma_1^o$ are identical to $\delta_1^p, \gamma_1^p$. Note that for all $\delta_1, \gamma_1$ which lie in region 2, $O_1$ will lie on the lines $O_0A$ or $O_0A'$. Note further that if $\delta_1, \gamma_1$ lies on the lines $O_0B$ or $O_0B'$, i.e. on the boundary between regions 2 and 3, the updating rules given above are consistent. The issue of consistency across the lines $O_0A$, $O_0A'$, i.e. the boundary between regions 1 and 2, is not as easily resolved as we note below.

Consider now the case where $\delta_1, \gamma_1$ lies in region 1. Three possible choices for the updating of the apex $O_1$ can be envisaged. The first choice is to put $\delta_1^o = \delta_1^p, \gamma_1^o = \gamma_1^p$, as shown in Figure 4.4a. This choice implies that the opening of the joint is the result of plastic deformation with the interface forces set to $N = S = 0$. The joint remains in contact; reversal of the deformation in the sense of opening and closing is not possible. The behaviour when $\delta_1, \gamma_1$ lies on $O_0A$ or $O_0A'$ is consistent, i.e. the updating rules for regions 1 and 2 give the same result.

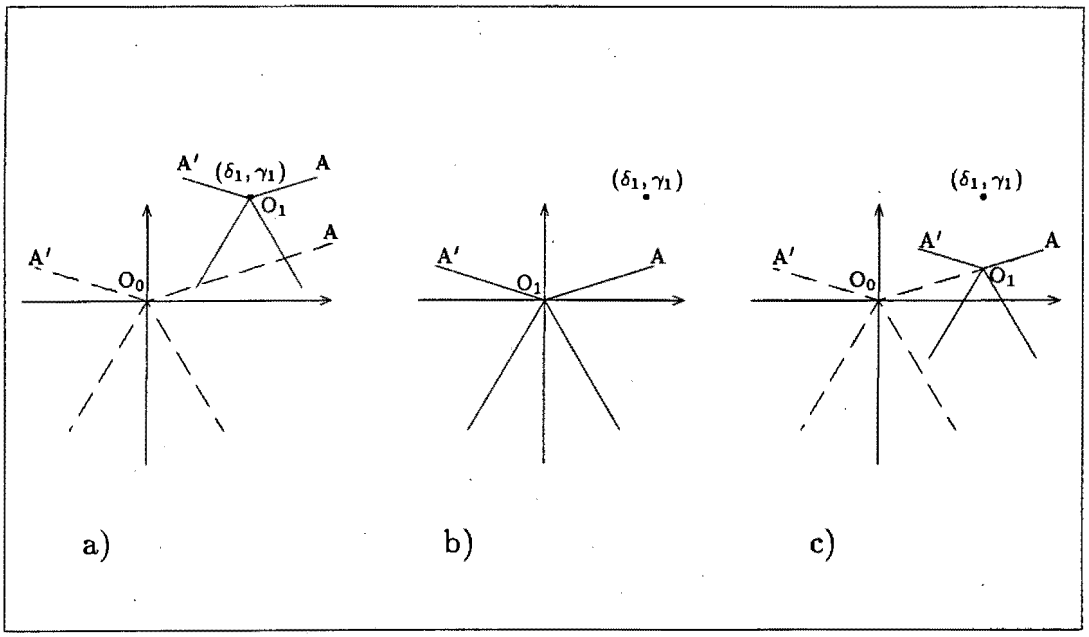



Figure 4.4: Possible updating rules associated with opening

A second possibility, which does permit loss of contact, is to leave the apex unchanged when $\delta_1, \gamma_1$ lies in region 1, with $\delta_1^o = \delta_0^o, \gamma_1^o = \gamma_0^o$, as shown in Figure 4.4b. Reversal of the deformation is possible, i.e. the joint can be opened and closed. However, if $\delta_1, \gamma_1$ lies on lines $O_0A$ or $O_0A'$, the updating rules are inconsistent since we obtain different results if we regard $O_0A$, $O_0A'$ as part of region 1 or as part of region 2. This inconsistency could have significant numerical consequences.

A third option is to adopt the rule shown in Figure 4.4c. Here we set $\gamma_1^o = \gamma_1^p, \delta_1^o = |\delta_1^p| \tan \mu$. The joint is open, in the sense that contact is lost, and further, we have consistency for $\delta_1, \gamma_1$ on $O_0A$ or $O_0A'$. If we open the joint without shear displacement ($\gamma_1 = 0$), reversal of the motion before contact is made is possible. If, however, we open the joint with shear displacement taking place, we can only partially reverse the motion before contact takes place.

The third option thus allows opening and maintains consistency, both of which are desirable attributes. However, it does predict some physically unreasonable behaviour. For example, if we open the joint without shear displacement, and then impose a programme of alternating shear displace-

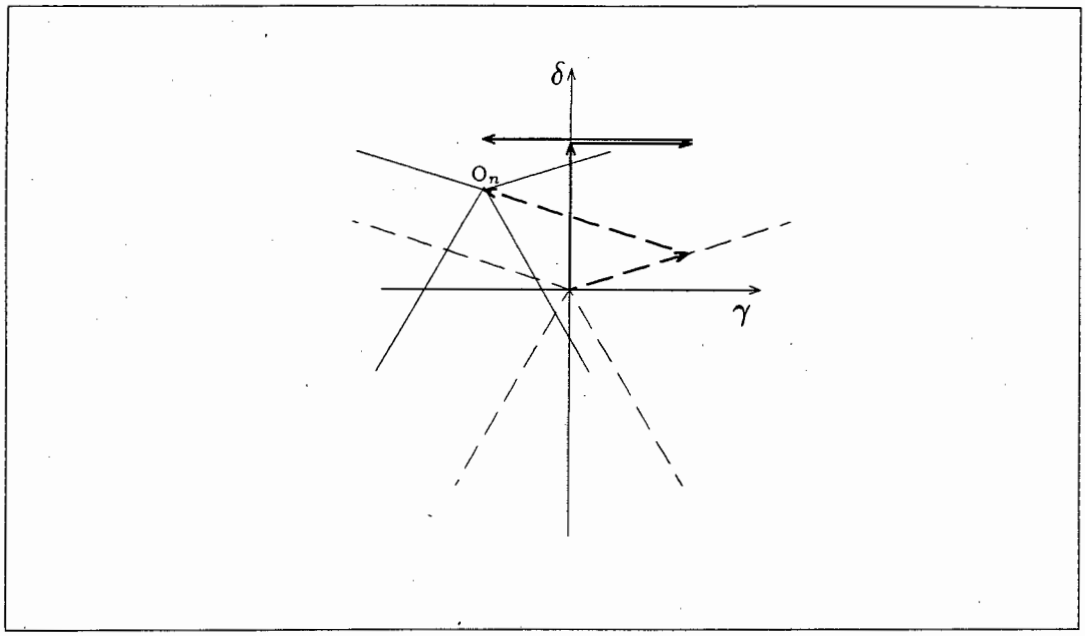



Figure 4.5: Alternating shear displacement programme at fixed $\delta$

ment at fixed $\delta$, the joint will close after a finite number of cycles. A shear displacement increment of any sign therfore causes a positive inelastic deformation of the joint. This is shown diagrammatically in Figure 4.5.

It would seem evident, from the comparison of the three possibilities, that the straightforward Coulomb model has an inherent problem associated with irreversible plastic deformation when $N = S = 0$. The third option represents a compromise which is internally consistent and which allows loss of contact to be represented to some extent. We shall choose the third option since it leads to a model which describes the joint reasonably well provided no significant reversal in shear displacement occurs.

### 4.1.3 Full Incremental Constitutive Equations

Including the position of vertex $\delta^\circ, \gamma^\circ$, and adopting the third option described in the previous section, we may now proceed to generalise equations (4.4), (4.5) and (4.6) directly to give explicit relations for an arbitrary increment.

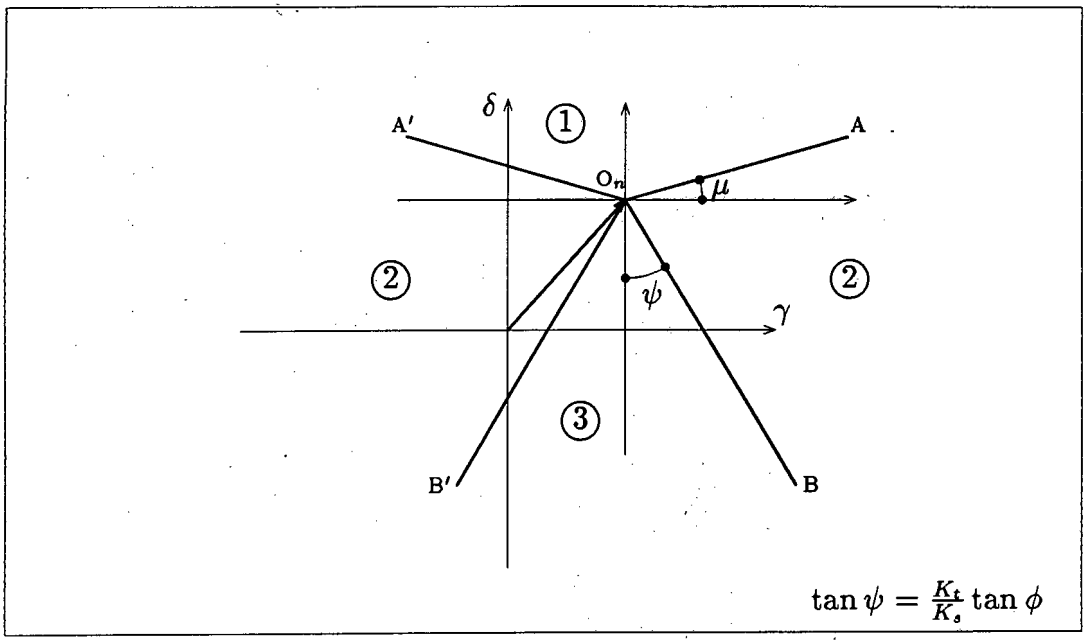



Figure 4.6: State diagram

In addition, we may relax the restriction that the elastic constraints $K_t$ and $K_s$ are equal. The lines $O_nB$ and $O_nB'$ will then be at angles $\psi$ to the negative $\delta$-axis, where

$$\tan\psi = \frac{K_t}{K_s}\tan\phi\,. \tag{4.7}$$

Generally, $K_t$ will be chosen to be larger than $K_s$ if they are different.

We may now write the explicit relations for the $(n+1)^{th}$ increment. At time $t_n$ we are given the displacements $\delta_n, \gamma_n$, the forces $N_n$, $S_n$, the plastic displacements $\delta_n^p$, $\gamma_n^p$ and the position of the vertex $\delta_n^o, \gamma_n^o$. We are further given increments in the total displacements $\Delta\delta_{n+1}, \Delta\gamma_{n+1}$, and hence new relative displacements $\delta_{n+1} = \delta_n + \Delta\delta_{n+1}$, $\gamma_{n+1} = \gamma_n + \Delta\gamma_{n+1}$. The relative displacement space is shown in Figure 4.6. The full set of relations are written as follows:
if

$$\delta_{n+1} \geq \delta_n^o + |\gamma_{n+1} - \gamma_n^o|\,\tan\mu\,,$$

then

$$\delta_{n+1}^e = \gamma_{n+1}^e = 0\,,$$

$$\begin{aligned}
\delta^p_{n+1} &= \delta_{n+1}\,, & \gamma^p_{n+1} &= \gamma_{n+1}\,,\\
\delta^o_{n+1} &= \delta^o_n + |\gamma_{n+1} - \gamma^o_n| \tan\mu, & \gamma^o_{n+1} &= \gamma_{n+1},\\
N_{n+1} &= S_{n+1} = 0. &&
\end{aligned} \tag{4.8}$$

If

$$\delta_{n+1} \le \delta^o_n - |\gamma_{n+1} - \gamma^o_n| \cot\psi\,,$$

then

$$\begin{aligned}
\delta^e_{n+1} &= \delta_{n+1} - \delta^o_n, & \gamma^e_{n+1} &= \gamma_{n+1} - \gamma^o_n,\\
\delta^p_{n+1} &= \delta^o_n\,, & \gamma^p_{n+1} &= \gamma^o_n\,,\\
\delta^o_{n+1} &= \delta^o_n\,, & \gamma^o_{n+1} &= \gamma^o_n\,,\\
N_{n+1} &= K_t\delta^e_{n+1}\,, & S_{n+1} &= K_s\gamma^e_{n+1}.
\end{aligned} \tag{4.9}$$

If

$$\delta^o_n + |\gamma_{n+1} - \gamma^o_n| \tan\mu \;>\; \delta_{n+1} \;>\; \delta^o_n - |\gamma_{n+1} - \gamma^o_n| \cot\psi\,,$$

put

$$a = \frac{(\delta_{n+1} - \delta^o_n)\sin\psi + |\gamma_{n+1} - \gamma^o_n|\cos\psi}{\cos(\psi - \mu)}\,, \qquad \text{and}$$

$$b = \frac{-(\delta_{n+1} - \delta^o_n)\cos\mu + |\gamma_{n+1} - \gamma^o_n|\sin\mu}{\cos(\psi - \mu)}\,,$$

then

$$\begin{aligned}
\delta^p_{n+1} &= \delta^o_n + a\sin\mu\,, & \gamma^p_{n+1} &= \gamma^o_n + a\cos\mu\ \mathrm{sgn}\,(\gamma_{n+1} - \gamma^o_n)\,,\\
\delta^e_{n+1} &= -b\cos\psi\,, & \gamma^e_{n+1} &= b\sin\psi\ \mathrm{sgn}\,(\gamma_{n+1} - \gamma^o_n)\,,\\
\delta^o_{n+1} &= \delta^p_{n+1}\,, & \gamma^o_{n+1} &= \gamma^p_{n+1}\,,\\
N_{n+1} &= K_t\delta^e_{n+1}\,, & S_{n+1} &= K_s\gamma^e_{n+1}\,.
\end{aligned} \tag{4.10}$$

### 4.1.4 Remarks

We have formulated a simple constitutive model where sliding is governed by Coulomb friction with an associated or non-associated flow rule in this section. The model is expressed in terms of three regions drawn in relative displacement space. Each region represents different behaviour, but no qual-

itative distinction is drawn between these regions. Such an approach greatly simplifies the contact problem.

Further, we realised in this section that the dilatant Coulomb model has an inherent problem associated with reversed shear deformation since an inelastic shear deformation increment of any sign is always accompanied by a positive increment of inelastic normal deformation. As a consequence, an initially open joint will close after a finite number of shear cycles when $\delta$ is held constant.

The explicit formulation includes the limiting cases $\mu = 0$, where sliding occurs without dilation, and the case $\mu = \phi$, where the flow rule is associated. In the case $\mu = 0$, it is clear that $\delta_n^o$ is always zero, so that the point $O_n$ simply move backwards and forwards along the $\gamma$-axis in relative displacement space. Further, the problems associated with reversal of shear displacement when the joint is open do not occur, since dilation does not take place. In the case of an associated flow rule, on the other hand, the consistent tangent stiffness matrix is symmetric and the incremental problem is convex. This will be demonstrated in the following chapter.

In the following section we show that the asperity model of Patton[34] can lead to a simple formulation which describes reversals of shear in a more reasonable manner; this approach has promise as a basis for a more sophisticated description of rock joints.

## 4.2 SAWTOOTH ASPERITY FORMULATION

### 4.2.1 The Yield Surfaces

We idealise the irregularities on the contact surfaces as sawtooth asperities which are perfectly seated in the virgin state, and oriented at angle $\mu$ to the plane of the joint. Figure 4.7b shows the contact surfaces after some inelastic deformation $\delta^p, \gamma^p$ has occurred. We note that once the joint has sheared some distance, its response to further sliding becomes sensitive to the direction of shearing, for it will now contract if the shear direction is reversed. Consider the case where sliding is in a direction which results in further dilation, and assume that the shear force on the asperity surfaces is

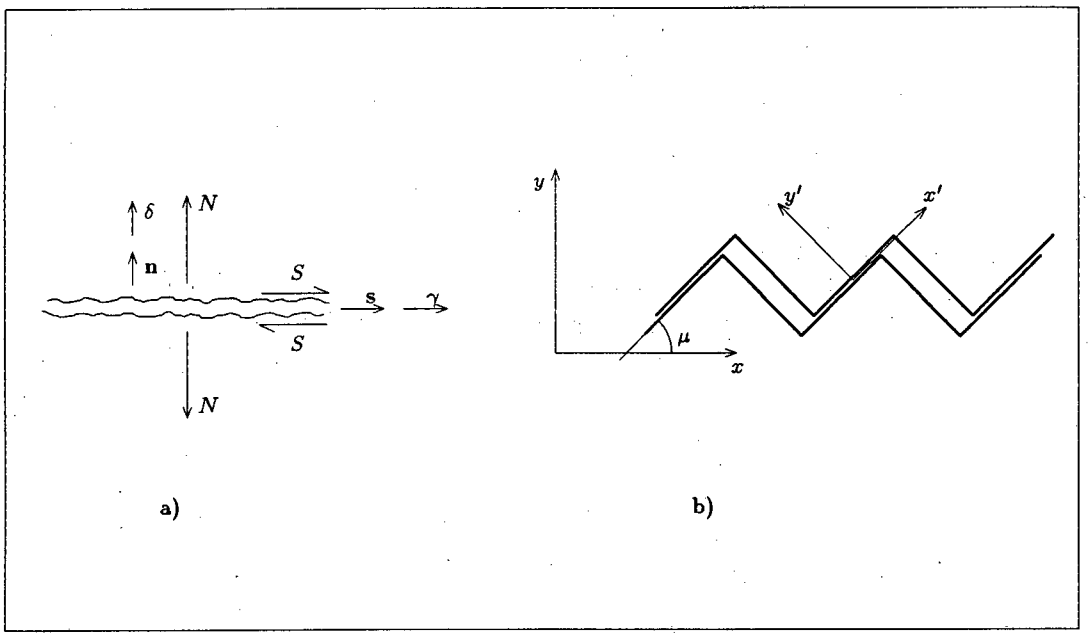



Figure 4.7: Joint surface idealisation

limited by a Coulomb friction law without any cohesion, hence

$$S' = -N' \tan \phi,$$

where $S'$ and $N'$ are respectively shear and normal forces in the asperity plane, with $\phi$ the angle of friction between the asperity surfaces.

After transforming $S'$ and $N'$ to the plane of the joint and some manipulation, we find the following relationship in terms of the conjugate forces across the plane of the joint:

$$S = -N \tan(\phi + \mu). \tag{4.11}$$

We recognise this relation as the basis of many other rock models, i.e. the models of Patton,[34] Barton[5,7] and Plesha,[35,36] and our choice is therefore fully justified. Consider next the case where the loading is applied in such a way that contraction takes place, and assume again that the shear force in the plane of the asperity surface is limited by a Coulomb friction law without cohesion, then

$$S' = N' \tan \phi \,,$$

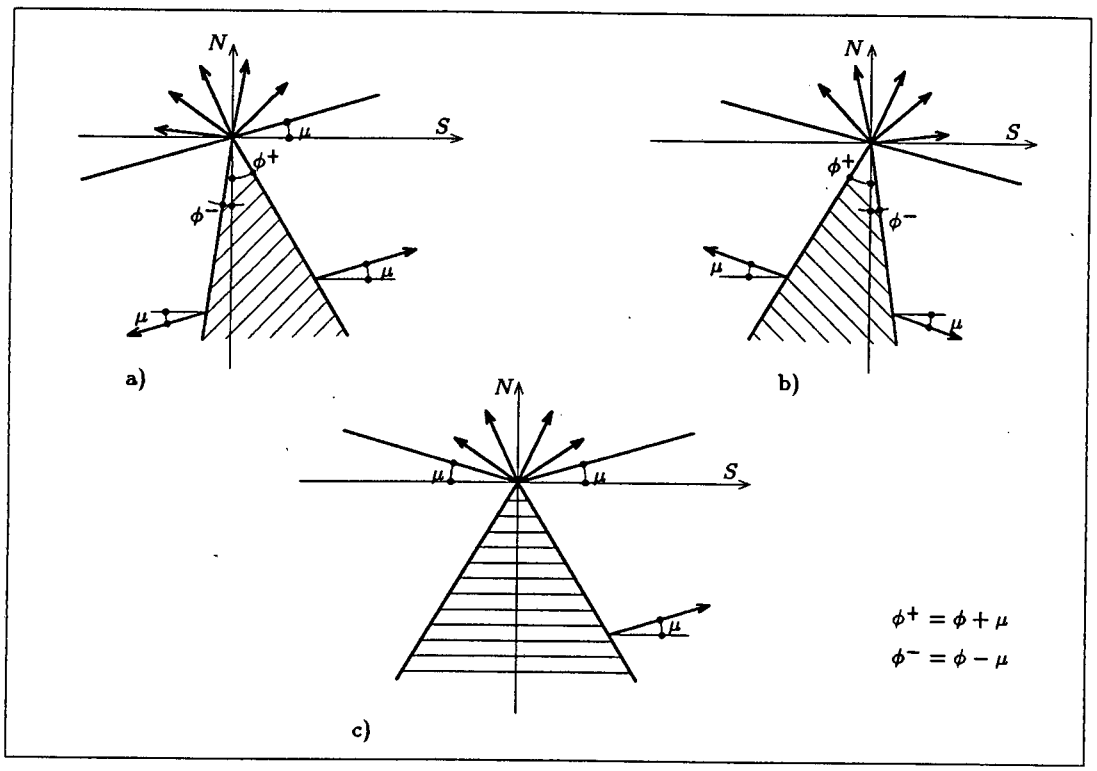



Figure 4.8: Sawtooth asperity yield surfaces

which simplifies to

$$S = N \tan(\phi - \mu) \tag{4.12}$$

after transformation to the plane of the joint. These yield surfaces, together with the direction of inelastic deformation, are shown in Figure 4.8a. The total angle between opposite sides of the yield surface is $2\phi$. The model is therefore equivalent to the Coulomb model except that it is oriented differently in force space; it is rotated through an angle $\mu$ with respect to the Coulomb model.

The state point $N, S$ is constrained to lie in the shaded area in Figure 4.8a, that is within or on the yield surface. If the state point lies within the yield surface, i.e. if

$$-N \tan(\phi + \mu) < S > N \tan(\phi - \mu) \,, \tag{4.13}$$

the inelastic relative displacement rates $\dot{\delta}^p, \dot{\gamma}^p$ are zero. If the state point lies

on any of these yield surfaces, i.e. if

$$S = \pm N \tan(\phi \mp \mu) \tag{4.14}$$

the inelastic relative displacement rates may be non-zero, and must satisfy the relation

$$\dot{\delta}^p = \dot{\gamma}^p \tan \mu \, . \tag{4.15}$$

We will refer to $\mu$ as the *dilation* or *geometric friction* angle. Physical considerations require that $0 \leq \mu \leq \phi$, and that $\phi + \mu < \frac{\pi}{2}$.

Figure 4.8b shows the yield surfaces for the case where the left hand side of the asperities are in contact, i.e. the mirror image of Figure 4.8a apply. We note that in the case where the joint is in its virgin state, contraction is impossible, and only the yield surfaces associated with dilation are active as shown in Figure 4.8c. The orientation of this yield surface is identical to the Coulomb model, but the angle between opposite sides of the yield surface is $2(\phi + \mu)$. However, once the joint has moved from the virgin state, either the yield surface shown in Figure 4.8a or Figure 4.8b applies, depending which sides of the asperities are in contact.

We note further that, as in the case of the Coulomb model, the inelastic deformation rate vector is not defined at the apices of these yield surfaces, but is constrained to lie in a fan bounded by lines drawn through the origin at angles $\pm\mu$ to the positive and negative $S$-axes, as shown in Figure 4.8.

### 4.2.2 Displacements in Relative Displacement Space

We now consider the corrector step problem. Suppose first that the joint is in its virgin state. This case, as we have pointed out, is equivalent to the Coulomb model and can thus be expressed in terms of three different regions drawn in relative displacement space (Fig 4.2). The only difference, assuming again that $K_t = K_s$ for the sake of geometric simplicity, is that lines $O_0B$ and $O_0B'$ are now drawn at an angle $\phi + \mu$ through the origin on either side of the negative $\delta$-axis as shown in Figure 4.9.

Each of these regions represent different contact conditions. For instance, if the state point $\delta_1, \gamma_1$ lies in region 1, bounded by lines $O_0A$, $O_0A'$ drawn through the origin at angle $\mu$ to the positive and negative $\gamma$-axes, separation of the surfaces will occur. The interface forces are zero, and hence the elastic components of $\delta_1, \gamma_1$ are both zero.

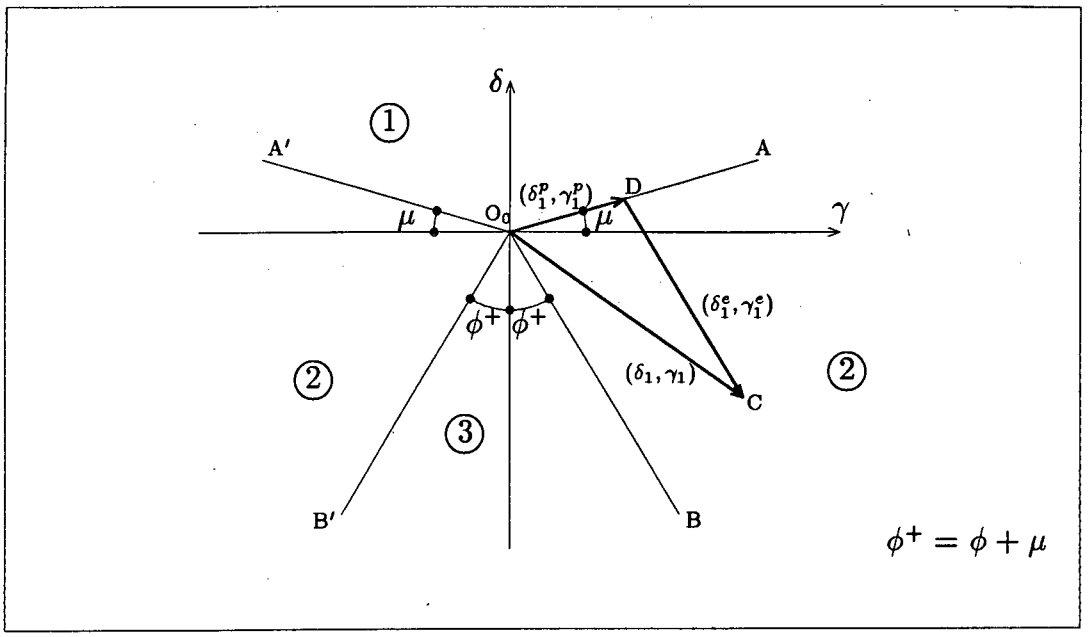



Figure 4.9: State diagram at joint virgin position

On the other hand, if the state point $\delta_1, \gamma_1$ lies in region 3, or on its boundaries, the response is entirely elastic, and the inelastic components of $\delta_1, \gamma_1$ are zero.

Finally, if the state point lies in the remaining areas of Figure 4.9, which we denote by region 2, both the elastic and inelastic components of $\delta_1, \gamma_1$ will be non-zero, and the interface forces $N_1, S_1$ will be non-zero and will lie on the yield surface. The joint has displaced from the virgin position, and either the left or right hand sides of the asperities remain in contact. If we consider the case where $\delta_1, \gamma_1$ is represented by the vector $O_0C$, we may locate a point D on line $O_0A$ by drawing DC parallel to $O_0B$. The inelastic component of $\delta_1, \gamma_1$ is then given by $O_0D$, and the elastic component by DC.

Displacements from the virgin position are therefore identical to those described for the Coulomb model. At the end of the first time-step the state diagram (Fig. 4.9) must be updated. The updating, however, differs from that of the Coulomb model. Suppose the joint displaces from the virgin position and that contact is now maintained only along the right hand side of the asperities. This state is defined by the yield surface shown in Figures 4.8a. Assuming still that $K_t = K_s$, we construct the updated state

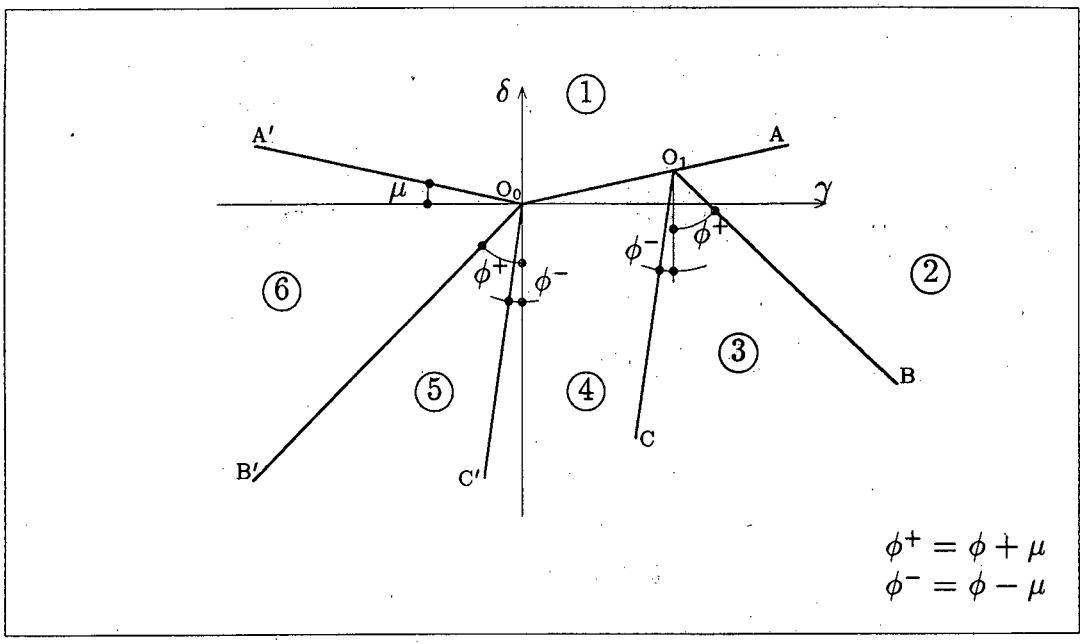



Figure 4.10: Updated State diagram

diagram (Fig. 4.10) by drawing lines $O_1B$ and $O_1C$ at angles $\phi+\mu$ and $\phi-\mu$ respectively. The apex $O_1$ of these lines, which we define by coordinates $\delta_1^o, \gamma_1^o$, is constrained to lie only along line $O_0A$, and must be updated at the end of each time-step. It then follows that $\delta_0^o = \gamma_0^o = 0$ for the virgin state. We note that in any subsequent time-step, the joint may, in addition to opening, further dilation or contraction, also move back to the virgin state or dilate with the opposite side of asperities in contact. We therefore complete the state diagram by drawing lines $O_0B'$ and $O_0C'$ through the origin at angles $\phi+\mu$ and $\phi-\mu$ respectively.

We may now consider the different updating rules for the case where the joint displaces from the virgin state. In the case where $\delta_1, \gamma_1$ lies in region 3 of Figure 4.9, the response is elastic, and the state diagram is unchanged. Hence $\delta_1^o = \delta_0^o, \gamma_1^o = \gamma_0^o$, and the yield surface associated with contraction remains inactive.

In the case where $\delta_1, \gamma_1$ lies in region 1, we set $\gamma_1^o = \gamma_1$ and $\delta_1^o = |\delta_1| \tan\mu$. That is, we move $\delta_1^o, \gamma_1^o$ to a point on line $O_0A$ which is located directly below $\delta_1, \gamma_1$. We choose this updating rule to ensure consistent updating rules when $\delta_1, \gamma_1$ lies on the boundary between regions 1 and 2 as explained

in the previous section.

Finally, consider the case where $\delta_1, \gamma_1$ lies in region 2. For the example shown in Figure 4.9, $\delta_1^o, \gamma_1^o$ are identical to $\delta_1^p, \gamma_1^p$, and we update the apex to point D along line $O_0$A.

We note that the joint can in now respond in any one of six possible ways, depending on the region in which the next update of state point lies. Consider an arbitrary increment, say the $(n+1)^{th}$ increment. Suppose that the joint has displaced from the virgin position, that the right hand side of the asperities are now in contact, and that the vertex $\delta_n^o, \gamma_n^o$ has been updated to a point $O_n$ along $O_0$A as indicated in Figure 4.11. If the state point $\delta_{n+1}, \gamma_{n+1}$, denoted by point $E$, lies in

- region 1; the faces of the joint are separated, the interface forces are zero, and hence the elastic components of $\delta_{n+1}, \gamma_{n+1}$ are both zero. We update the apex $\delta_{n+1}^o, \gamma_{n+1}^o$ to a point $D$ directly below $\delta_{n+1}, \gamma_{n+1}$ as indicated in Figure 4.12a.

- region 2; further dilation takes place with the right hand sides of the asperities in contact. If $\delta_{n+1}, \gamma_{n+1}$ is represented by the vector $O_0$E, we may locate a point D on line $O_0$A by drawing DE parallel to $O_0$B. The inelastic component of $\delta_{n+1}, \gamma_{n+1}$ is then given by $O_0$D, and the elastic component by DE as illustrated in Figure 4.12b. The apex of the state diagram is updated to point D.

- region 3; the response is elastic. No changes in the inelastic part of deformation, and the apex $\delta_{n+1}^o, \gamma_{n+1}^o$ remains at the current position $\delta_n^o, \gamma_n^o$ (Fig. 4.12c).

- region 4; the faces of the joint contract with the right hand sides of the asperities in contact. We find the elastic and inelastic parts of deformation as before, except, we now construct a line through the state point parallel to line $O_n$C (Fig. 4.12d). The apex is updated to the intercept of line DE and line $O_0$A.

- region 5; the joint moves back to the virgin state. The inelastic parts of deformation are set to zero as shown in Figure 4.12e, the apex $(\delta_{n+1}^o, \gamma_{n+1}^o)$ is updated to the origin $O_0$, and the yield surface associated with contraction (lines $O_0$C′ and $O_n$C) is inactive. Region 4 thus contracts to a line that lies within the elastic region, and the Coulomb model with a total friction angle $(\phi + \mu)$ is recovered.

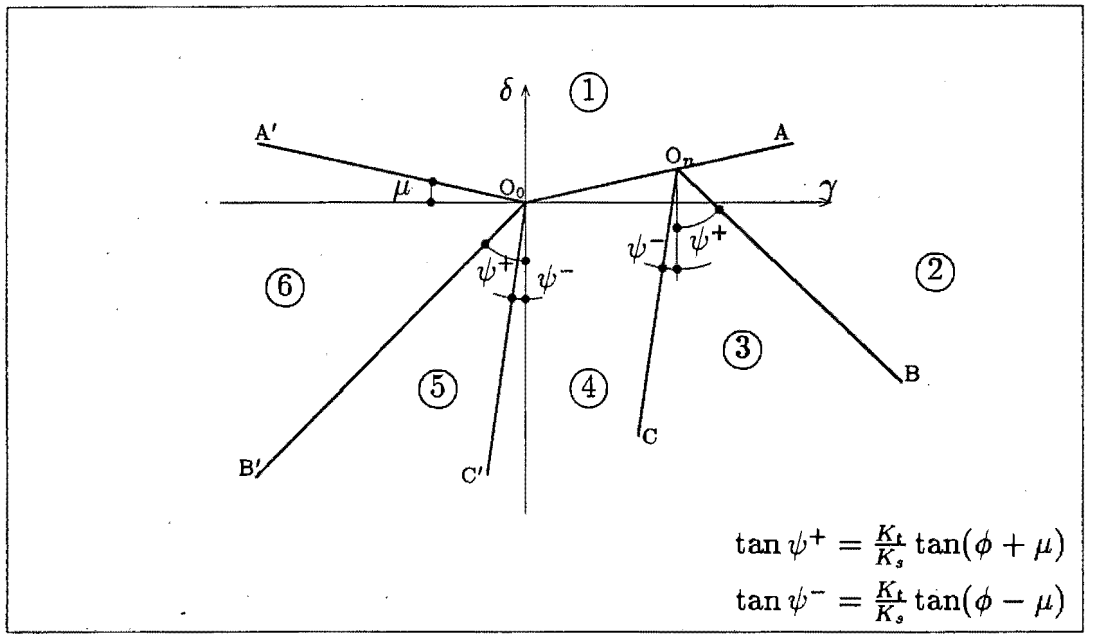



Figure 4.11: State diagram

- region 6; the joint first contracts to the virgin state, the left hand sides of the asperities then come into contact, and dilation takes place. The elastic and inelastic deformation are obtained as for region 2 (Fig. 4.12f). Since the left hand side of the asperities are now in contact, the mirror image of Figure 4.11 applies, with the apex $O_{n+1}$ updated to point D.

### 4.2.3 Full Incremental Equations

We may now relax the restriction that the elastic constraints $K_t$ and $K_s$ are equal, and formally write the equations for the corrector step problem. The lines $O_nB$ and $O_nB'$ will then be at angles $\psi^+$ to a vertical line, where

$$\tan\psi^+ = \frac{K_t}{K_s}\tan(\phi+\mu)\,, \tag{4.16}$$

and lines $O_nC$ and $O_nC'$ at angles $\psi^-$ to a vertical line, where

$$\tan\psi^- = \frac{K_t}{K_s}\tan(\phi-\mu)\,. \tag{4.17}$$

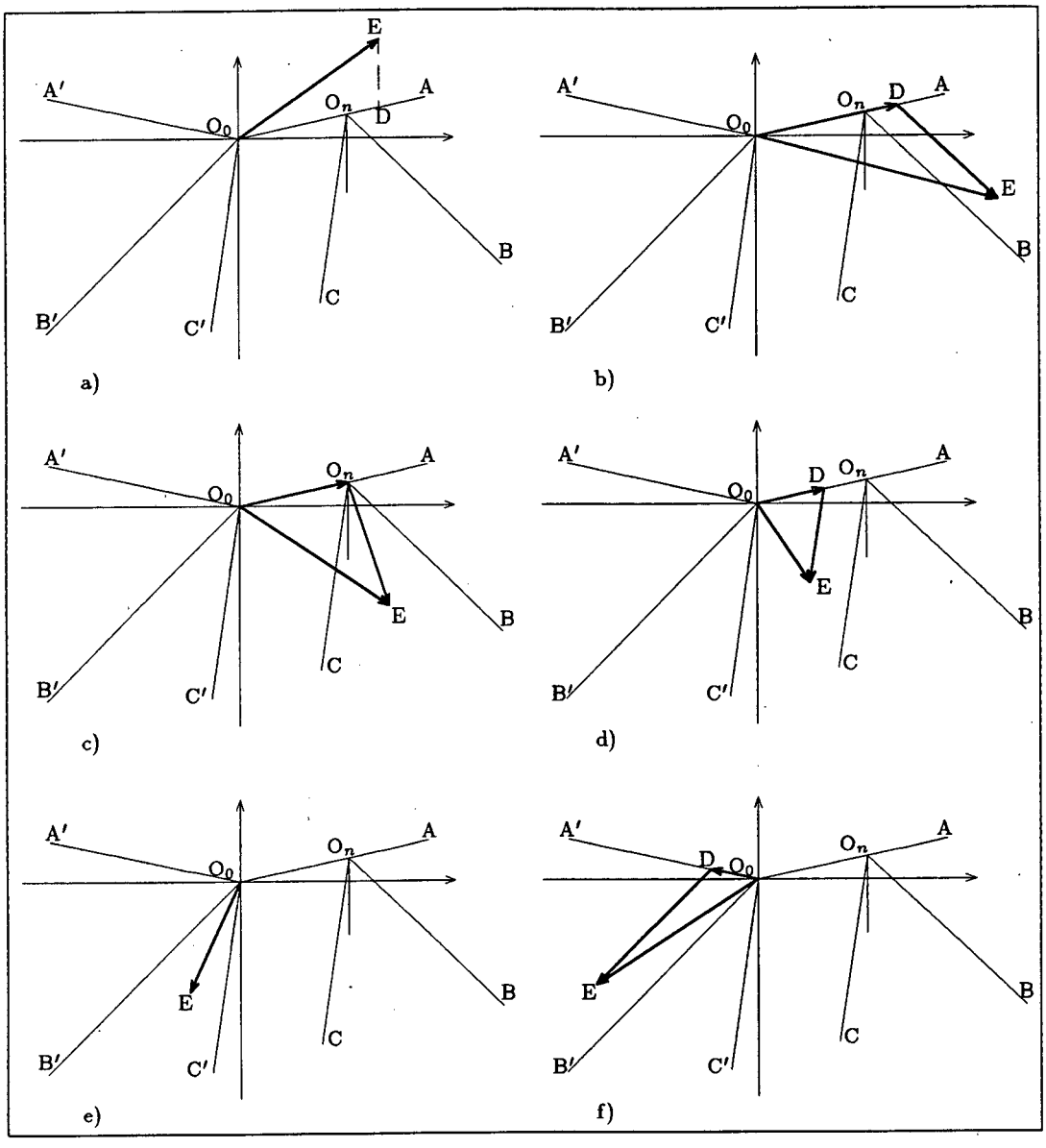



Figure 4.12: Joint response in the six different regions

The corrector step problem for an arbitrary increment, say the $(n+1)^{th}$ increment, is posed as follows: At the $n^{th}$ increment we are given relative displacements $\delta_n, \gamma_n$, the conjugate forces $N_n, S_n$, and the inelastic deformations $\delta_n^p, \gamma_n^p$. Suppose that the right hand side of the asperities are in contact, and that the vertex $\delta_n^o, \gamma_n^o$ has been updated to point $O_n$ as indicated in Figure 4.11. We are further given increments in the total displacements $\Delta\delta_{n+1}, \Delta\gamma_{n+1}$, and hence new relative displacements $\delta_{n+1} = \delta_n + \Delta\delta_{n+1}$, $\gamma_{n+1} = \gamma_n + \Delta\gamma_{n+1}$. Adopting a positive sign for zero, we write the full set of relations, which cater for the two mirror images as well as displacements from the virgin state, as follows:
if

$$\delta_{n+1} \geq |\gamma_{n+1}| \tan\mu \,,$$

the surfaces are separated. Then

$$\begin{aligned}
\delta_{n+1}^e &= \gamma_{n+1}^e = 0 \,, \\
\delta_{n+1}^p &= \delta_{n+1} \,, & \gamma_{n+1}^p &= \gamma_{n+1} \,, \\
\delta_{n+1}^o &= |\gamma_{n+1}| \tan\mu \,, & \gamma_{n+1}^o &= \gamma_{n+1} \,, \\
N_{n+1} &= S_{n+1} = 0 \,.
\end{aligned} \tag{4.18}$$

If

$$\begin{aligned}
\delta_{n+1} &< \gamma_{n+1}\,\mathrm{sgn}(\gamma_n^o)\,\tan\mu \,, & &\text{and} \\
\delta_{n+1} &> \delta_n^o - (\gamma_{n+1} - \gamma_n^o)\,\mathrm{sgn}(\gamma_n^o)\,\cot\psi^+ \,,
\end{aligned}$$

further dilation takes place. Put

$$\begin{aligned}
a &= \frac{(\delta_{n+1} - \delta_n^o)\,\sin\psi^+ + |\gamma_{n+1} - \gamma_n^o|\,\cos\psi^+}{\cos(\psi^+ - \mu)} \,, \quad \text{and} \\
b &= \frac{-(\delta_{n+1} - \delta_n^o)\,\cos\mu + |\gamma_{n+1}|\,\sin\mu}{\cos(\psi^+ - \mu)} \,,
\end{aligned}$$

then

$$\begin{aligned}
\delta_{n+1}^p &= \delta_n^o + a\sin\mu \,, & \gamma_{n+1}^p &= \gamma_n^o + a\cos\mu\ \mathrm{sgn}\,(\gamma_{n+1} - \gamma_n^o) \,, \\
\delta_{n+1}^e &= -b\cos\psi^+ \,, & \gamma_{n+1}^e &= b\sin\psi^+\ \mathrm{sgn}\,(\gamma_{n+1} - \gamma_n^o) \,, \\
\delta_{n+1}^o &= \delta_{n+1}^p \,, & \gamma_{n+1}^o &= \gamma_{n+1}^p \,, \\
N_{n+1} &= K_t\delta_{n+1}^e \,, & S_{n+1} &= K_s\gamma_{n+1}^e \,.
\end{aligned} \tag{4.19}$$

If

$$\begin{aligned} \delta_{n+1} &\leq \delta_n^o - (\gamma_{n+1} - \gamma_n^o)\ \mathrm{sgn}(\gamma_n^o)\ \cot\psi^+ , \qquad \text{and} \\ \delta_{n+1} &\leq \delta_n^o + (\gamma_{n+1} - \gamma_n^o)\ \mathrm{sgn}(\gamma_n^o)\ \cot\psi^- , \end{aligned}$$

the response is elastic, hence

$$\begin{aligned} \delta_{n+1}^e &= \delta_{n+1} - \delta_n^o , & \gamma_{n+1}^e &= \gamma_{n+1} - \gamma_n^o , \\ \delta_{n+1}^p &= \delta_n^o , & \gamma_{n+1}^p &= \gamma_n^o , \\ \delta_{n+1}^o &= \delta_n^o , & \gamma_{n+1}^o &= \gamma_n^o , \\ N_{n+1} &= K_t \delta_{n+1}^e , & S_{n+1} &= K_s \gamma_{n+1}^e . \end{aligned} \tag{4.20}$$

If

$$\begin{aligned} \delta_{n+1} &< |\gamma_{n+1}|\ \tan\mu , \\ \delta_{n+1} &> \delta_n^o + (\gamma_{n+1} - \gamma_n^o)\ \mathrm{sgn}(\gamma_n^o)\ \cot\psi^- , \qquad \text{and} \\ \delta_{n+1} &< \gamma_{n+1}\ \mathrm{sgn}(\gamma_n^o)\ \cot\psi^- , \end{aligned}$$

the joint contracts, hence put

$$\begin{aligned} a &= \frac{(\delta_{n+1} - \delta_n^o)\ \sin\psi^- + |\gamma_{n+1} - \gamma_n^o|\ \cos\psi^-}{\cos(\psi^- + \mu)} , \qquad \text{and} \\ b &= \frac{-(\delta_{n+1} - \delta_n^o)\ \cos\mu - |\gamma_{n+1} - \gamma_n^o|\ \sin\mu}{\cos(\psi^- + \mu)} , \end{aligned}$$

then

$$\begin{aligned} \delta_{n+1}^p &= \delta_n^o - a\sin\mu , & \gamma_{n+1}^p &= \gamma_n^o + a\cos\mu\ \mathrm{sgn}\,(\gamma_{n+1} - \gamma_n^o) , \\ \delta_{n+1}^e &= -b\cos\psi^- , & \gamma_{n+1}^e &= b\sin\psi^-\ \mathrm{sgn}\,(\gamma_{n+1} - \gamma_n^o) , \\ \delta_{n+1}^o &= \delta_{n+1}^p , & \gamma_{n+1}^o &= \gamma_{n+1}^p , \\ N_{n+1} &= K_t \delta_{n+1}^e , & S_{n+1} &= K_s \gamma_{n+1}^e . \end{aligned} \tag{4.21}$$

If

$$\begin{aligned} \delta_{n+1} &\geq \gamma_{n+1}\ \mathrm{sgn}(\gamma_n^o)\ \cot\psi^- , \qquad \text{and} \\ \delta_{n+1} &\leq \gamma_{n+1}\ \mathrm{sgn}(\gamma_n^o)\ \cot\psi^+ , \end{aligned}$$

the joint moves back to the virgin position, hence

$$\delta_{n+1}^e = \delta_{n+1} , \qquad \gamma_{n+1}^e = \gamma_{n+1} ,$$

$$\begin{aligned} \delta^p_{n+1} &= \gamma^p_{n+1} = 0 \,, \\ \delta^o_{n+1} &= \gamma^o_{n+1} = 0 \,, \\ N_{n+1} &= K_t \delta^e_{n+1} \,, \qquad\qquad S_{n+1} = K_s \gamma^e_{n+1}. \end{aligned} \tag{4.22}$$

If

$$\begin{aligned} \delta_{n+1} &< |\gamma_{n+1}| \ \tan\mu \,, \qquad\qquad \text{and} \\ \delta_{n+1} &> \gamma_{n+1} \ \mathrm{sgn}(\gamma^o_n) \ \cot\psi^+ \,, \end{aligned}$$

the opposite sides of the asperities come in contact, hence put

$$\begin{aligned} a &= \frac{\delta_{n+1} \ \sin\psi^+ + |\gamma_{n+1}| \ \cos\psi^+}{\cos(\psi^+ - \mu)} \,, \qquad \text{and} \\ b &= \frac{-\delta_{n+1} \ \cos\mu + |\gamma_{n+1}| \ \sin\mu}{\cos(\psi^+ - \mu)} \,, \end{aligned}$$

then

$$\begin{aligned} \delta^p_{n+1} &= a \sin\mu \,, & \gamma^p_{n+1} &= -a \cos\mu \ \mathrm{sgn}\,(\gamma^o_n) \,, \\ \delta^e_{n+1} &= -b \cos\psi^+ \,, & \gamma^e_{n+1} &= -b \sin\psi^+ \ \mathrm{sgn}\,(\gamma^o_n) \,, \\ \delta^o_{n+1} &= \delta^p_{n+1} \,, & \gamma^o_{n+1} &= \gamma^p_{n+1} \,, \\ N_{n+1} &= K_t \delta^e_{n+1} \,, & S_{n+1} &= K_s \gamma^e_{n+1} \,. \end{aligned} \tag{4.23}$$

### 4.2.4 Remarks

The corrector step equations for the sawtooth asperity model also include the limiting case, $\mu = 0$, where sliding occurs without dilation. In this case $\delta^o_n$ is always zero, so that the point $O_n$ simply moves backwards and forwards along the $\gamma$-axis in relative displacement space. Furthermore, the three yield surfaces which constitute the asperity model are identical, and also identical to the Coulomb model without dilation.

We therefore deduce that the main difference between the Coulomb model and the asperity model is the way in which inelastic normal deformation is treated. In the asperity model the joint is sensitive to the direction of sliding, with dilation or contraction a possibility. In the Coulomb model an

increment in inelastic shear deformation of any sign is always accompanied by a positive increment in inelastic normal deformation. Moreover, the apex of the asperity model is constrained to lie on lines $O_0A'$,$O_0A$ only, whereas the apex of the Coulomb model may lie anywhere in region one (defined for the virgin state), as well as on its boundaries.

In both formulations the inelastic relative displacements rates are governed by a linear relation of the form $\dot{\delta}^p = \dot{\gamma}^p \tan\mu$. The dilation $\delta^p$ therefore increases monotonically with increasing inelastic shear deformation. This relation further implies a linear boundary in relative displacement space between contact and no-contact conditions. Rock asperities, however, have finite amplitude, finite strength and degrade with continuing sliding.

We formulate a model in the following section in which the dilation ceases when the normal deformation reaches the height of the asperities.

## 4.3 LOGARITHMIC SPIRAL MODEL

We extend the concepts posed in the previous section by adapting the friction laws formulated for the asperity model to include the dependence of the dilation angle $\mu$ on the amount of shearing, i.e. dilation at time-step $t_n$ takes place when

$$|S| = -N\tan(\phi + \mu_n)\,, \tag{4.24}$$

and contraction when

$$|S| = -N\tan(\phi - \mu_n)\,. \tag{4.25}$$

As before, physical considerations require that $0 \leq \mu_n \leq \phi$, and that $\phi + \mu_n < \frac{\pi}{2}$. We further replace the straight line boundary between contact and no-contact regions with a curve which represents a non-linear relation between $\dot{\delta}^p$ and $\dot{\gamma}^p$. We choose a curve so that further inelastic normal deformation $\delta^p$ ceases when the maximum dilation $\delta^m$ is reached. The dilation angle $\mu_n$ at time-step $t_n$ is then defined as the gradient of the inelastic displacement rate relation at $\dot{\delta}^p_n, \dot{\gamma}^p_n$. This means that the orientation of the yield surface changes from one time-step to the next. Moreover, the direction of the elastic deformation vector (oriented at an angle $\phi \pm \mu_n$) changes, as shown in the relative displacement diagram of Figure 4.13. Since the direction of neither the elastic nor the inelastic displacement vectors are known at the beginning of the time-step, an explicit formulation is no longer possible. Furthermore, the formulation leads to an ambiguity as indicated; the total deformation is

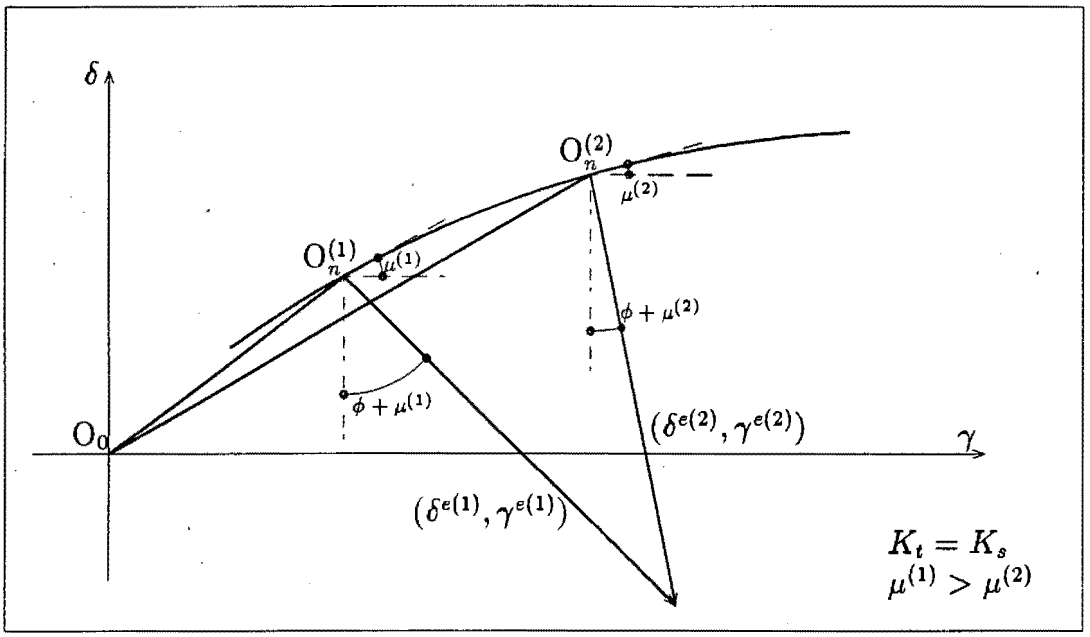



Figure 4.13: Possible solutions associated with non-linear $\delta^p, \gamma^p$ relation

not always uniquely divided into elastic and inelastic parts.

Suppose we choose a $\dot{\delta}^p, \dot{\gamma}^p$ relation such that the resulting family of elastic vectors, oriented at an angle $\phi \pm \mu_n$, lie along lines which intersect at one point. If this is the case, the ambiguity is restricted to one point only. We will choose this point to lie far away from the region of interest, i.e. if the state point approaches this point, the behaviour becomes physically unrealisable. Further, this point will be defined on the initial yield surface. An explicit model can then be formulated since the direction of elastic vectors are known; they lie along radial lines emanating from one point.

These ideas are possible if we choose a segment of a logarithmic or equiangular spiral to represent the boundary between open and contact states in relative displacement space. The origin of the spiral will be defined to lie on the initial yield surface, and the elastic vectors will then lie along lines emanating from the spiral origin. This approach, however, requires that the angle between a line drawn through the origin of the spiral and a line tangent to the spiral be a constant, which is only true if we restrict the elastic constants $K_t$ and $K_s$ to be equal. The formulation nevertheless illustrates some of the features inherent to any model in which both the yield surface

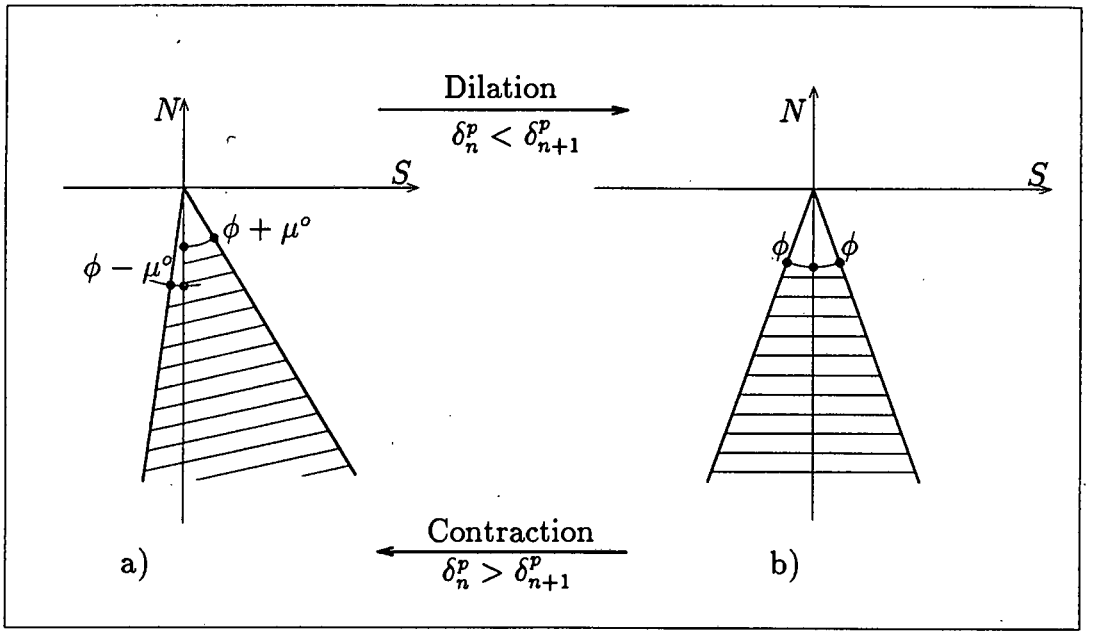



Figure 4.14: Limiting yield surfaces

and flow potential change from one time-step to the next.

Figure 4.14a shows the orientation of the yield surface at the onset of yielding with $\gamma^p > 0$. The dilation angle at this point, which we define as the initial dilation angle $\mu^o$, is a maximum, hence $\mu_n = \mu^o$. The yield surface at this stage is equivalent to the sawtooth asperity yield surface with the right hand sides of the asperities in contact (Fig. 4.8a). Figure 4.14b shows the orientation of the yield surface when the point of maximum dilation $\delta^m$ is reached (i.e. $\delta^p = \delta^m$ and $\mu_n = 0$), and the Coulomb yield surface is recovered. Since the angle between the radial generator and a line tangent to the logarithmic spiral remains constant, the gradient $\mu_n$ can be expressed as a linear function of $\theta$ only; $\theta$ is the angular component of polar coordinates $R, \theta$. The yield surface therefore rotates in such a way that $\mu_n$ decreases as a linear function of $\theta$, from $\mu_n = \mu^o$ to $\mu_n = 0$, between the two limit cases depicted.

It is evident from these diagrams that both dilation and contraction result in behaviour which is characteristic of a softening material. For instance, shearing a joint under conditions of constant normal stress (i.e. the joint is free to dilate) results in shear force-shear deformation response *oabc* shown in Figure 4.15. In the limiting case, where a relative displacement estimate

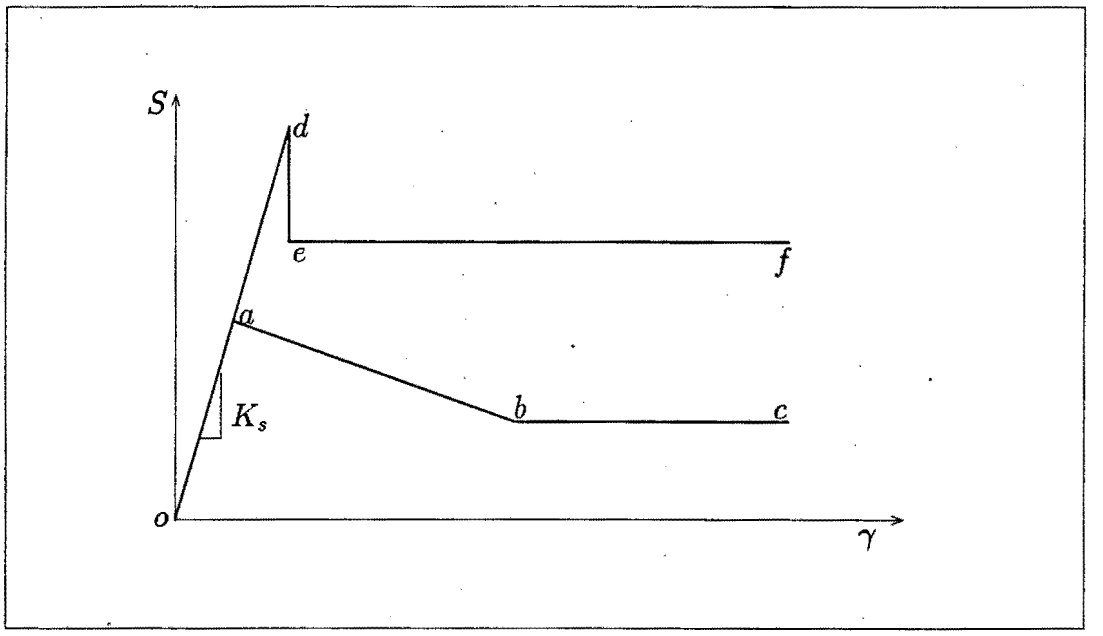



Figure 4.15: Shear stress-shear deformation response

$\delta_{n+1}, \gamma_{n+1}$ at any time-step $t_{n+1}$ coincides with the origin of the logarithmic spiral, an infinite number of solutions are possible, resulting in response *ode f*. However, we can avoid predictions of the state point near the origin of the spiral by choosing the elastic constants $K_t, K_s$ sufficiently large.

### 4.3.1 Construction of State Diagram

In constructing the state diagram in relative displacement space, we first consider the joint in its virgin position, with the yield surface associated with contraction inactive.

Setting $K_t = K_s$, we draw line $O^oB$,$O^oB'$ at an angle $\phi + \mu^o$ through the origin on either side of the negative $\delta$-axis, as shown in Figure 4.16. We then locate a point $D^b$,$D'^b$ (defined by coordinates $\delta^b, \pm\gamma^b$) on line $O^oB$,$O^oB'$ at such a position that a segment of the logarithmic spiral, constructed with its origin at $D^b$,$D'^b$, intersects line $O^oB$, $O^oB'$ at $\delta = \gamma = 0$, and draw a local coordinate system $x', y'$ at this point. With respect to this local coordinate system, we draw a segment of the logarithmic spiral curve through point $O^o$

and $O^a$,$O'^a$, where $O^a$,$O'^a$ (defined by coordinates $\delta^a, \pm\gamma^a$) is the point where the gradient of the spiral becomes zero. The remaining boundary between the contact and no-contact regions consists of a horizontal line segment drawn at $\delta = \delta^a$. The logarithmic spiral is given by the following relation in polar coordinates:

$$R = e^{s\theta} \,, \tag{4.26}$$

where $R$ is the radial component and $\theta$ the angular component of a point on the spiral with respect to the $x', y'$ coordinate system. Positive $\theta$ is measured anti-clockwise if $\gamma^b > 0$ and clockwise if $\gamma^b < 0$. $s = \cot\varphi$ is a constant, with $\varphi$ the angle between lines drawn through the origin and tangent to the spiral as shown in Figure 4.16. From geometric considerations it is evident that this angle is always $\varphi = \frac{\pi}{2} - \phi$ at any point $R, \theta$ along the spiral segment. It then follows that

$$s = \cot(\frac{\pi}{2} - \phi) = \tan\phi \,. \tag{4.27}$$

Furthermore, the gradient of the spiral is a linear function of $\theta$, given by

$$\begin{aligned} \mu &= \theta - \phi \\ &= \mu^o \qquad \text{when} \qquad \delta^p = 0 \\ &= 0 \qquad \text{when} \qquad \delta^p = \delta^a \,. \end{aligned} \tag{4.28}$$

Referring to Figure 4.16, we note that line $O^oB$ is located at angle

$$\theta^{ob} = \phi + \mu^o \tag{4.29}$$

with respect to the $x', y'$ coordinate system, and hence at distance

$$R^{ob} = e^{s\theta^{ob}} = e^{(\phi+\mu^o)\tan\phi} \tag{4.30}$$

from the origin of this coordinate system. The origin $D^b$,$D'^b$ of the spiral is then given by the following Cartesian coordinates in relative displacement space:

$$\begin{aligned} \delta^b &= -R^{ob}\cos(\phi+\mu^o) &&= -e^{[(\phi+\mu^o)\tan\phi]}\cos(\phi+\mu^o) \,, \\ \gamma^b &= \pm R^{ob}\sin(\phi+\mu^o) &&= \pm e^{[(\phi+\mu^o)\tan\phi]}\sin(\phi+\mu^o) \,. \end{aligned} \tag{4.31}$$

We next draw a line through the origin of the spiral and point $O^a$,$O'^a$, and denote the radial component $R^{ab}$, and the angular component $\theta^{ab}$, with

$$\begin{aligned} \theta^{ab} &= \phi \,, \qquad \text{and} \\ R^{ab} &= e^{\theta^{ab}\tan\phi} = e^{\phi\tan\phi} \,. \end{aligned} \tag{4.32}$$

The position of this point in terms of Cartesian coordinates in relative dis-

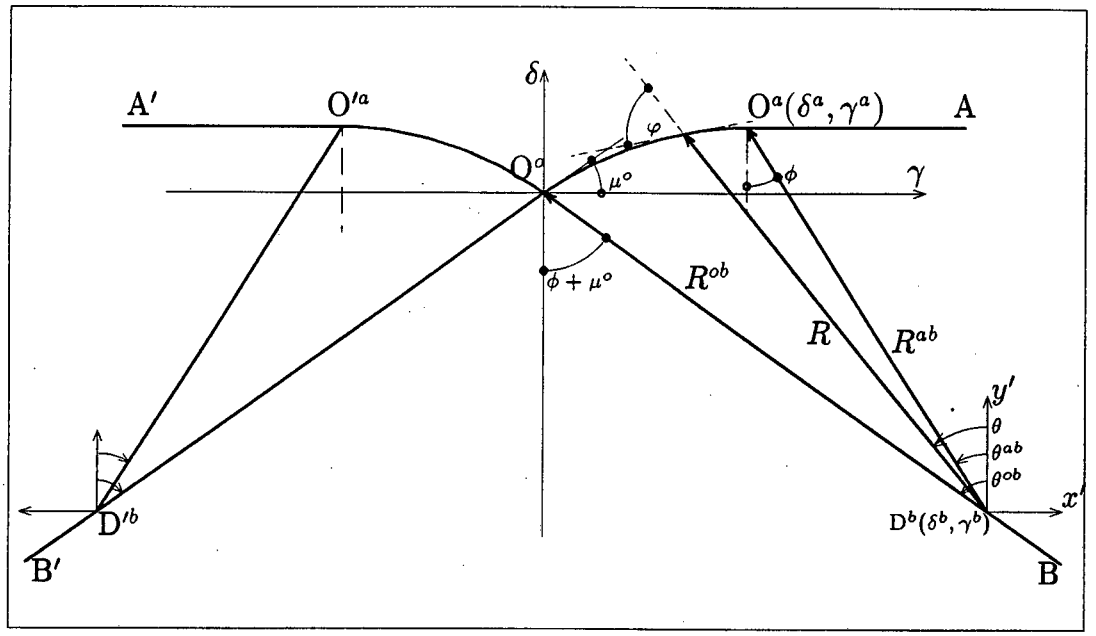



Figure 4.16: Construction of state diagram

placement space is

$$\begin{aligned}
\delta^a &= \delta^b + R^{ab}\cos\phi \\
&= -e^{[(\phi+\mu^o)\tan\phi]}\cos(\phi+\mu^o) + e^{[\phi\tan\phi]}\cos\phi\,, \\
\gamma^a &= \gamma^b \pm R^{ab}\sin\phi \\
&= \pm\left(e^{[(\phi+\mu^o)\tan\phi]}\sin(\phi+\mu^o) - e^{[\phi\tan\phi]}\sin\phi\right)\,. \qquad (4.33)
\end{aligned}$$

The shape of the spiral, its origin $D'^b$,$D^b$, and the *inflection point* $O'^a$,$O^a$ (i.e. the point where the spiral and horizontal line meets) are thus completely expressed in terms of the friction angle $\phi$ and the initial dilation angle $\mu^o$ only. In other words, different combinations of $\phi$ and $\mu^o$ will yield different spiral segments, and hence different positions $\delta^a, \pm\gamma^a$. We constrain the shape of the spiral segment further by scaling the segment so that $\delta^a$ coincides with the point of maximum dilation $\delta^m$. Then, introducing a dimensionless scaling factor $\zeta^b$, given by the ratio

$$\begin{aligned}
\zeta^b &= \frac{\delta^m}{\delta^a} \\
&= \delta^m[e^{\phi\tan\phi}\cos\phi - e^{(\phi+\mu^o)\tan\phi}\cos(\phi+\mu^o)]^{-1}\,, \qquad (4.34)
\end{aligned}$$

we redefine the origin of the spiral $D^b$,$D'^b$ at $\zeta^b\delta^b, \pm\zeta^b\gamma^b$, and the inflection

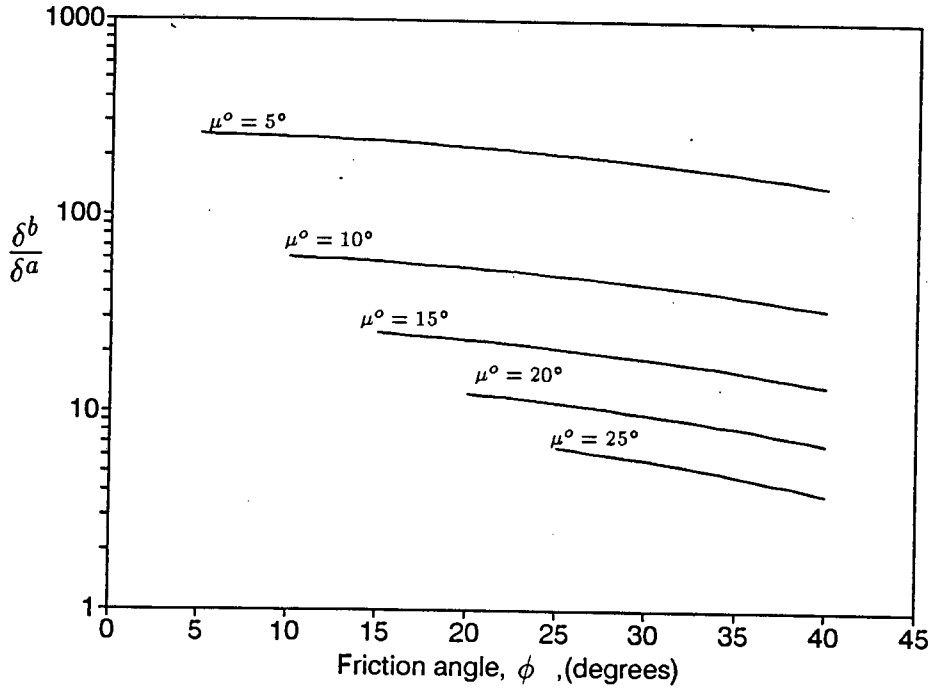



Figure 4.17: Influence of $\phi$ and $\mu^o$ on the vertical component of the spiral origin

point $O'^a$,$O^a$ at $\zeta^b\delta^a = \delta^m, \pm\zeta^b\gamma^a$. The segment of the spiral is then given by

$$R = \zeta^b e^{[\theta \tan\phi]}, \qquad \phi \leq \theta \leq \phi + \mu^o \tag{4.35}$$

in polar coordinates with respect to the local coordinate system, or by

$$\begin{aligned} \delta &= \zeta^b\delta^b + R\cos\theta \\ &= -\zeta^b\left(e^{[(\phi+\mu^o)\tan\phi}\cos(\phi+\mu^o) - e^{[\theta\tan\phi]}\cos\theta\right), \\ \gamma &= \zeta^b\gamma^b \pm R\sin\theta \\ &= \pm\left(\zeta^b e^{[(\phi+\mu^o)\tan\phi]}\sin(\phi+\mu^o) - \zeta^b e^{[\theta\tan\phi]}\sin\theta\right), \end{aligned} \tag{4.36}$$

in terms of Cartesian coordinates with respect to the $\delta, \gamma$ coordinate system. All subsequent segments $k$ of the spiral

$$\phi + 2\pi k \leq \theta \leq \phi + \mu^o + 2\pi k \qquad k = 1, 2, 3\ldots \tag{4.37}$$

will coincide once scaled by the dimensionless parameter $\zeta^b$.

Figure 4.17 shows that $\delta^b$ (the normal component of the origin of the spiral) is an order of magnitude larger than $\delta^a$ (the normal component of the inflection point) for a wide range of properties $\phi, \mu^o$. This, together with a careful choice

of elastic constants $K_t, K_s$, will avoid predictions of the state point near the origin of the spiral $\delta^b, \gamma^b$. Note also that the formulation is only valid if the state point lies above $\delta^b, \gamma^b$.

### 4.3.2 Displacements in Relative Displacement Space

We may now consider the corrector step problem of determining the elastic and inelastic parts of the total relative displacement of an arbitrary increment, say the $(n+1)^{th}$ increment. The displacements $\delta_n, \gamma_n$, the forces $N_n, S_n$, the inelastic displacements $\delta_n^p, \gamma_n^p$, and the position of the vertex $\delta_n^o, \gamma_n^o$, are all known at the $n^{th}$ increment. We are further given increments in the total displacements $\Delta\delta_{n+1}, \Delta\gamma_{n+1}$, and hence new relative displacements $\delta_{n+1} = \delta_n + \Delta\delta_{n+1}$, $\gamma_{n+1} = \gamma_n + \Delta\gamma_{n+1}$.

Figure 4.18 shows the relative displacement space for the case where the vertex $\delta_n^o, \gamma_n^o$ is located at a point along the segment of the logarithmic spiral curve with $\gamma_n^o > 0$. The case where the vertex is located at a point along the straight line segment is shown in Figure 4.19. The mirror image of these figures apply when $\gamma_n^o < 0$.

Determining the elastic and inelastic parts of this new estimate of total relative displacements $\delta_{n+1}, \gamma_{n+1}$ is similar to the way in which they are determined for the asperity model. The only significant difference is that the regions associated with dilation and contraction are divided into two subregions each, with the elastic vector oriented differently in each subregion. Consider, for instance, the case where the state point $\delta_{n+1}, \gamma_{n+1}$ lies in region 2a. If this is the case, it is more convenient to transform the state point into polar coordinates $R_{n+1}^b, \theta_{n+1}^b$, measured with respect to point $D^b$:

$$\begin{aligned} R_{n+1}^b &= \sqrt{(\delta^b - \delta_{n+1})^2 + (\gamma^b - \gamma_{n+1})^2}\,, \\ \theta_{n+1}^b &= \sin^{-1}\left(\frac{\gamma^b - \gamma_{n+1}}{R_{n+1}^b}\right). \end{aligned} \tag{4.38}$$

The dilation angle at time $t_{n+1}$ is then simply given by

$$\mu_{n+1} = \theta_{n+1} - \phi\,. \tag{4.39}$$

The elastic vector is thus oriented at an angle $\phi + \mu_{n+1}$ with respect to a vertical line (i.e. along a radial line emanating from point $D^b$). The inelastic displacements satisfy a point which lies along curve $O^oO^a$. On the other hand, if the state point lies in region 2b, the dilation angle is zero, and the

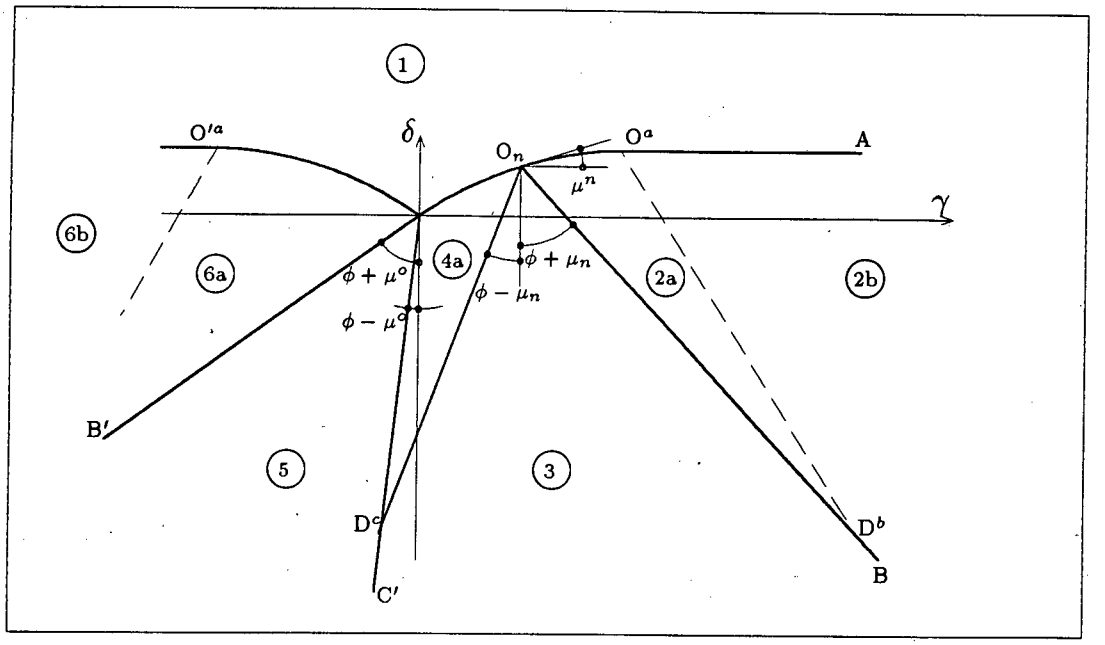



Figure 4.18: State diagram with apex updated along segment $O_oO^a$

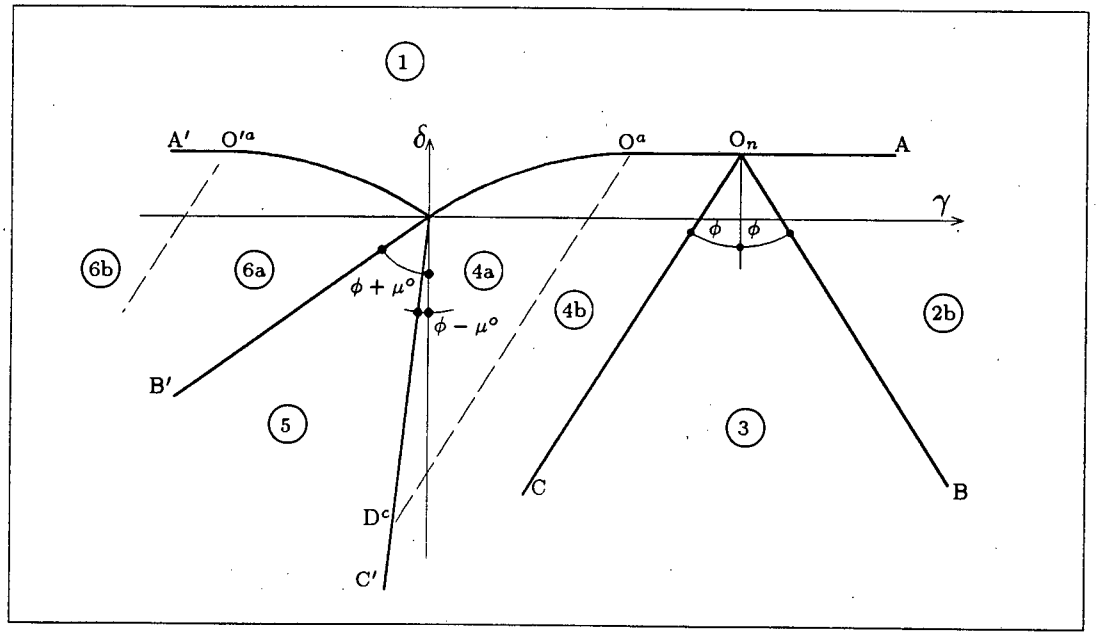



Figure 4.19: State diagram with apex updated along segment $O^aA$

elastic vector is oriented at an angle $\phi$ with respect to a vertical line (i.e. parallel to line $O^aD^b$). The inelastic displacements must then satisfy a point which lies along the horizontal line segment $O^aA$.

The subregions of region 6, which represents behaviour where the direction of shearing is reversed in the $(n+1)^{th}$ increment and dilation taking place with $\gamma^p_{n+1} < 0$, are treated exactly like those of region 2.

Similarly, if the state point lies in subregion 4b, the elastic vector lies at an angle $\phi$ with respect to a vertical line, but parallel to line $O^aD^c$. If the state point, however, lies in region 4a, the elastic and inelastic parts of the total deformation can no longer be found explicitly since the family of elastic vectors, drawn at angle $\phi - \mu_{n+1}$ with respect to a vertical line, do not intersect at one point. We can, however, approximate the segment of the logarithmic spiral by constructing a segment of the logarithmic spiral from point $D^c$, located along line $O^oC'$, as indicated in Figure 4.20. The origin of the spiral in this case lies on the initial yield surface associated with contraction. The elastic vectors then lie along a radial line emanating from point $D^c$, and the inelastic displacements satisfy a point along the spiral constructed from this point. The approximation is adequate as long as the spiral segment is small. This is the case for most practical problems where the initial dilation angle $\mu^o$, which determines the length of segment $O^oO^a$, rarely exceeds 30°.

The location of point $D^c$ is given by

$$\begin{aligned} \delta^c &= -\zeta^c \, e^{[(\phi-\mu^o)\tan\phi]} \cos(\phi-\mu^o) \,, \\ \gamma^c &= -\zeta^c \, e^{[(\phi-\mu^o)\tan\phi]} \sin(\phi-\mu^o) \,, \end{aligned} \tag{4.40}$$

and the scale factor $\zeta^c$ by

$$\zeta^c = \delta^m [e^{[\phi\tan\phi]}\cos\phi - e^{[(\phi-\mu^o)\tan\phi]}\cos(\phi-\mu^o)]^{-1} \,. \tag{4.41}$$

These relations can be derived using arguments similar to those posed in the previous section, and are thus not repeated here. The position of the state point $\delta_{n+1}, \gamma_{n+1}$ in terms of polar coordinates with respect to point $D^c$ is given by

$$\begin{aligned} R^c_{n+1} &= \sqrt{(\delta_{n+1}-\delta^c)^2 + (\gamma_{n+1}-\gamma^c)^2} \,, \\ \theta^c_{n+1} &= \sin^{-1}\left(\frac{\gamma_{n+1}-\gamma^c}{R^c_{n+1}}\right) \,, \end{aligned} \tag{4.42}$$

and the dilation angle at time $t_{n+1}$ is then given by

$$\mu_{n+1} = \phi - \theta_{n+1} \,. \tag{4.43}$$

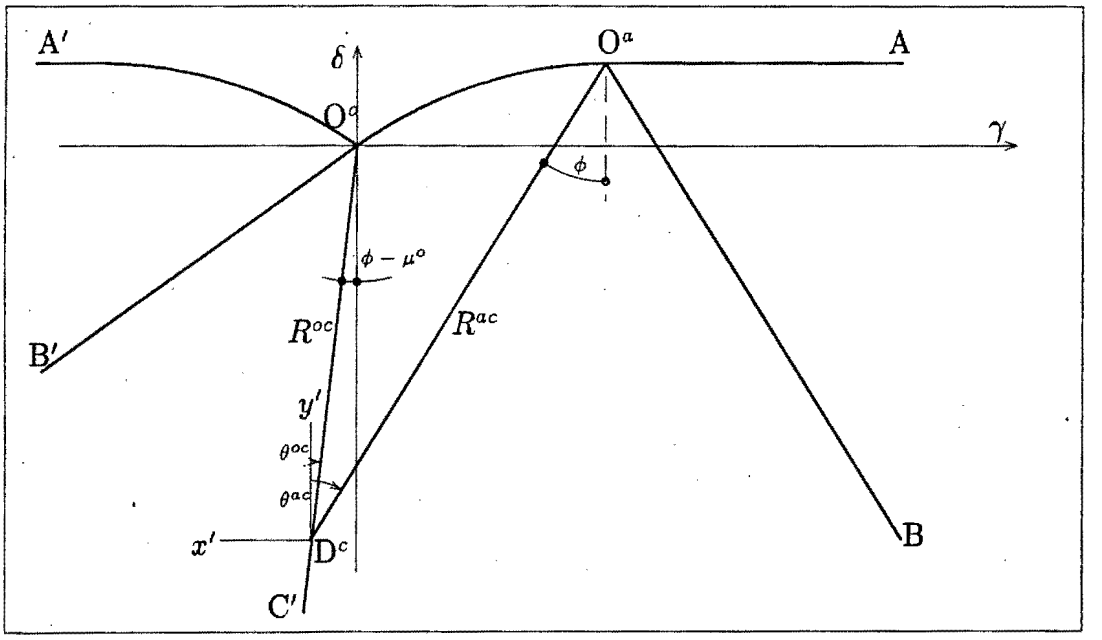



Figure 4.20: Approximation of spiral associated with contraction

It then follows that the elastic vector is oriented at an angle $\phi - \mu_{n+1}$ with respect to a vertical line (i.e. along a radial line emanating from point $D^c$).

Finding the elastic and inelastic parts of the total deformation in any of the remaining regions is straightforward, and were fully discussed in Section 4.2.2. With the exception of region 1, the updating rules of the vertex $\delta^o_{n+1}, \gamma^o_{n+1}$ are also identical to those of the asperity model.

If the state point lies in region 1, we update the apex to a point which lies on curve $O^oA$, $O^oA'$, and directly below the state point. Updating the apex when the state point lies directly above the straight line segment is straightforward. However, if the state point lies above the curved segment the updated position is not explicitly known, and the updating is done using a predictor-corrector iterative scheme. In the predictor step of the first iteration, the apex is updated to a point on curve $O^oO^a$,$O^oO'^a$ which lies along a radial line with origin $D'^b$,$D^b$. This estimate lies on the curve segment, but not directly below the state point. In the corrector step we fix the normal component of this estimate $\delta^{o(1)}_{n+1}$, and set the shear component to $\gamma^{o(1)}_{n+1} = \gamma_{n+1}$. This point then lies below the state point, but no longer on the curve segment. The procedure is repeated until the apex satisfies the requirements that it must

be located on segment $O^oO^a$,$O^oO'^a$ directly below the state point, within an accepted tolerance. The convergence is of the scheme is rapid: the criteria are normally satisfied within a 1 % tolerance after three or four iterations.

### 4.3.3 Full Set of Incremental Equations

The full set of relations are given below, and caters for the two mirror images, as well as for displacements from the virgin state.

We first find the position of the local coordinate system $D^b$,$D'^b$, and the position of the inflection point $O^a$,$O'^a$ in terms of Cartesian coordinates. This point, as well as the shape of the spiral segment, is completely defined by the friction angle $\phi$, the initial dilation angle $\mu^o$, and the maximum dilation $\delta^m$. Point $D^b$,$D'^b$ is located at

$$\begin{aligned} \delta^b &= -\zeta^b\, e^{[(\phi+\mu^o)\tan\phi]}\, \cos(\phi+\mu^o)\,, \\ \gamma^b &= \pm\zeta^b\, e^{[(\phi+\mu^o)\tan\phi]}\, \sin(\phi+\mu^o)\,, \end{aligned} \tag{4.44}$$

and point $O^a$,$O'^a$ at

$$\begin{aligned} \delta^a &= -\zeta^b\, e^{[(\phi+\mu^o)\tan\phi]}\, \cos(\phi+\mu^o) + \zeta^b\, e^{[\phi\tan\phi]}\, \cos\phi\,, \\ \gamma^a &= \pm\zeta^b \left( e^{[(\phi+\mu^o)\tan\phi]}\, \sin(\phi+\mu^o) - e^{[\phi\tan\phi]}\, \sin\phi \right)\,. \end{aligned} \tag{4.45}$$

The scale factor $\zeta^b$ is given by

$$\zeta^b = \frac{\delta^m}{e^{[\phi\tan\phi]}\cos\phi - e^{[(\phi+\mu^o)\tan\phi]}\cos(\phi+\mu^o)}\,.$$

The signs of $\gamma^a$ and $\gamma^b$ depend on the context in which they are used, and will therefore be decided upon later.

The position of the vertex $\delta_n^o, \gamma_n^o$, in terms of polar coordinates with respect to point $D^b$,$D'^b$, is then given by

$$\begin{aligned} R_n^o &= \sqrt{(\delta^b - \delta_n^o)^2 + (\gamma^b - |\gamma_n^o|)^2}\,, \\ \theta_n^o &= \sin^{-1}\left(\frac{\gamma^b - |\gamma_n^o|}{R_n^o}\right)\,. \end{aligned} \tag{4.46}$$

It then follows that the dilation angle at time $t_n$ is given by

$$\begin{aligned} \mu_n &= 0 \qquad \text{if} \qquad \theta_n^o \le \phi\,, \qquad \text{and} \\ \mu_n &= \theta_n^o - \phi \qquad \text{if} \qquad \phi < \theta_n^o < \phi + \mu^o\,. \end{aligned}$$

Similarly, the position of the state point $\delta_{n+1}, \gamma_{n+1}$, in terms of polar coordinates with respect to point $D^b$,$D'^b$, is given by

$$\begin{aligned} R^b_{n+1} &= \sqrt{(\delta^b - \delta_{n+1})^2 + (\gamma^b - |\gamma_{n+1}|)^2}\,, \\ \theta^b_{n+1} &= \sin^{-1}\left(\frac{\gamma^b - |\gamma_{n+1}|}{R^b_{n+1}}\right)\,. \end{aligned} \tag{4.47}$$

The full set of relations can then be written as follows:
if

$$\begin{aligned} \delta_{n+1} &\geq \delta^m && \text{or}\,, \\ R^b_{n+1} &\geq \zeta^b\, e^{[\theta^o_{n+1}\tan\phi]} && \text{with} \qquad \phi < \theta^b_{n+1} < \phi + \mu^o\,, \end{aligned}$$

the joint surfaces are separated, hence

$$\begin{aligned} \delta^p_{n+1} &= \delta_{n+1}\,, \\ \gamma^p_{n+1} &= \gamma_{n+1}\,, \\ \delta^e_{n+1} &= \gamma^e_{n+1} = 0\,, \\ N_{n+1} &= S_{n+1} = 0\,. \end{aligned} \tag{4.48}$$

The position of the apex is updated as follows:
If

$$\delta_{n+1} \geq \delta^m \qquad \text{with} \qquad \theta^b_{n+1} \leq \phi\,,$$

then

$$\begin{aligned} \delta^o_{n+1} &= \delta^m\,, \\ \gamma^o_{n+1} &= \gamma_{n+1}\,. \end{aligned} \tag{4.49}$$

On the other hand, if

$$R^{o(i)}_{n+1} \geq \zeta^b\, e^{[\theta^{o(i)}_{n+1}\tan\phi]} \qquad \text{with} \qquad \phi < \theta^{o(i)}_{n+1} < \phi + \mu^o\,,$$

the position of the apex is found using an iterative scheme where

$$\begin{aligned} \delta^{o(i+1)}_{n+1} &= \zeta^b\, e^{[\theta^{o(i)}_{n+1}\tan\phi]}\,\cos\theta^{o(i)}_{n+1}\,, \\ \gamma^{o(i+1)}_{n+1} &= \gamma_{n+1}\,, \\ R^{o(i+1)}_{n+1} &= \sqrt{\left(\delta^b - \delta^{o(i+1)}_{n+1}\right)^2 + \left(\gamma^b - |\gamma^{o(i+1)}_{n+1}|\right)^2}\,, \\ \theta^{o(i+1)}_{n+1} &= \sin^{-1}\left(\frac{\gamma^b - |\gamma^{o(i+1)}_{n+1}|}{R^{o(i+1)}_{n+1}}\right)\,. \end{aligned} \tag{4.50}$$

The first estimate is given by $\delta_{n+1}^{o(1)} = \delta^m$, $\gamma_{n+1}^{o(1)} = \gamma_{n+1}$. The iteration scheme is terminated when $|R_{n+1}^{o(i+1)} - R_{n+1}^{o(i)}|$ is smaller than a specified tolerance.

If

$$\begin{aligned} \delta_{n+1} &< \delta^m \quad \text{or} \quad R_{n+1}^b < \zeta^b \, e^{[\theta_{n+1}^b \tan\phi]}, \quad \text{and} \\ \delta_{n+1} &> \delta_n^o - (\gamma_{n+1} - \gamma_n^o)\,\mathrm{sgn}(\gamma_n^o)\,\cot(\phi + \mu_n), \end{aligned}$$

further dilation takes place. The orientation of the elastic vector depends on the subregion in which the state point lies. If

$$\delta_{n+1} \geq \delta^a - (|\gamma_{n+1}| - \gamma^a)\,\cot\phi,$$

the state point lies in region 2a of Figure 4.18. The elastic vector is then oriented at an angle $\phi$ with respect to a vertical line, hence

$$\begin{aligned} \delta_{n+1}^p &= \delta^m, \\ \gamma_{n+1}^p &= (\delta_{n+1} - \delta^m)\mathrm{sgn}(\gamma_n^o)\,\tan\phi + \gamma_{n+1}, \\ \delta_{n+1}^e &= \delta_{n+1} - \delta^m, \\ \gamma_{n+1}^e &= (\delta^m - \delta_{n+1})\,\mathrm{sgn}(\gamma_n^o)\,\tan\phi, \\ \delta_{n+1}^o &= \delta_{n+1}^p, \qquad \gamma_{n+1}^o = \gamma_{n+1}^p, \\ N_{n+1} &= K_t \delta_{n+1}^e, \qquad S_{n+1} = K_s \gamma_{n+1}^e. \end{aligned} \tag{4.51}$$

If, however,

$$\delta_{n+1} < \delta^a - (|\gamma_{n+1}| - \gamma^a)\,\cot\phi,$$

the state point lies in region 2b of Figure 4.18. The elastic vector then lies along a radial line emanating from pont $\mathrm{D}^b$,$\mathrm{D}'^b$, hence

$$\begin{aligned} \delta_{n+1}^p &= \delta^b + \zeta^b e^{[\theta_{n+1}^b \tan\phi]} \cos\theta_{n+1}^b, \\ \gamma_{n+1}^p &= (\gamma^b - \zeta^b \, e^{[\theta_{n+1}^b \tan\phi]} \sin\theta_{n+1}^b)\,\mathrm{sgn}(\gamma_n^o), \\ \delta_{n+1}^e &= -(\zeta^b \, e^{[\theta_{n+1}^b \tan\phi]} - R_{n+1}^b)\cos\theta_{n+1}^b, \\ \gamma_{n+1}^e &= (\zeta^b \, e^{[\theta_{n+1}^b \tan\phi]} - R_{n+1}^b)\sin\theta_{n+1}^b, \\ \delta_{n+1}^o &= \delta_{n+1}^p, \qquad \gamma_{n+1}^o = \gamma_{n+1}^p, \\ N_{n+1} &= K_t \delta_{n+1}^e, \qquad S_{n+1} = K_s \gamma_{n+1}^e. \end{aligned} \tag{4.52}$$

If

$$\begin{aligned} \delta_{n+1} &\leq \delta_n^o - (\gamma_{n+1} - \gamma_n^o)\ \mathrm{sgn}(\gamma_n^o)\ \cot(\phi + \mu_n)\ , \qquad \text{and} \\ \delta_{n+1} &\leq \delta_n^o + (\gamma_{n+1} - \gamma_n^o)\ \mathrm{sgn}(\gamma_n^o)\ \cot(\phi - \mu_n)\ , \end{aligned}$$

the response is elastic, hence

$$\begin{aligned} \delta_{n+1}^e &= \delta_{n+1} - \delta_n^o, & \gamma_{n+1}^e &= \gamma_{n+1} - \gamma_n^o, \\ \delta_{n+1}^p &= \delta_n^o\ , & \gamma_{n+1}^p &= \gamma_n^o\ , \\ \delta_{n+1}^o &= \delta_n^o\ , & \gamma_{n+1}^o &= \gamma_n^o\ , \\ N_{n+1} &= K_t \delta_{n+1}^e\ , & S_{n+1} &= K_s \gamma_{n+1}^e . \end{aligned} \tag{4.53}$$

If

$$\begin{aligned} \delta_{n+1} &< \delta^m \qquad \text{or} \qquad R_{n+1}^b < \zeta^b\ e^{[\theta_{n+1}^b \tan\phi]}\ , \\ \delta_{n+1} &> \delta_n^o + (\gamma_{n+1} - \gamma_n^o)\ \mathrm{sgn}(\gamma_n^o)\ \cot(\phi - \mu_n)\ , \qquad \text{and} \\ \delta_{n+1} &< \gamma_{n+1}\ \mathrm{sgn}(\gamma_n^o)\ \cot(\phi - \mu^o)\ , \end{aligned}$$

the joint contracts. The orientation of the elastic vector depends on the subregion in which the state point lies. If

$$\delta_{n+1} \geq \delta^a - (|\gamma_{n+1}| - \gamma^a)\ \cot\phi\ ,$$

the state point lies in region 4a of Figure 4.19, with the elastic vector oriented at an angle $\phi$ with respect to a vertical line, hence

$$\begin{aligned} \delta_{n+1}^p &= \delta^m\ , \\ \gamma_{n+1}^p &= (\delta^m - \delta_{n+1})\ \mathrm{sgn}(\gamma_n^o)\ \tan\phi + \gamma_{n+1}\ , \\ \delta_{n+1}^e &= \delta_{n+1} - \delta^m\ , \\ \gamma_{n+1}^e &= (\delta_{n+1} - \delta^m)\ \mathrm{sgn}(\gamma_n^o)\ \tan\phi\ , \\ \delta_{n+1}^o &= \delta_{n+1}^p\ , & \gamma_{n+1}^o &= \gamma_{n+1}^p\ , \\ N_{n+1} &= K_t \delta_{n+1}^e\ , & S_{n+1} &= K_s \gamma_{n+1}^e\ . \end{aligned} \tag{4.54}$$

If, on the other hand

$$\delta_{n+1} < \delta^a - (|\gamma_{n+1}| - \gamma^a)\ \cot\phi\ ,$$

the state point lies in region 4b of Figure 4.19. Explicit relations for the inelastic and elastic deformation are then no longer possible. A good approximation can, however, be obtained assuming that the elastic vector lies along a radial line emanating from a point $D^c$,$D'^c$, defined by coordinates

$$\delta^c = -\zeta^c\ e^{[(\phi - \mu^o)\tan\phi]}\ \cos(\phi - \mu^o)\ ,$$

$$\gamma^c = -\zeta^c \, e^{[(\phi-\mu^o)\,\tan\phi]} \, \mathrm{sgn}(\gamma_n^o) \, \sin(\phi-\mu^o) \,, \tag{4.55}$$

with

$$\zeta^c = \frac{\delta^m}{e^{[\phi\tan\phi]}\cos\phi - e^{[(\phi-\mu^o)\tan\phi]}\cos(\phi-\mu^o)} \,.$$

The position of the state point in terms of polar coordinates is

$$\begin{aligned} R_{n+1}^c &= \sqrt{(\delta_{n+1}-\delta^c)^2 + (\gamma_{n+1}-\gamma^c)^2} \,, \\ \theta_{n+1}^c &= \sin^{-1}\left[\frac{(\gamma_n^o-\gamma^c)\mathrm{sgn}(\gamma_n^o)}{R_{n+1}^c}\right] . \end{aligned} \tag{4.56}$$

Then

$$\begin{aligned} \delta_{n+1}^p &= \delta^c + \zeta^c \, e^{[\theta_{n+1}^c\tan\phi]} \, \cos\theta_{n+1}^c \,, \\ \gamma_{n+1}^p &= \gamma^c + \zeta^c \, e^{[\theta_{n+1}^c\tan\phi]} \, \mathrm{sgn}(\gamma_n^o) \, \sin\theta_{n+1}^c \,, \\ \delta_{n+1}^e &= -(\zeta^c \, e^{[\theta_{n+1}^c\tan\phi]} - R_{n+1}^c)\cos\theta_{n+1}^c \,, \\ \gamma_{n+1}^e &= -(\zeta^c \, e^{[\theta_{n+1}^c\tan\phi]} - R_{n+1}^c)\sin\theta_{n+1}^c \,, \\ \delta_{n+1}^o &= \delta_{n+1}^p \,, \qquad \gamma_{n+1}^o = \gamma_{n+1}^p \,, \\ N_{n+1} &= K_t\delta_{n+1}^e \,, \qquad S_{n+1} = K_s\gamma_{n+1}^e \,. \end{aligned} \tag{4.57}$$

If

$$\begin{aligned} \delta_{n+1} &\geq \gamma_{n+1}\,\mathrm{sgn}(\gamma_n^o)\,\cot(\phi-\mu^o) \,, \qquad \text{and} \\ \delta_{n+1} &\leq \gamma_{n+1}\,\mathrm{sgn}(\gamma_n^o)\,\cot(\phi+\mu^o) \,, \end{aligned}$$

the joint moves back to the virgin position, and

$$\begin{aligned} \delta_{n+1}^e &= \delta_{n+1}, \qquad \gamma_{n+1}^e = \gamma_{n+1}, \\ \delta_{n+1}^p &= \gamma_{n+1}^p = 0 \,, \\ \delta_{n+1}^o &= \gamma_{n+1}^o = 0 \,, \\ N_{n+1} &= K_t\delta_{n+1}^e \,, \qquad S_{n+1} = K_s\gamma_{n+1}^e. \end{aligned} \tag{4.58}$$

If

$$\begin{aligned} \delta_{n+1} &< \delta^m \quad \text{or} \quad R_{n+1}^b < \zeta^b \, e^{[\theta_{n+1}^b\tan\phi]} \,, \qquad \text{and} \\ \delta_{n+1} &> \gamma_{n+1}\,\mathrm{sgn}(\gamma_n^o)\,\cot(\phi+\mu^o) \,, \end{aligned}$$

the opposite sides of the asperities come in contact. As before, the orientation of the elastic vector depends on the subregion in which the state point lies. If

$$\delta_{n+1} \geq \delta^a - (|\gamma_{n+1}| - \gamma^a) \cot\phi \,,$$

the state point lies in region 6a of Figure 4.18, and the elastic vector lies at an angle $\phi$ with respect to a vertical line. Then

$$\begin{aligned}
\delta^p_{n+1} &= \delta^m \,, \\
\gamma^p_{n+1} &= -(\delta_{n+1} - \delta^m)\,\mathrm{sgn}(\gamma^o_n)\,\tan\phi + \gamma_{n+1} \,, \\
\delta^e_{n+1} &= \delta_{n+1} - \delta^m \,, \\
\gamma^e_{n+1} &= -(\delta^m - \delta_{n+1})\,\mathrm{sgn}(\gamma^o_n)\,\tan\phi \,, \\
\delta^o_{n+1} &= \delta^p_{n+1} \,, \qquad\qquad \gamma^o_{n+1} = \gamma^p_{n+1} \,, \\
N_{n+1} &= K_t \delta^e_{n+1} \,, \qquad\qquad S_{n+1} = K_s \gamma^e_{n+1} \,.
\end{aligned} \tag{4.59}$$

Or, if

$$\delta_{n+1} < \delta^a - (|\gamma_{n+1}| - \gamma^a) \cot\phi \,,$$

the state point lies in region 6b of Figure 4.18, and the elastic vector lies along a radial line emanating from pont $\mathrm{D}^b$,$\mathrm{D}'^b$, hence

$$\begin{aligned}
\delta^p_{n+1} &= \delta^b + \zeta^b\, e^{[\theta^b_{n+1}\tan\phi]}\,\cos\theta^b_{n+1} \,, \\
\gamma^p_{n+1} &= -(\gamma^b + \zeta^b\, e^{[\theta^b_{n+1}\tan\phi]}\,\sin\theta^b_{n+1})\mathrm{sgn}(\gamma^o_n) \,, \\
\delta^e_{n+1} &= -(\zeta^b\, e^{[\theta^b_{n+1}\tan\phi]} - R^b_{n+1})\cos\theta^b_{n+1} \,, \\
\gamma^e_{n+1} &= (\zeta^b\, e^{[\theta^b_{n+1}\tan\phi]} - R^b_{n+1})\sin\theta^b_{n+1} \,, \\
\delta^o_{n+1} &= \delta^p_{n+1} \,, \qquad\qquad \gamma^o_{n+1} = \gamma^p_{n+1} \,, \\
N_{n+1} &= K_t \delta^e_{n+1} \,, \qquad\qquad S_{n+1} = K_s \gamma^e_{n+1} \,.
\end{aligned} \tag{4.60}$$

### 4.3.4 Remarks

We have shown in this section that a simple extension of the sawtooth model can lead to a more realistic rock joint model. The model assumes a non-linear relation, expressed in terms of a logarithmic spiral curve, between inelastic normal and shear relative displacement rates. The curve is constructed so that the inelastic normal deformation ceases after some shearing. The dilation angle, defined as the gradient of this inelastic displacement rate relation,

thus varies with shear deformation; in particular, the dilation angle decreases linearly (from a maximum value to zero) with increasing shear deformation, and hence, contains as limit cases the sawtooth and Coulomb model respectively. The model limits the choice of elastic constants $K_t, K_s$, but nevertheless illustrates some of the features inherent to any model in which both the yield surface and flow potential change from one time-step to the next. A model can be formulated where this restriction is avoided. However, such an extension will lead to a formulation which is no longer explicit. Further, care must be taken since there are generally regions in relative displacement space in which the total deformation can not be uniquely divided into elastic and inelastic components .

## 4.4 SUMMARY

Three plane joint models were formulated in this chapter. Particular attention was given to cater for all possible histories of opening, closing and sliding, in any direction. The advantage of these formulations is that no qualitative distinction need be made between opening and closing on the one hand, and frictional sliding on the other. This is a consequence of the backward difference discretisation assumption which is used to formulate the non-linear incremental equations.

A very simple model in which sliding is governed by classical Coulomb friction, with no cohesion and a non-associated flow rule, was formulated first. This model, however, has an inherent problem associated with reversed shear deformation. An inelastic shear deformation increment of any sign is always accompanied by a positive increment of inelastic normal deformation. Consequently, an initially open joint will close after a finite number of shear cycles.

This problem is avoided in the sawtooth asperity model. The model idealises the rough contact surfaces as interlocking sawtooth asperities which match in the virgin state. Sliding causes asperities on the one side of the joint surface to ride over the asperities on the other surface. The result is surfaces which dilate or contract, depending on the direction of sliding. This model provided important insights into the dilatant Coulomb concept, i.e. the Coulomb model is only applicable in the virgin state, and must be modified when the joint is displaced.

The continuing dilation accompanied by monotonic shearing predicted by the sawtooth model is, however, a limitation. The logarithmic spiral model overcomes this limitation by assuming a non-linear relation between inelastic relative displacement rates, with inelastic normal deformation ceasing after some shearing. As a consequence, the friction laws change from one time-step to the next in a way which is characteristic of a softening material. The model largely retains the simple explicit nature of the Coulomb and sawtooth asperity models, but offers a far more realistic model for dilatant rock joints.

These models represent highly idealised rock joint behaviour. A number of extensions are possible in order to cater for more representative behaviour. These include degrading asperities, and hence only partial contraction before dilation takes place when the direction of shearing is reversed, friction laws which depend on the level of normal stress and a non-linear normal stress-closure relation. Most extensions will, however, result in increasing computational effort.

# CHAPTER 5
# NUMERICAL IMPLEMENTATION

In this chapter we complete the formulation of the model by first deriving a local tangent stiffness matrix for each of the three constitutive models. We further derive relations for an isoparametric interface element of arbitrary orientation and length. The constitutive equations and tangent stiffness matrix are then substituted into these relations which can be assembled into the global incremental finite element equations.

## 5.1 TANGENT STIFFNESS MATRIX

In the predictor step, in say the $(i+1)^{th}$ iteration in the solution of the $(n+1)^{th}$ increment, displacement estimates $\delta_{n+1}^{(i)}, \gamma_{n+1}^{(i)}$ are replaced by improved estimates $\delta_{n+1}^{(i+1)}, \gamma_{n+1}^{(i+1)}$. In order to achieve this, we must formulate a local tangent stiffness based on the holonomic incremental constitutive equations. The local tangent stiffness matrix $\mathbf{D}_{n+1}^{(i)}$ relates the internal force rates $\dot{N}, \dot{S}$ and the displacement rates $\dot{\delta}, \dot{\gamma}$ at the state defined at the end of the $i^{th}$ iteration. The rate equations are written in the form

$$\left\{ \begin{array}{c} \dot{N} \\ \dot{S} \end{array} \right\} = \mathbf{D}_{n+1}^{(i)} \left\{ \begin{array}{c} \dot{\delta} \\ \dot{\gamma} \end{array} \right\} . \tag{5.1}$$

The components of $\mathbf{D}_{n+1}^{(i)}$ are then

$$\mathbf{D}_{n+1}^{(i)} = \left[ \begin{array}{cc} \left. \frac{\partial N}{\partial \delta} \right|_{\delta_{n+1}^{(i)}, \gamma_{n+1}^{(i)}} & \left. \frac{\partial N}{\partial \gamma} \right|_{\delta_{n+1}^{(i)}, \gamma_{n+1}^{(i)}} \\ \left. \frac{\partial S}{\partial \delta} \right|_{\delta_{n+1}^{(i)}, \gamma_{n+1}^{(i)}} & \left. \frac{\partial S}{\partial \gamma} \right|_{\delta_{n+1}^{(i)}, \gamma_{n+1}^{(i)}} \end{array} \right] \tag{5.2}$$

The components of $\mathbf{D}_{n+1}^{(i)}$ depend only on the region in which the state point $\delta_{n+1}^{(i)}, \gamma_{n+1}^{(i)}$ lies. We derive these components, first for the Coulomb model, and then extend them in the second section for the sawtooth and logarithmic

spiral models.

### 5.1.1 The Coulomb Friction Model

The Coulomb model is represented in terms of three different regions (describing opening, elastic and frictional sliding states) in relative displacement space (Fig. 4.6). Two of these cases can be evaluated very simply. Firstly, if the state point lies in region 1, where there is no contact, the gradients in equation (5.2) are each zero. Hence,
if

$$\delta_{n+1}^{(i)} \geq \delta_n^o + |\gamma_{n+1}^{(i)} - \gamma_n^o| \tan\mu \,,$$

then

$$\mathbf{D}_{n+1}^{(i)} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} . \tag{5.3}$$

Secondly, if the state point lies in region 3, the behaviour is linear elastic; there is contact but no sliding, hence,
if

$$\delta_{n+1}^{(i)} \leq \delta_n^o - |\gamma_{n+1}^{(i)} - \gamma_n^o| \cot\psi \,,$$

then

$$\mathbf{D}_{n+1}^{(i)} = \begin{bmatrix} K_t & 0 \\ 0 & K_s \end{bmatrix} . \tag{5.4}$$

In order to derive the components of $\mathbf{D}_{n+1}^{(i)}$ when the state point lies in region 2, we note that

$$\begin{aligned} N_{n+1} &= K_t \delta_{n+1}^e \,, \\ S_{n+1} &= K_s \gamma_{n+1}^e \,. \end{aligned}$$

We may form the derivatives required in equation (5.2) by making use of the expressions for $\delta_{n+1}^e, \gamma_{n+1}^e$ in equation (4.10). After some straightforward manipulation, in which we use equation (4.7), we find that
if

$$\delta_n^o + |\gamma_{n+1}^{(i)} - \gamma_n^o| \tan\mu \; > \; \delta_{n+1}^{(i)} \; > \; \delta_n^o - |\gamma_{n+1}^{(i)} - \gamma_n^o| \cot\psi \,.$$

then

$$\mathbf{D}_{n+1}^{(i)} = \frac{K_t K_s}{K_s + K_t \tan\phi \tan\mu} \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix} \tag{5.5}$$

with the components $D_{ij}$ given by

$$\begin{aligned} D_{11} &= 1 , \\ D_{12} &= -\mathrm{sgn}(\gamma_{n+1}^{(i)} - \gamma_n^o) \tan\mu , \\ D_{21} &= -\mathrm{sgn}(\gamma_{n+1}^{(i)} - \gamma_n^o) \tan\phi , \\ D_{22} &= \tan\phi \tan\mu . \end{aligned}$$

It is evident that $\mathbf{D}_{n+1}^{(i)}$ is non-symmetric in region 3, except in the special case of an associated flow rule where $\phi = \mu$. The expressions for the sawtooth asperity model and logarithmic spiral model are simply extensions of the Coulomb model, and are given in the following section.

### 5.1.2 The Sawtooth and Logarithmic Spiral Asperity Models

The two asperity models are represented by six different regions in relative displacement space (Figures 4.11, 4.18 and 4.19). The only difference between the two asperity models is that the constant dilation angle of the sawtooth model is replaced by a varying dilation angle, which is either zero or given by equation (4.39) or (4.43). The spiral model further limits the choice of elastic constants, i.e. $K_t = K_s$.

If the state point lies in region 1, where there is no contact, the gradients in equation (5.2) are each zero. Hence,
if

$$\delta_{n+1}^{(i)} \geq |\gamma_{n+1}^{(i)}| \tan\mu ,$$

then

$$\mathbf{D}_{n+1}^{(i)} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} . \tag{5.6}$$

If the state point lies in region 3 or 5, the behaviour is linear elastic; there is contact but no sliding.
If

$$\delta_{n+1}^{(i)} \leq \delta_n^o - (\gamma_{n+1}^{(i)} - \gamma_n^o) \,\mathrm{sgn}(\gamma_n^o) \cot\psi^+ , \qquad \text{and}$$

$$\delta_{n+1}^{(i)} \leq \delta_n^o + (\gamma_{n+1}^{(i)} - \gamma_n^o)\,\mathrm{sgn}(\gamma_n^o)\,\cot\psi^- ,$$

or if

$$\delta_{n+1}^{(i)} \geq \gamma_{n+1}^{(i)}\,\mathrm{sgn}(\gamma_n^o)\,\cot\psi^- , \qquad \text{and}$$
$$\delta_{n+1}^{(i)} \leq \gamma_{n+1}^{(i)}\,\mathrm{sgn}(\gamma_n^o)\,\cot\psi^+ ,$$

then

$$\mathbf{D}_{n+1}^{(i)} = \begin{bmatrix} K_t & 0 \\ 0 & K_s \end{bmatrix} . \tag{5.7}$$

If the state point lies in region 2 inelastic deformation occurs, with the joint surfaces dilating. Hence,
if

$$\delta_{n+1}^{(i)} < \gamma_{n+1}^{(i)}\,\mathrm{sgn}(\gamma_n^o)\,\tan\mu , \qquad \text{and}$$
$$\delta_{n+1}^{(i)} > \delta_n^o - (\gamma_{n+1}^{(i)} - \gamma_n^o)\,\mathrm{sgn}(\gamma_n^o)\,\cot\psi^+ ,$$

then

$$\mathbf{D}_{n+1}^{(i)} = \frac{K_t K_s}{K_s + K_t \tan(\phi+\mu)\tan\mu} \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix} \tag{5.8}$$

with the components $D_{ij}$ given by

$$\begin{aligned} D_{11} &= 1 , \\ D_{12} &= -\mathrm{sgn}(\gamma_{n+1}^{(i)} - \gamma_n^o)\tan\mu , \\ D_{21} &= -\mathrm{sgn}(\gamma_{n+1}^{(i)} - \gamma_n^o)\tan(\phi+\mu) , \\ D_{22} &= \tan(\phi+\mu)\tan\mu . \end{aligned}$$

If the state point lies in region 6 inelastic deformation takes place with the joint surfaces dilating, but with the opposite side of asperities in contact. Hence, if

$$\delta_{n+1}^{(i)} < |\gamma_{n+1}^{(i)}|\,\tan\mu , \qquad \text{and}$$
$$\delta_{n+1}^{(i)} > \gamma_{n+1}^{(i)}\,\mathrm{sgn}(\gamma_n^o)\,\cot\psi^+ ,$$

then

$$\mathbf{D}_{n+1}^{(i)} = \frac{K_t K_s}{K_s + K_t \tan(\phi+\mu)\tan\mu} \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix} \tag{5.9}$$

with the components $D_{ij}$ given by

$$\begin{aligned} D_{11} &= 1 \\ D_{12} &= \text{sgn}(\gamma_n^\circ)\tan\mu\,, \\ D_{21} &= \text{sgn}(\gamma_n^\circ)\tan(\phi+\mu)\,, \\ D_{22} &= \tan(\phi+\mu)\tan\mu\,. \end{aligned}$$

Finally, the joint contracts if the state point lies in region 4. If

$$\begin{aligned} \delta_{n+1}^{(i)} &< |\gamma_{n+1}^{(i)}|\,\tan\mu\,, \\ \delta_{n+1}^{(i)} &> \delta_n^\circ + (\gamma_{n+1}^{(i)} - \gamma_n^\circ)\,\text{sgn}(\gamma_n^\circ)\,\cot\psi^-\,, \qquad \text{and} \\ \delta_{n+1}^{(i)} &< \gamma_{n+1}^{(i)}\,\text{sgn}(\gamma_n^\circ)\,\cot\psi^-\,, \end{aligned}$$

then

$$\mathbf{D}_{n+1}^{(i)} = \frac{K_t K_s}{K_s - K_t\tan(\phi-\mu)\tan\mu}\begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix} \tag{5.10}$$

with the components $D_{ij}$ given by

$$\begin{aligned} D_{11} &= 1 \\ D_{12} &= \text{sgn}(\gamma_{n+1}^{(i)} - \gamma_n^\circ)\tan\mu\,, \\ D_{21} &= -\text{sgn}(\gamma_{n+1}^{(i)} - \gamma_n^\circ)\tan(\phi-\mu)\,, \\ D_{22} &= -\tan(\phi-\mu)\tan\mu\,. \end{aligned}$$

## 5.2 FINITE ELEMENT DISCRETISATION

The formulation of the isoparametric interface element which follows is based on the work reported by Beer[9], Plesha[35] and Rodic and Owen[38].

Consider a two-dimensional problem with two bodies in contact. We assume that the contact surfaces are sufficiently smooth so that the surface tangent and normal are not ill-defined. Each of the bodies is then discretised with quadratic continuum finite elements, and the contact surfaces are discretised with quadratic 6-noded interface elements. One such interface element, together with the geometry of the reference element, is shown in Figure 5.1. Compatibility on the contact surface is ensured if the end nodes of the interface element correspond to the corner nodes of the continuum elements, and

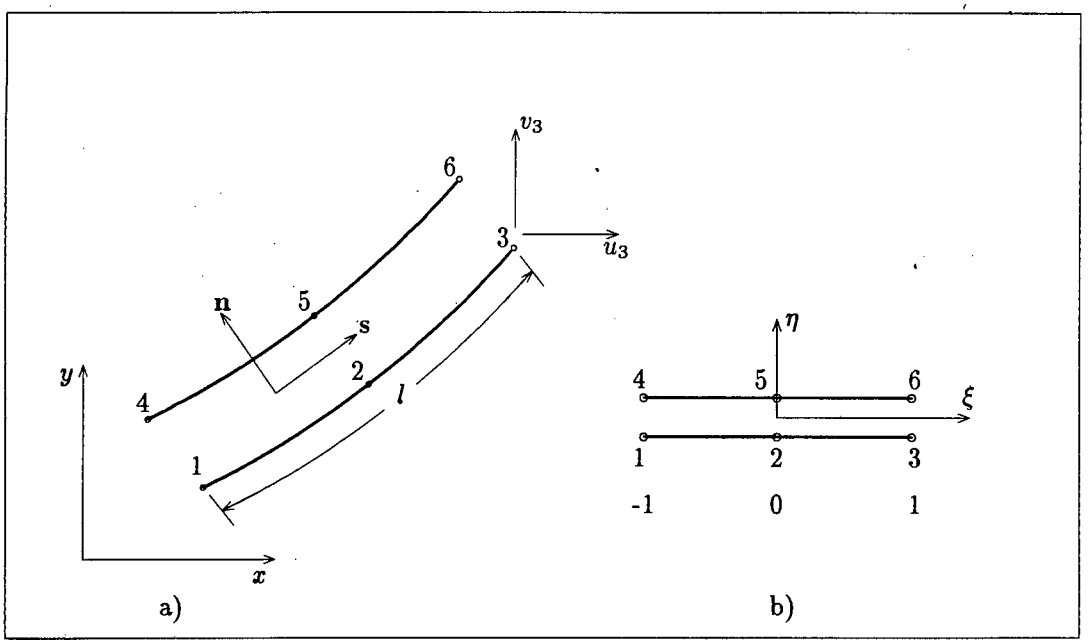



Figure 5.1: (a) Geometry of joint element and (b) Geometry of reference element

the midside nodes of the interface elements correspond to the midside nodes of the continuum elements. We also note that a four noded interface element, which is compatible with linear continuum elements, can be recovered by selecting linear shape functions.

Since the interface element is restricted to small displacement analysis, only one face of the element is required to define the geometry of the element. The global coordinates of the bottom face of the element are given in terms of the reference coordinate $\xi$, by the mapping

$$x_b = (N_1 \quad N_2 \quad N_3) \begin{Bmatrix} x_1 \\ x_2 \\ x_3 \end{Bmatrix} = \mathbf{N}_b \mathbf{x}_b \ , \tag{5.11}$$

and

$$y_b = (N_1 \quad N_2 \quad N_3) \begin{Bmatrix} y_1 \\ y_2 \\ y_3 \end{Bmatrix} = \mathbf{N}_b \mathbf{y}_b \ , \tag{5.12}$$

with the shape functions given by

$$N_1 = \frac{1}{2}\xi(\xi - 1)\,, \quad N_2 = 1 - \xi^2 \quad \text{and} \quad N_3 = \frac{1}{2}\xi(\xi + 1)\,, \tag{5.13}$$

where the subscript numbers refer to the node numbers.

The unit vectors normal and tangent to the contact surface are respectively given by

$$\mathbf{n} = \frac{1}{J}\left[\frac{\partial x_b}{\partial \xi}\,\mathbf{i} + -\frac{\partial y_b}{\partial \xi}\,\mathbf{j}\right],$$

and

$$\mathbf{s} = \frac{1}{J}\left[\frac{\partial y_b}{\partial \xi}\,\mathbf{i} + \frac{\partial x_b}{\partial \xi}\,\mathbf{j}\right],$$

where the Jacobian $J$ is related to the arc-length $l$

$$J = \left[\left(\frac{\partial x_b}{\partial \xi}\right)^2 + \left(\frac{\partial y_b}{\partial \xi}\right)^2\right]^{\frac{1}{2}} = \frac{l}{2}\,. \tag{5.14}$$

The transformation from global to local coordinates is then given by the following matrix

$$\boldsymbol{\Theta} = \begin{bmatrix} \mathbf{n} \\ \mathbf{s} \end{bmatrix}. \tag{5.15}$$

The nodal displacements are expressed in the same way as the nodal coordinates. The displacement at any point at the bottom of the joint is given by the approximation

$$\begin{Bmatrix} u_b \\ v_b \end{Bmatrix} = \begin{bmatrix} N_1 & 0 & N_2 & 0 & N_3 & 0 \\ 0 & N_1 & 0 & N_2 & 0 & N_3 \end{bmatrix} \begin{Bmatrix} u_1 \\ v_1 \\ u_2 \\ v_2 \\ u_3 \\ v_3 \end{Bmatrix}$$

$$= \mathbf{N}_b \mathbf{u}_b\,, \tag{5.16}$$

and similarly for a point on the top surface

$$\begin{Bmatrix} u_t \\ v_t \end{Bmatrix} = \mathbf{N}_t \mathbf{u}_t\,, \tag{5.17}$$

where the shape functions are the same as those used in the geometric representation.

The nodal displacements in the normal and tangential directions are obtained by the following transformation

$$\left\{ \begin{array}{c} u_i' \\ v_i' \end{array} \right\} = \Theta \left\{ \begin{array}{c} u_i \\ v_i \end{array} \right\}, \qquad i = t, b. \tag{5.18}$$

The relative displacements in the local coordinate system are defined as the difference between the displacements of the top and bottom surfaces

$$\left\{ \begin{array}{c} \delta \\ \gamma \end{array} \right\} = \left\{ \begin{array}{c} u_t' \\ v_t' \end{array} \right\} - \left\{ \begin{array}{c} u_b' \\ v_b' \end{array} \right\} \tag{5.19}$$

Substitution of equations (5.16) (5.17) and (5.18) into above leads to the following relationship

$$\left\{ \begin{array}{c} \delta \\ \gamma \end{array} \right\} = \mathbf{B} \left\{ \begin{array}{c} \mathbf{u}_b \\ \mathbf{u}_t \end{array} \right\}, \qquad \text{where} \quad \mathbf{B} = \Theta[-\mathbf{N}_b \quad \mathbf{N}_t]\,. \tag{5.20}$$

We note that the $\mathbf{B}$ matrix is analogous to those of continuum elements except that derivatives of shape functions are not required. The element stiffness matrix ${}^e\mathbf{K}$ and internal force vector ${}^e\mathbf{F}$ are obtained by standard procedures, and are given by the following relations

$${}^e\mathbf{K} = \int_l \mathbf{B}^T\mathbf{D}\mathbf{B}\, ds = \int_{-1}^{1} \mathbf{B}^T\mathbf{D}\mathbf{B}\, J\, d\xi\,, \tag{5.21}$$

and

$${}^e\mathbf{F} = \int_l \mathbf{B}\,\boldsymbol{\sigma}\, ds = \int_{-1}^{1} \mathbf{B}\,\boldsymbol{\sigma}\, J\, d\xi\,. \tag{5.22}$$

The matrix $\mathbf{D}$ and the conjugate forces or stresses $\boldsymbol{\sigma}$ for an element of unit length were formulated in Section 5.1 and Chapter 4 respectively. The correct relations are simply substituted into equations (5.21) and (5.22) to complete the formulation.

The components of ${}^e\mathbf{K}$ and ${}^e\mathbf{F}$ are usually integrated numerically by using either Gaussian-quadrature or Newton-Cotes quadrature. A Gauss-quadrature integration, with the choice of full or reduced integration, was adopted in this formulation. The performance of the joint can, however, be seriously affected by the choice of integration order[30] as well as the magnitude of the elastic constants $K_t, K_s$. These aspects are considered in more detail in Example 6.1.1.

## 5.3 IMPLEMENTATION INTO A FINITE ELEMENT CODE

The joint element was implemented into a general purpose finite element code, ABAQUS. ABAQUS[23] is based on an incremental strategy and uses a Newton-Raphson method as numerical technique to solve the non-linear equilibrium equations within each increment. The user is provided with a library of interfaces for user subroutines. The interface to define an element was used to implement the joint element.

Within each iteration, at a particular increment of the Newton-Raphson scheme, the predicted displacements $\mathbf{u}_{n+1}^{(i)}$, as well as the state variables associated with each element, are passed to the user subroutine. The state variables are defined to be the position of the vertex $\delta_n^o, \gamma_n^o$, and are only updated once equilibrium within an increment is achieved. The relative displacements $\delta_{n+1}^{(i)}, \gamma_{n+1}^{(i)}$ are then calculated according to equation (5.20), and are used, together with the position of the vertex $\delta_n^o, \gamma_n^o$, to decide which contact conditions apply. The internal force vector ${}^e\mathbf{F}_{n+1}^{(i)}$ and stiffness matrix ${}^e\mathbf{K}_{n+1}^{(i)}$ of each element are then evaluated according to equations (5.21) and (5.22). These are passed back to the main program, assembled into the global equations, and the next displacement increment predicted:

$$\Delta\mathbf{u}^{(i)} = [\mathbf{K}_{n+1}^{(i)}]^{-1} [\mathbf{R}_{n+1} - \mathbf{F}_{n+1}^{(i)}], \tag{5.23}$$

where $\mathbf{R}_{n+1}$ represents the external force vector. The improved total displacement is then given by

$$\mathbf{u}_{n+1}^{(i+1)} = \mathbf{u}_{n+1}^{(i)} + \Delta\mathbf{u}^{(i)}. \tag{5.24}$$

This new estimate of displacement is passed to the user subroutine and the next correction is again performed at element (or Gauss point) level. The procedure is continued until the force residual $\boldsymbol{\Phi}_{n+1}^{(i+1)} = \mathbf{R}_{n+1} - \mathbf{F}_{n+1}^{(i+1)}$ is reduced to an acceptable value, at which stage the solution moves to the next load increment.

## 5.4 SUMMARY

We considered in this chapter the implementation of the interface element into an existing finite element program ABAQUS. We first derived an element tangent stiffness matrix for the different joint models. These relations, together with the constitutive relations, are incorporated into the spatially

discrete interface element, which can be assembled into the global finite element equations. The interface element is fully compatible with other plane elements in the code.

# CHAPTER 6
# EXAMPLES

In this chapter we consider a series of applications, mainly in the field of rock mechanics, to illustrate the predictions of the formulation. The aim with these applications is to verify the interface formulation and to highlight the features of the different constitutive models.

The first example, which concerns a wedge being forced between two blocks, directly compares the penalty method used in this formulation with the Lagrange-multiplier method used by ABAQUS. Issues such as convergence and accuracy are addressed. In this example the dilation angle is set zero, and all the constitutive models are therefore equivalent.

The second problem investigates the behaviour of a vertical fracture which intersects a hangingwall beam. This problem is intended to emphasise the importance of a joint which can model any possible history of opening, sticking and sliding states. The problem further serves as an illustration of the Coulomb formulation.

Three examples are presented in Section 6.3 to compare the sawtooth asperity and logarithmic spiral models. The first two examples are numerical simulations of laboratory direct shear tests on initially fully mated rock joints. The properties of the two models are matched as far as possible to the simulations reported by Plesha[36]. The third example is a highly idealised model of a deep tabular excavation in rock.

The following section, Section 6.4, contains two comprehensive examples of a deep tabular excavation in unbounded rock. The first example considers mining through a number of existing vertical fractures which dilate when sheared. The response predicted by the Coulomb model and sawtooth asperity models are compared. The second application includes both parting planes and vertical fractures. This example is mainly intended to illustrate that a large number of discontinuities can be handled effectively by the formulation, and that complex behaviour can be captured. The Coulomb model with an associated flow rule is chosen to show this.

Before we consider these applications, it would be appropriate to first deal with some of the issues relevant to deep level gold mining in South Africa. There are essentially two features, the horizontal parting planes and vertical shear fractures, that are of importance in this study. The parting planes are geological discontinuities which separate the rock into approximately 1 m thick layers. The vertical fractures on the other hand are mining induced fractures. They initiate ahead of the face at regular intervals and may eventually extend several meters into the hangingwall and footwall material. We will not consider the formation of these fractures, but rather concentrate on the deformation along existing vertical fractures. The vertical shear fractures and the horizontal parting planes, divide the rock mass into relatively intact blocks. The distribution of stress and deformation in the proximity of the excavation is influenced by deformation occurring along these discontinuities.

An elastic analysis, which ignores the presence of any of these discontinuities, suggests that the rock surrounding the excavation (stope) is supported by three different mechanisms. The first mechanism exists far from the stope where material is undisturbed and hence held up by the geostatic stress state. The direction of the major principal stress closer to the excavation follows a trajectory which indicates that material is supported by an arching effect across the stope span. Finally, the horizontal component of the stress vector at the stope face is in a direction away from the face and causes stretching in the layers near the excavation, and hence a horizontal tensile stress zone develops between the arch and stope.

The weak parting planes allow layers near the excavation to separate from adjacent layers. Each layer then acts as a separate beam supporting only its self weight. The behaviour of a beam above the stope differs from one below the stope. In addition to separation, slip may occur along layers that are still in contact. In the case where the friction coefficient along the plane is high, little or no slip occurs. On the other hand, when the coefficient of friction is assumed to be low, large shearing takes place. The slip is induced by the outward thrust of the arching effect and relieves material confined by parting planes from the stretching action. In the case where shearing is accompanied by dilation, the vertical stress across the bedding plane increases which in turn arrests further sliding.

The hangingwall beam, consisting of the assumed 1 m thin layer between the hangingwall face and the first bedding plane above the stope, experiences complex behaviour in the theoretical case where no support is provided in the mined out region. The beam is traversed by vertical shear fractures

which separate the beam into relatively intact blocks. These shear fractures typically undergo sequences of open, stick and slip states as the excavation advances. A beam so composed has a remarkable degree of stability; this is a result of arching action within the depth of the beam. There are two regimes of behaviour. The beam first acts as a shallow arch, and in the absence of support, will eventually snap through until closure takes place at the centre of the beam. Once contact with the footwall is made, a steady state, stable regime is established, with the closure point following the advance of the face.

In addition to these mechanisms, another important mechanism (associated with vertical fractures) is present. It would appear that shear reversal takes place on vertical cracks as the stope advances through the crack.[44] This suggests that the behaviour of a crack under reversed shearing is a potentially important factor in the analysis. In particular, dilation is expected to occur as sliding takes place. If reversed sliding takes place different physical models of the crack suggest that, on the one hand, further dilation occurs, or that contraction, on the other, may occur. These limiting cases lead to significantly different stress distributions in the rock mass above the hangingwall. This mechanism will be illustrated in Section 6.4.

## 6.1 VERIFICATION OF THE INTERFACE ELEMENT

### 6.1.1 Wedge Embedded Between Blocks

The first example consists of a wedge embedded between two blocks.[16] The two blocks are initially in contact, with the bottom sides fixed, as shown in Figure 6.1. The continuum, as well as the contact surfaces, are discretised with quadratic elements. A horizontal pressure of 10 MPa is applied to the sides of the two blocks and the wedge is displaced downwards a distance of 10 mm.

Figures 6.2 and 6.3 show the variation of normal and shear stress along the contact surfaces, measured along a vertical line, from the bottom of the blocks to the top of the wedge. These results indicate that contact is maintained only along the bottom 12 m of the blocks, and between the wedge and blocks, with inelastic shear deformation occurring along the entire wedge contact surface.

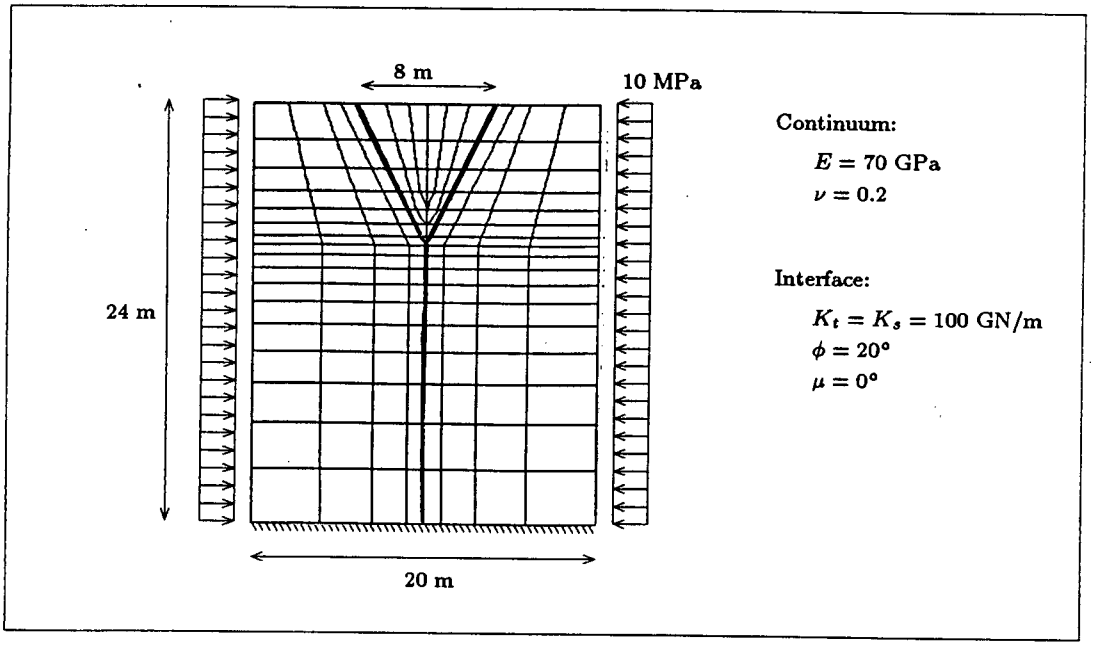



Figure 6.1: Wedge embedded between blocks

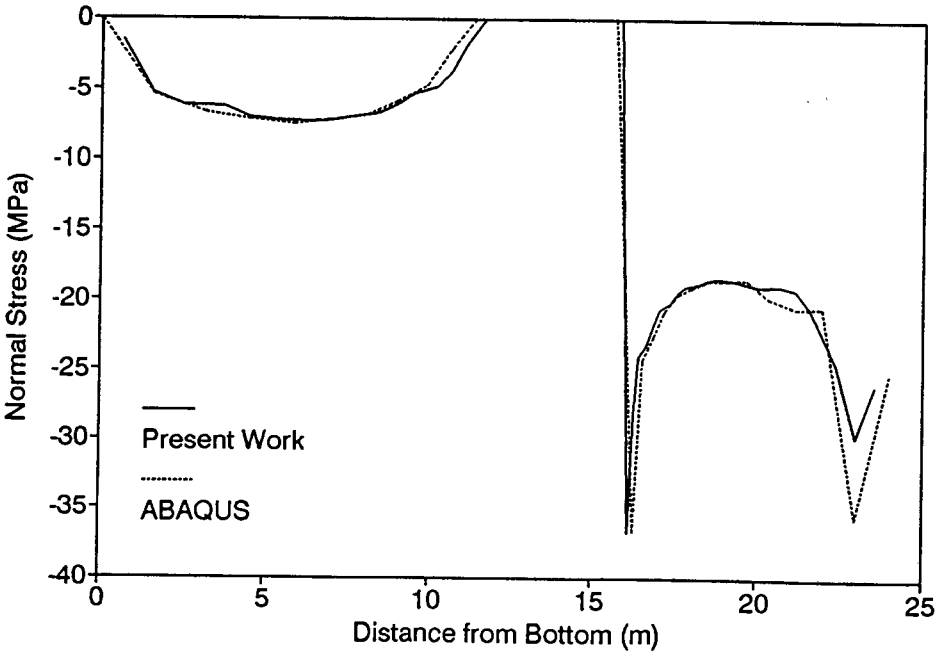



Figure 6.2: Normal Stress along Contact Area

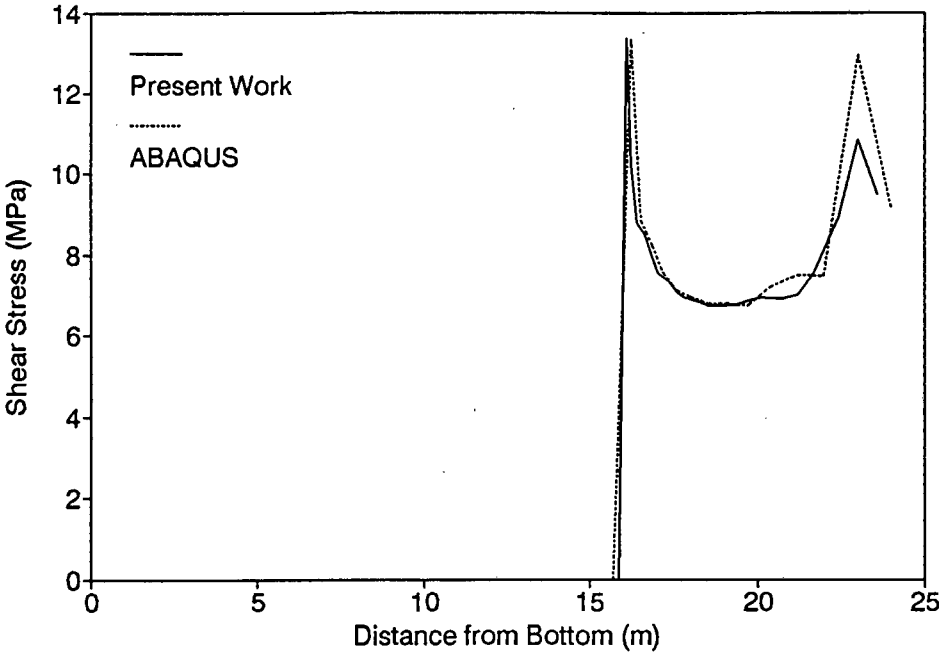



Figure 6.3: Shear Stress along Contact Area

Figures 6.2 and 6.3 also compare stresses obtained with one of the interface elements in the ABAQUS library with the interface element formulated in this thesis. ABAQUS uses a mixed formulation where the contact conditions normal to the faces are enforced by Lagrange multipliers in a rigid sense, and the tangential conditions by an elastic compliance or shear stiffness in an elasto-plastic sense. The shear stress is limited by a non-associated Coulomb friction law with no dilation.

It is interesting to note that the normal stress obtained with the present formulation agrees well with that of the rigid formulation. A rigid contact condition can therefore be modelled accurately by elasto-plastic conditions, even when a relatively low normal stiffness (100 GN/m) is used. This is encouraging since the penalty formulation is computationally less expensive than the Lagrangian formulation because fewer equations have to be solved, and the rate of convergence is better, provided that elastic constants are not too large. It was further found that a high normal stiffness does not necessarily improve the accuracy of results (oscillations in contact stress may occur), and that it adversely influences the rate of convergence. Oden and Kikuchi[30] also reported oscillations in stress, and suggested the use of reduced integration to eliminate the problem.

## 6.2 VERIFICATION OF COULOMB MODEL

### 6.2.1 The Hangingwall Beam Problem

We consider here the idealised case where the hangingwall beam is intersected by only one vertical fracture. The hangingwall beam model, as well as the discretisation near the fracture, is shown in Figure 6.4. The half span is 30 m with the fracture 10 m from the centreline. We assume that the beam is initially supported along its length and fixed at the end. The mining action is simulated by removing the support in 0.5 m steps, starting from the centreline. Equilibrium is achieved after each step, and the behaviour of the fracture plane is monitored as it goes through a sequence of opening, closing and sliding states.

Figure 6.5 shows the displaced shape of the beam near the fracture at four different stages during the excavation process. The different stages are numbered alphabetically, and will be referred to in subsequent figures. Figure 6.5a depicts the displaced shape when the mining face is 0.5 m behind the fracture. The fracture is open at the top and closed at the bottom, and is accompanied by a positive relative shear displacement due to the curvature of the beam (Figures 6.6 and 6.7). The point denoted 1 in these figures indicates the position of the apex at this stage of the excavation; dilation has thus taken place at the fracture top although the surfaces are separated. Figure 6.8 indicates that the normal stress at the bottom is initially zero and increases, accompanied by positive shear stress, as the mining face approaches the fracture. No yielding takes place during this stage.

Figure 6.5b shows the displaced shape when the mining face is just ahead of the fracture. The unsupported side of the hangingwall displaces downwards and the relative shear displacement changes sign as indicated by Figures 6.6 and 6.7. Consequently the shear stress at the bottom changes sign, and yielding takes place (Fig. 6.8). The normal stress at the bottom will now reduce as the span increases further and the fracture moves into the sagging area of the beam. The fracture surfaces will eventually separate at the bottom and come in contact at the top (Fig. 6.5c and d). The total relative normal displacement at the top of the fracture is still positive at stage c of the excavation, as indicated by Figure 6.7, and the contact surfaces in Figure 6.5c are therefore drawn separated. The apex finally moves to the point denoted 2 in Figures 6.6 and 6.7. No yielding will occur (Fig. 6.9) since the fracture is now in the low shear stress region of the beam.

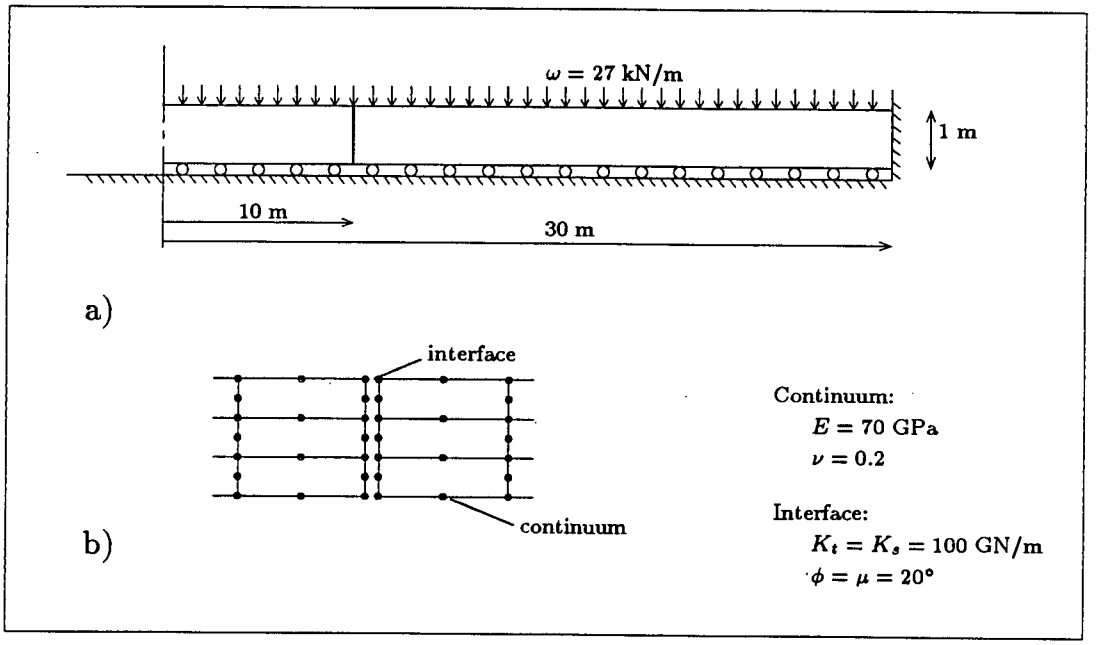



Figure 6.4: a) Hangingwall beam, and b) Discretisation near Fracture

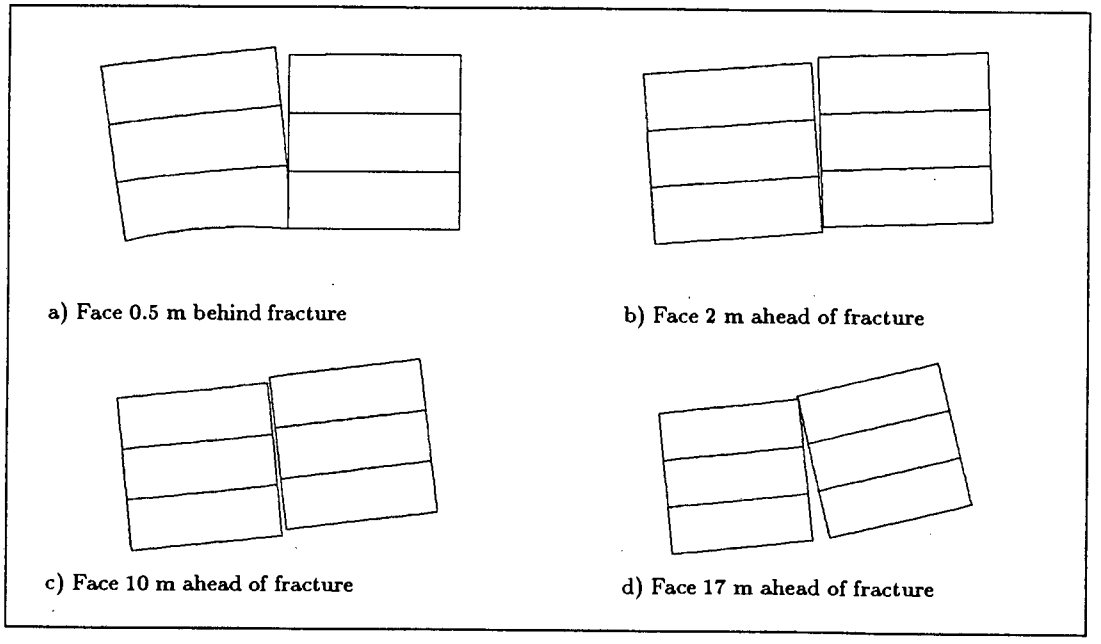



Figure 6.5: Fracture Deformation

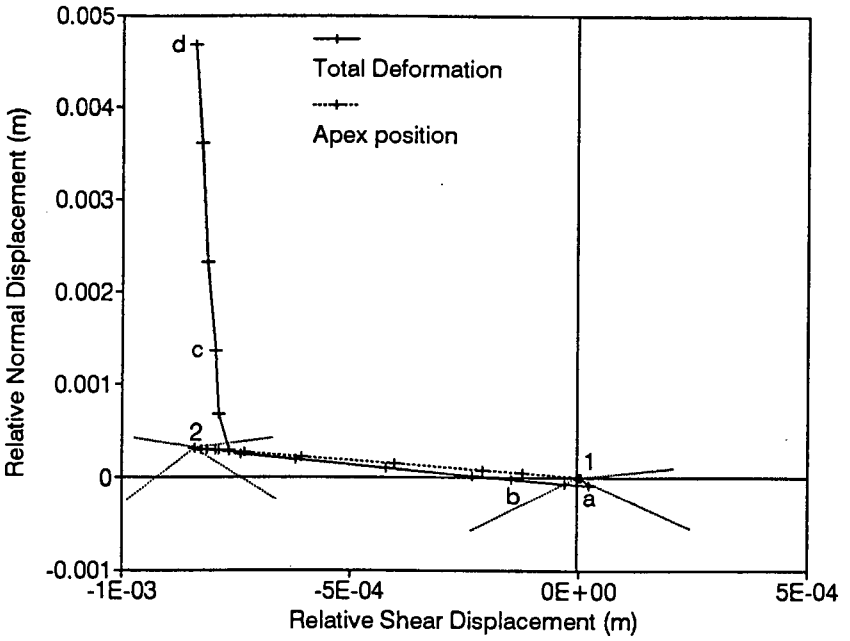



Figure 6.6: Total Deformation and Apex Position at Bottom of Fracture

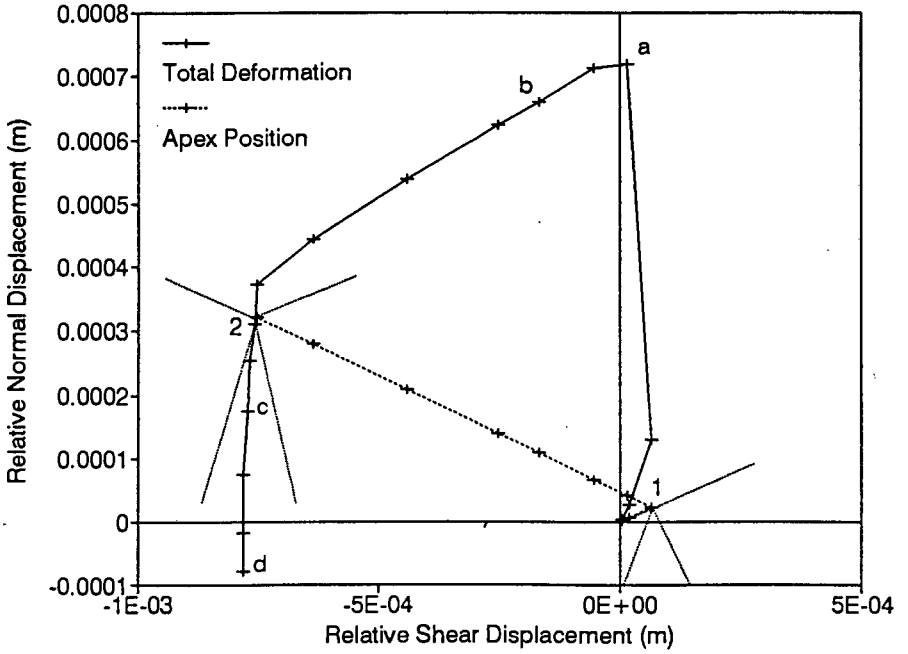



Figure 6.7: Total Deformation and Apex Position at Top of Fracture

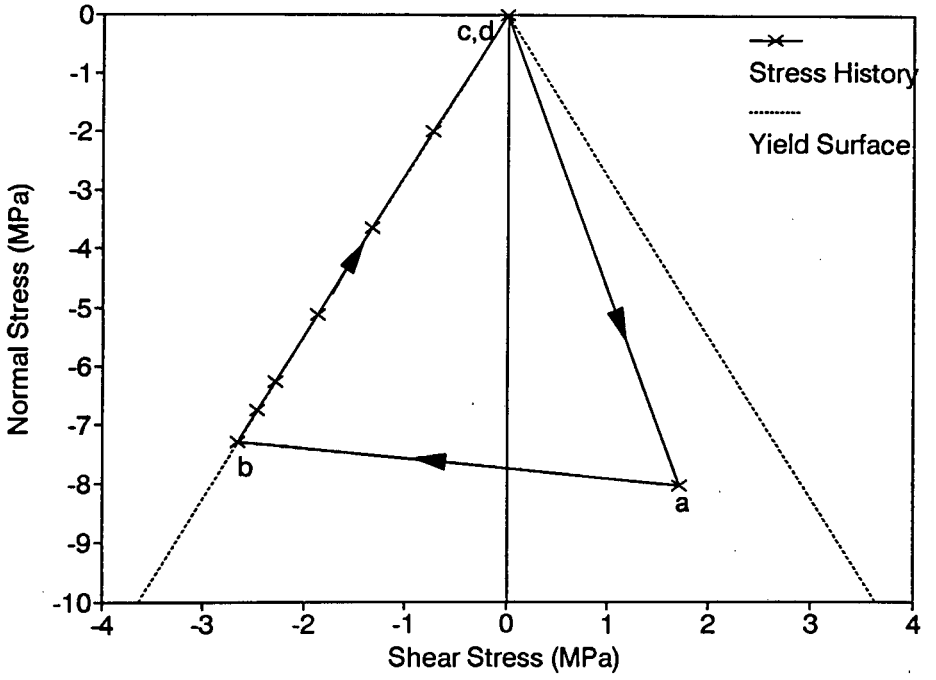



Figure 6.8: Stress State at Bottom of Fracture

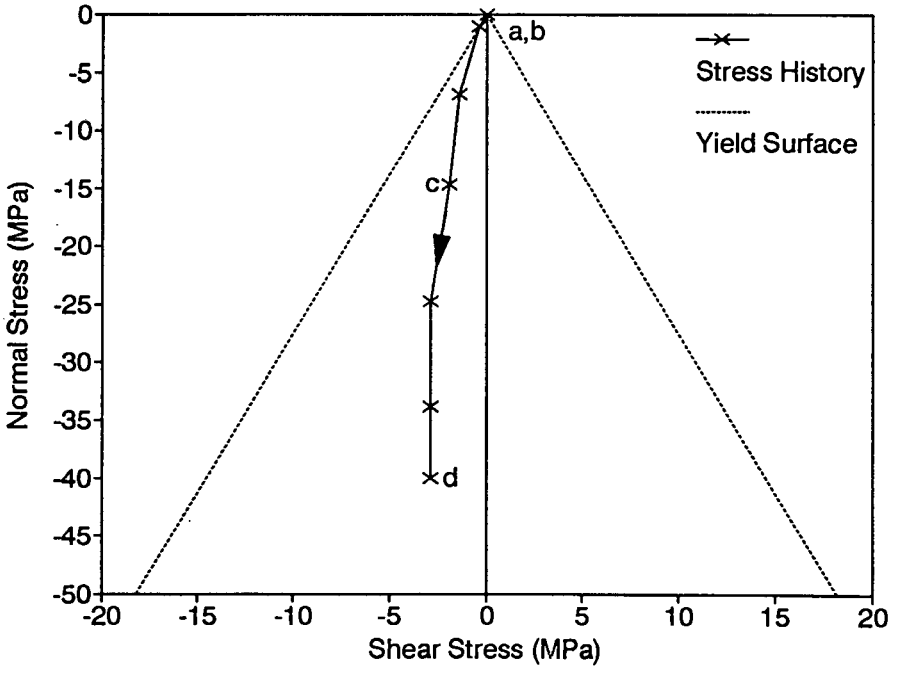



Figure 6.9: Stress State at Top of Fracture

## 6.3 VERIFICATION OF THE ASPERITY MODELS

### 6.3.1 Unconstrained Direct Shear Test

Direct shear tests on rock joints are usually carried out under constant normal stress conditions which allow the joint to dilate freely. Experimental results for these tests are readily available in literature.[7,24,27] The example reported here is a simulation of such a test and considers the behaviour predicted by the sawtooth and spiral models, as well as the response of the spiral model at different normal stress levels. The properties of the model are matched as far as possible to the properties used by Plesha[36] to simulate Kutter's[27] cyclic shear tests. The rock specimen and finite element model are shown in Figure 6.10. The specimen is 10 cm long and 4 cm thick, with a discontinuity running along the center parallel to the bottom and top plates.

The rock is assumed to behave elastically with Young's modulus $E = 1$ GPa and Poisson's ratio $\nu = 0.3$. The interface on the other hand behaves according to either the sawtooth or spiral model. The properties common to both joint models are $K_t = K_s = 0.08$ GN/m and $\phi = 32°$. An initial dilation angle of $\mu° = 16°$ is assumed in the case of the spiral model. This gives a total friction angle of 48° at the onset of yielding, which reduces to a final value of 32° as shearing takes place and the maximum normal deformation of $\delta^m = 2$ mm is reached. In the case of the sawtooth model the dilation angle is assumed to be $\mu = 16°$, with the total friction angle remaining constant at 48° throughout the entire loading history.

The loading consists of a uniform vertical pressure of 50 kPa, 90 kPa or 200 kPa, as well as a prescribed horizontal displacement history, both applied to the top plate of the model. The bottom of the model is constrained against horizontal and vertical displacements. The response of the interface at one of the Gauss points is monitored as shearing takes place.

A complete shear stress-shear deformation cycle, at a constant normal stress level of 90 kPa, is shown in Figure 6.11. The diagram shows elastic-perfectly plastic response predicted by the sawtooth model, and softening behaviour predicted by the logarithmic spiral model. The inelastic normal deformation (Fig. 6.12) increases linearly with increasing shear deformation in the case of the sawtooth model, but reaches a finite value of 2 mm in the case of the spiral model. The figure further shows that the unloading path differs from the loading path due to the approximation associated with contraction, but

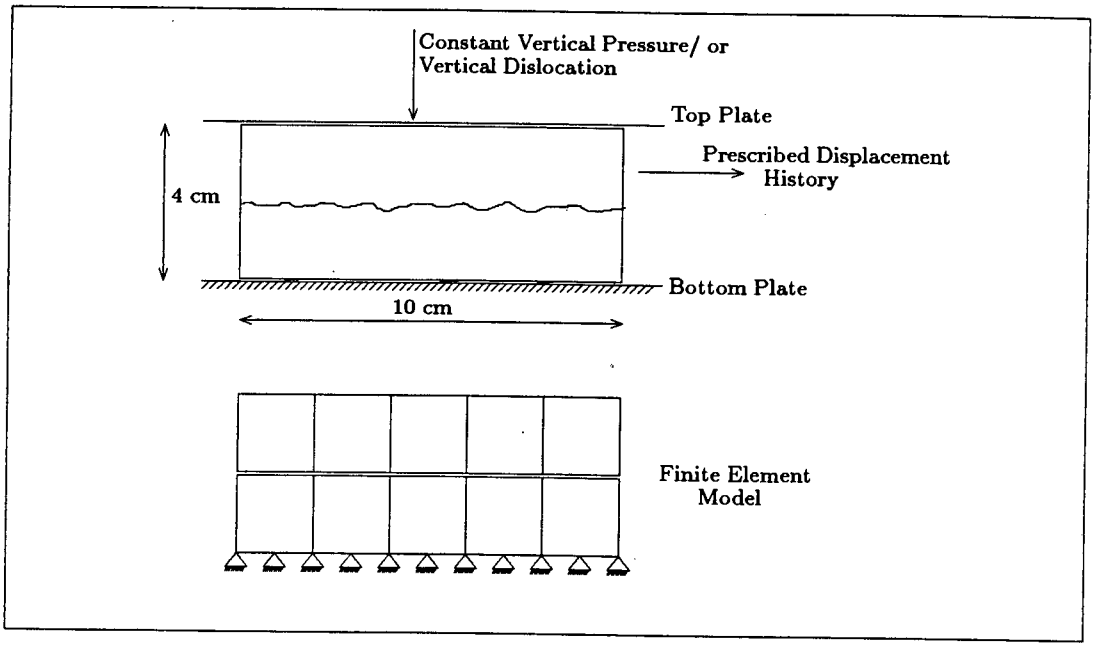



Figure 6.10: Direct shear test model

within an acceptable tolerance. The behaviour predicted by the spiral model agrees favourably, in a qualitative sense, with the cyclic shear test reported by Hutson and Dowding[24] and Kutter and Weisbach.[27]

Figure 6.13 compares the shear response predicted by the spiral model at different levels of normal stress, showing *brittle* behaviour at high normal stress with almost *ductile* response at lower normal stress levels. This behaviour is also in accordance with laboratory observations.[5,7,24] The incidence of peak behaviour is, however, not predicted adequately: a joint under high normal stress conditions is expected to reach peak behaviour at smaller shear deformation than a joint at low normal stress as indicated in Figure 2.2 of Chapter 2. A model that allows such behaviour must be based on a curved failure envelope.

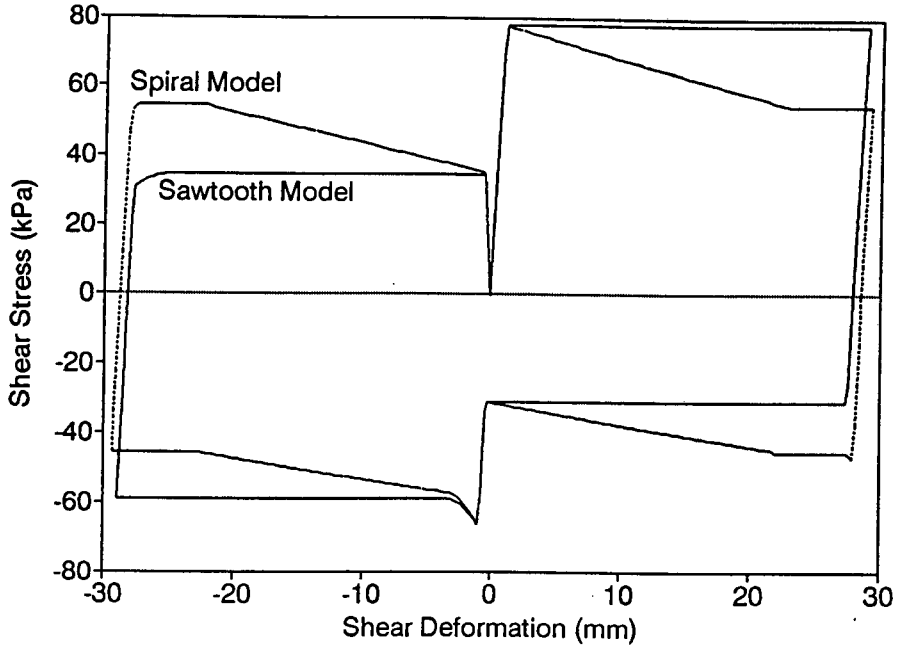



Figure 6.11: Complete shear stress-shear deformation cycle

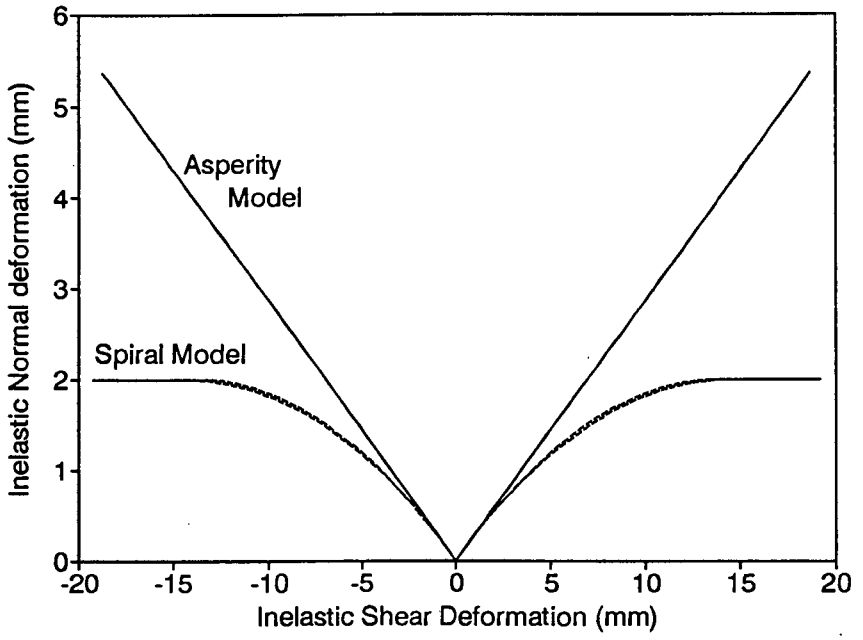



Figure 6.12: Inelastic deformation history

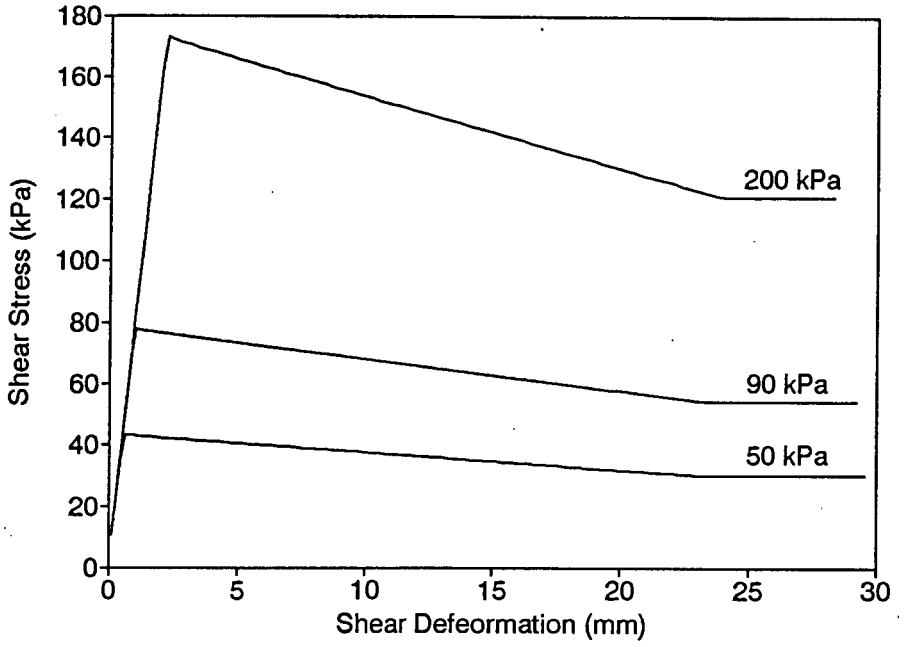



Figure 6.13: Response at different normal stress levels

### 6.3.2 Constrained Direct Shear Test

The next example is direct shear test on the rock specimen shown in Figure 6.10, with the top plate constrained against vertical displacements so that the specimen is no longer free to dilate. Inelastic normal deformation along the joint then leads to an increase in vertical stress through the thickness of the specimen, which in turn arrests further sliding.

The geometry of the specimen and the finite element model are the same as that of the first example. The elastic constants of the rock are assumed to be $E = 10$ GPa and $\nu = 0.3$, and the interface properties are $K_t = K_s = 10$ GN/m and $\phi = 26°$. A dilation angle of $\mu = 10°$ is chosen for the sawtooth model and an initial dilation angle of $\mu^o = 10°$ is chosen for the spiral model. Dilation in the case of the spiral model ceases when $\delta^m = 1$ mm.

The top plate of the model is displaced downwards an amount sufficient to produce an initial vertical stress of approximately 10 MPa across the interface. Prescribed horizontal displacement increments are applied subsequently, and the response of the interface monitored at a Gauss point.

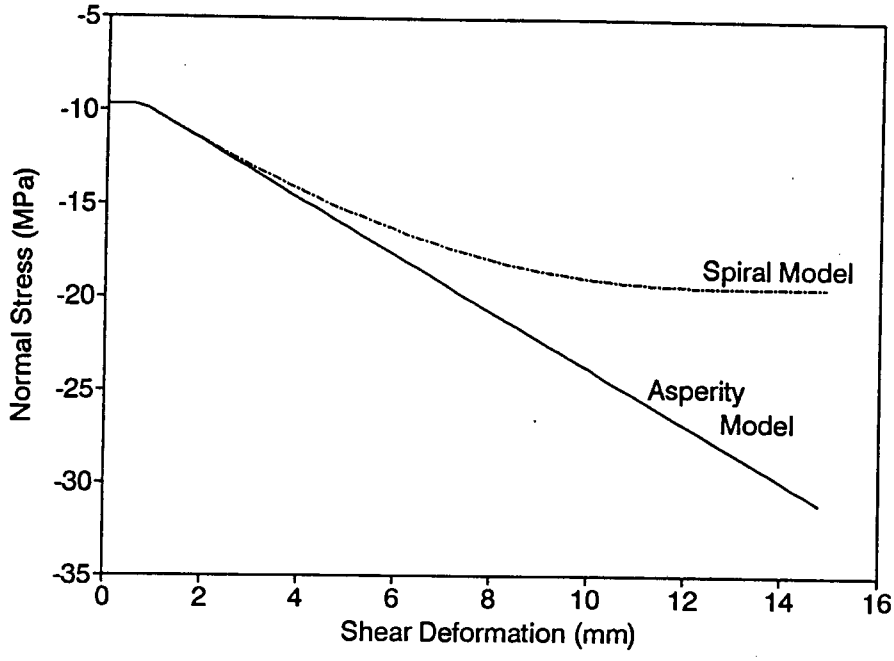



Figure 6.14: Variation of normal stress with shear deformation

Figures 6.14 and 6.15 show the variation of the normal and shear stress, respectively, as shearing takes place. After some elastic behaviour, inelastic deformation commences with the normal and shear stress increasing monotonically in the case of the sawtooth model, but reaching finite values of $\sigma \approx 20$ MPa and $\tau \approx 10$ MPa in the case of the spiral model. The evolution of the yield surface, as well as the locus of stress points $\sigma, \tau$ during the loading history, are shown in Figure 6.16. In the case of the sawtooth model the stress points remain on the yield surface after the initial elastic deformation, but move further away from the origin as shearing takes place. In the case of the spiral model the yield surface rotates from its initial position (which coincides with the sawtooth yield surface) to the final position as indicated. Once the final position has been reached, the stress point remains at a fixed point on the yield surface. Shearing then takes place without further dilation as shown in Figure 6.17, resulting in finite normal and shear stress across the joint.

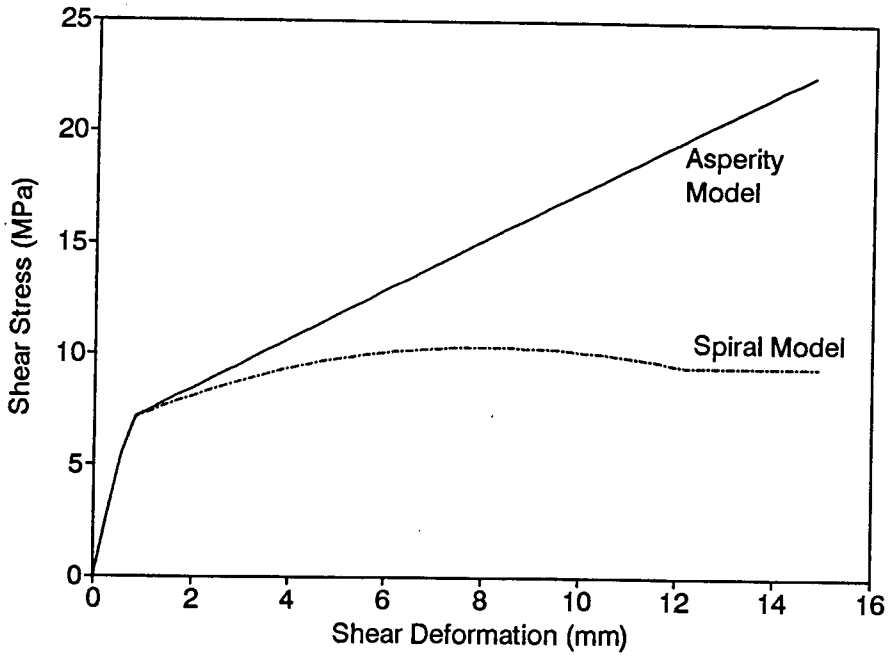



Figure 6.15: Variation of shear stress with shear deformation

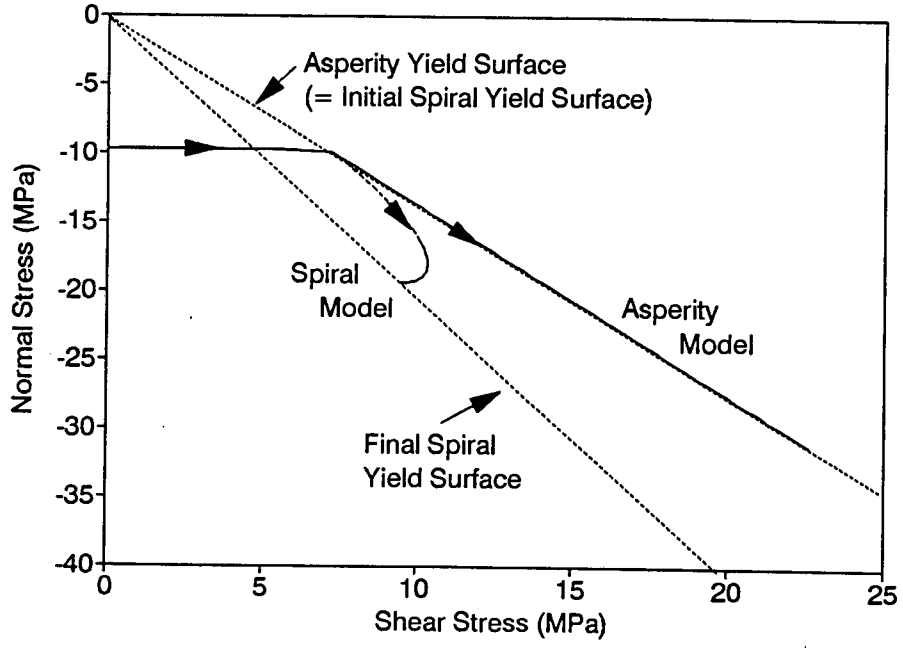



Figure 6.16: Response in stress space

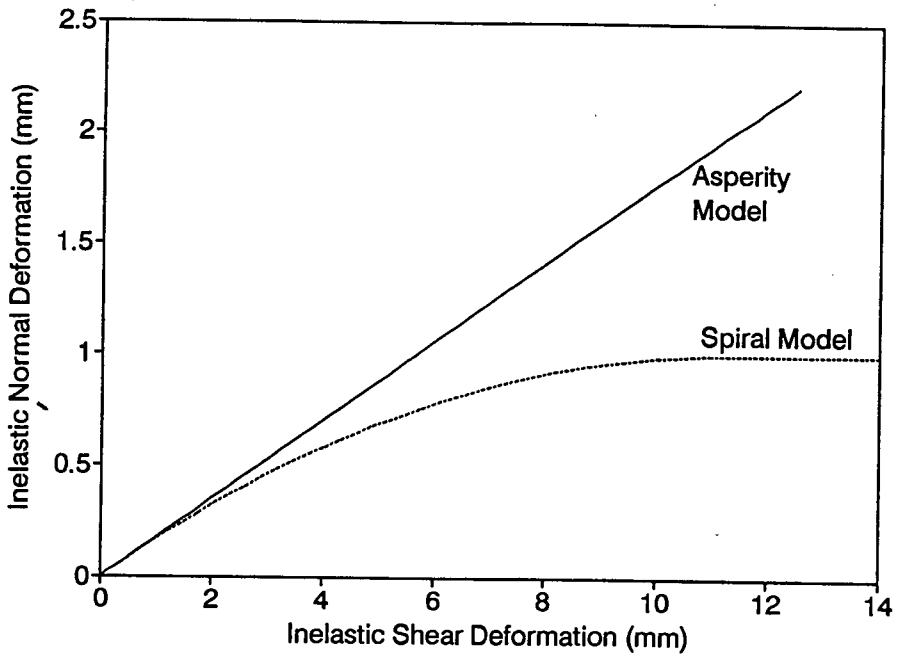



Figure 6.17: Variation of inelastic deformation

### 6.3.3 Deep Tabular Excavation in Rock

The last example in this section is an idealised model of a deep tabular excavation in rock in which only vertical shear fractures are present. Significant shearing may occur along these fractures. Due to the confined nature of the rock, dilation manifests itself through increasing horizontal stress, which in turn has a stabilising effect on the rock mass near the excavation.

The excavation is modelled as a narrow slit advancing horizontally and symmetrically about a vertical centreline. The surrounding rock is assumed to be unbounded and vertical gravitational forces are present resulting in a stress field of $\sigma_x = -15$ MPa and $\sigma_y = -60$ MPa at the level of the excavation. We simplify the problem further by assuming that the geotechnical stress field is constant throughout the depth. This enables us to assume horizontal symmetry through the centreline of the excavation, and hence to model only a quarter of the problem. The idealised model is shown in Figure 6.18. The rock mass is represented by one continuum element and one infinite domain element, and the vertical fracture with one interface element. Mining is simulated by removing the bottom row of vertical supports.

There are two actions present: a driving action (due to the self weight) which causes the rock mass to displace downwards, and a resisting action due to frictional effects along the interface. The response of the structure, monitored at the midside Gauss point of the interface element, is shown in Figures 6.19 to 6.22. The response predicted by the sawtooth model is labeled *abc* and the response predicted by the spiral model *abdef* in these figures. The initial shear stress along the joint is zero and the normal stress 15 MPa. Vertical displacement of the rock mass causes the shear stress along the joint to increase, first elastically with the normal stress decreasing due to Poisson's ratio effect, until the yield surface is reached at point *b*. Further shearing is then accompanied by dilation as shown in Figure 6.22, but depends on the friction law chosen.

The inelastic response predicted by the sawtooth model is straightforward. The locus of stress points $\sigma, \tau$ remains on the yield surface but moves further away from the origin until equilibrium is reached after a vertical displacement of 8 mm (point c). The rock mass is now partly supported by the shear force along the joint, and partly by the farfield material.

Since the inelastic normal deformation is limited in the spiral model, and since the dilation angle decreases with continuing shear deformation, the rock mass displaces downwards much further before equilibrium is reached. In this case the normal stress across the joint is higher than that predicted by the sawtooth model, but the shear stress is lower.

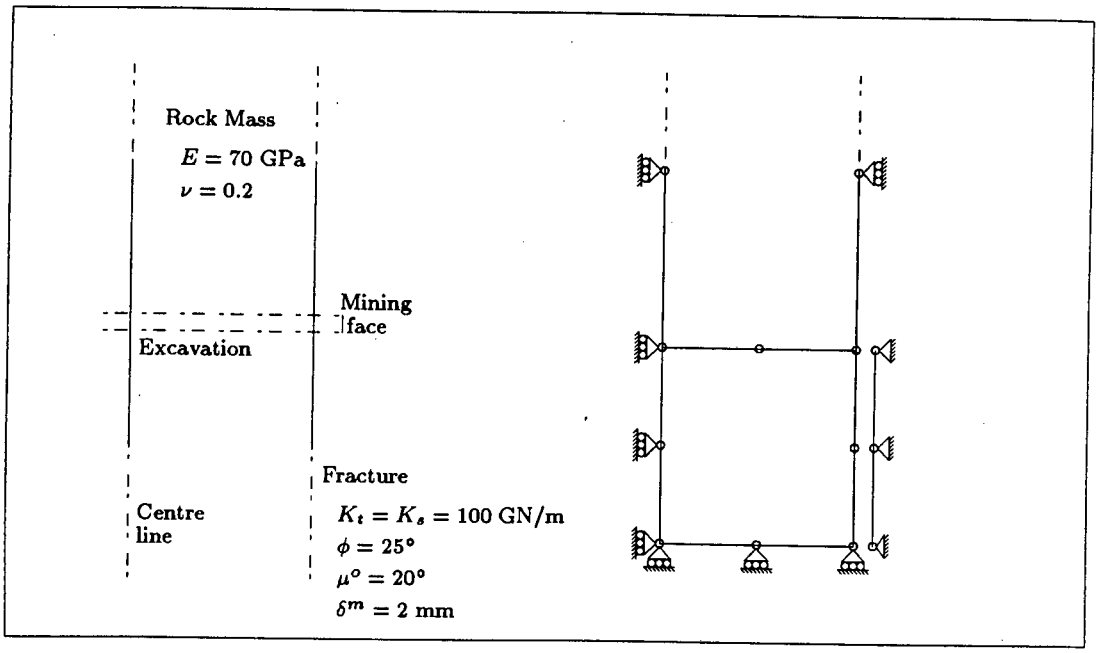



Figure 6.18: Idealised excavation in rock

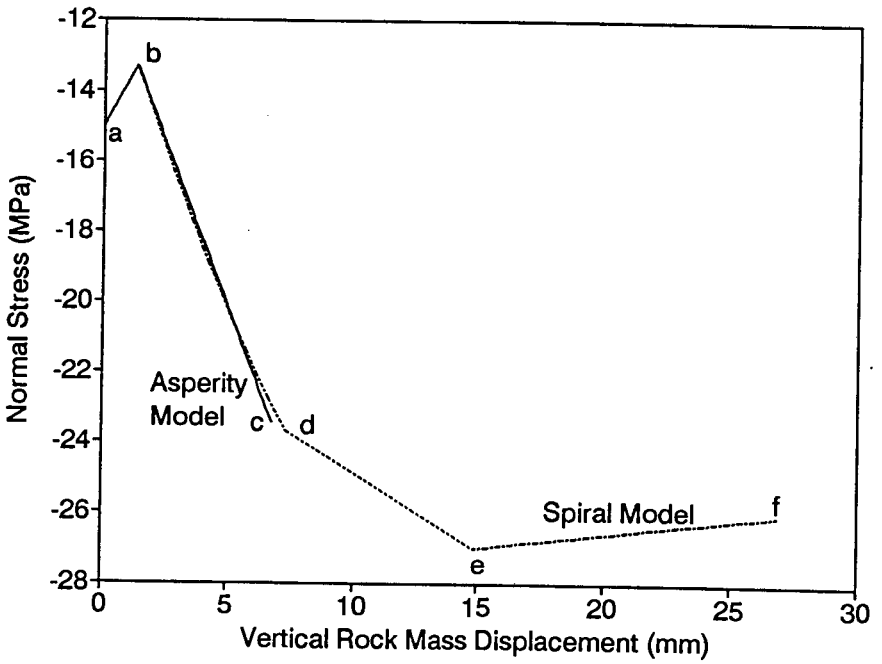



Figure 6.19: Variation of normal stress with shear deformation

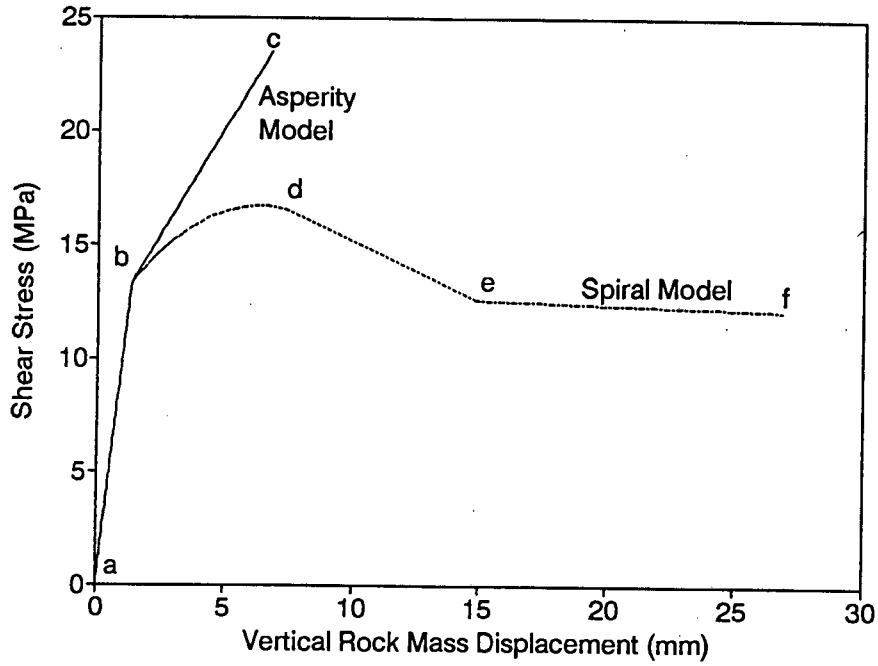



Figure 6.20: Variation of shear stress with shear deformation

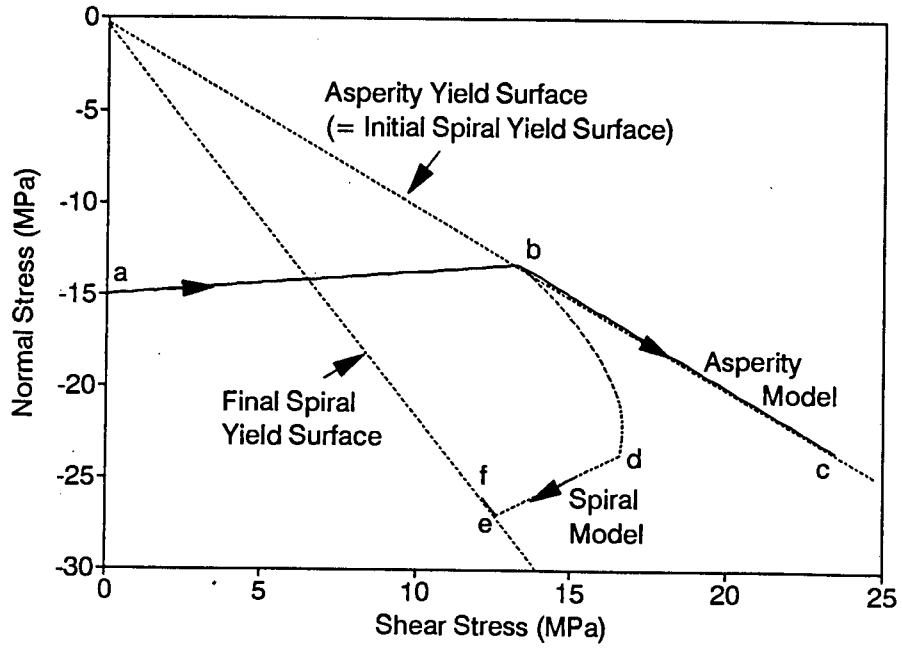



Figure 6.21: Response in stress space

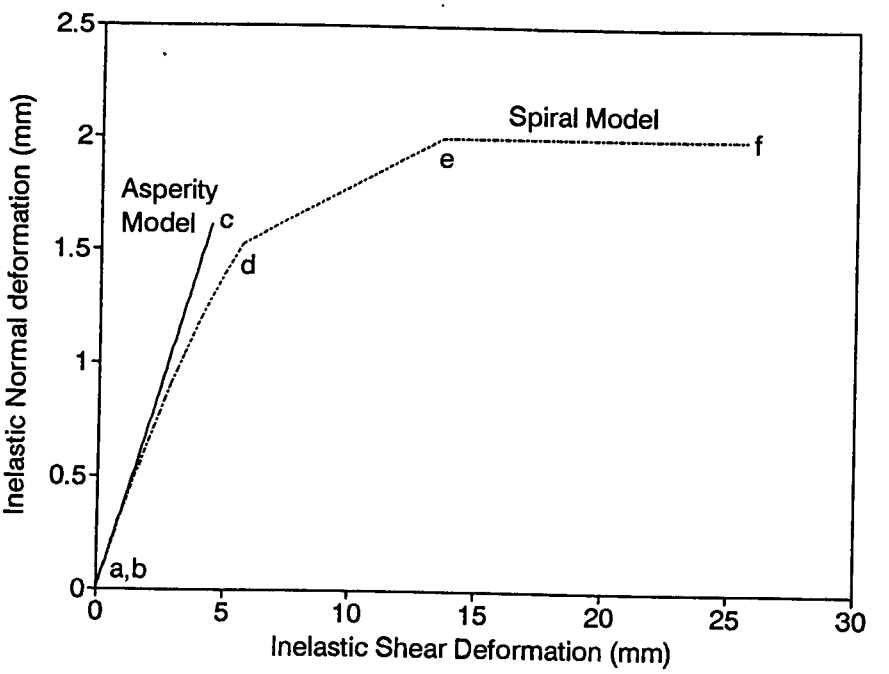



Figure 6.22: Variation of inelastic deformation

## 6.4 ROCK MASS PROBLEMS

The first example in this section considers mining through a number of existing vertical fractures which may dilate or contract when sheared. The response predicted by the Coulomb model and sawtooth asperity models are compared. The second application includes both horizontal parting planes and vertical fractures. This example is mainly intended to illustrate that a large number of discontinuities can be handled effectively by the formulation presented in this thesis, and that complex behaviour can be captured.

In both examples the finite element mesh extends 50 m above and below the stope, and is 90 m wide with one side representing a vertical symmetry axis, and infinite domain elements at the other boundaries. The rock mass is assumed to be elastic with Young's modulus 70 GPa and Poisson's ratio is 0.2. The elastic constants $K_t, K_s$ on all the discontinuities are taken as 100 GN/m. Both models are discretised with linear continuum and interface elements, except for the region near the excavation in example 2 which is discretised with quadratic elements.

A stress field due to gravitational and tectonic forces exists through the depth of the rock. It is assumed that this stress varies linearly with depth, and that the ratio between horizontal and vertical stress components is $k = 0.5$. We consider mining at a level where the stress field in the horizontal and vertical direction are -30 MPa and -60 MPa respectively. This linear varying stress state is achieved by superimposing a constant stress field of $\sigma_v = -60, \sigma_h = -30$ MPa onto a vertical body force of 27 KN/m$^3$. ABAQUS provides the facility where predefined elements can be removed from the model, and later be replaced if necessary. This feature is useful in simulating the advance of the excavation and placing backfill material in the mined out region.

### 6.4.1 Example 1

The first example comprises 5 vertical fractures spaced 6 m apart and extending 6 m into the hangingwall and footwall material. The stope is 1 m deep, and is advanced sequentially in steps of 6 m each by removing elements. Friction angles of 40° and 20° are chosen for the Coulomb and sawtooth models respectively. The dilation angle is assumed to be 20° in both cases.

Each of the fractures will experience the same behaviour during the loading history. We will consider the response at a point on the second vertical fracture, 1 m above the excavation and 12 m from the centreline, to illustrate this. Figures 6.24, 6.25 and 6.26 show the state of the fracture in stress space and state space, respectively, at the five loading or excavation steps. The point labelled 0 is the initial stress state (with a normal stress of 30 MPa across the joint), and the point labelled 5 the final excavation with the mining face 30 m from the centreline.

Both joint models predict an increase in normal stress (Fig. 6.24) across the joint during the first excavation when the fracture is still ahead of the mining face (i.e. the mining face is 6 m from the centreline), as indicated by the point denoted 1 in Figure 6.24. This increase is due to dilation occurring along the first fracture. Little shearing occurs along fractures ahead of the face at this stage. However, once the face and fracture coincides (the second excavation in this case), large shearing takes place. The position of the updated state diagram at this stage is labelled A in Figures 6.25 and 6.26. The behaviour predicted by the two models is still similar at this point. The slight difference is due to reversed sliding which now occurs along the first fracture. This reversed sliding, accompanied by either further dilation

or contraction along joints in the mined region, results in a different stress distribution in the rock mass, which in turn influences the behaviour along fractures at, and ahead of the face. The reversed sliding behaviour becomes evident when the next 6 m is excavated (excavation 3) and fracture 2 moves into the mined-out region. The Coulomb model predicts further dilation, and contact is therefore maintained along the fracture. The sawtooth model, on the other hand, allows contraction to occur. Since these fractures lie in the material that is subjected to stretching due to the arching effect, the fracture simply opens. The opening occurs during the fourth excavation.

Figure 6.27 shows the horizontal stress state at the end of the excavation with the face 30 m from the centreline. The stress is measured along a horizontal line 1 m above the excavation, from the centreline to a point 10 m ahead of the face. The figure shows the behaviour predicted by an elastic analysis without any of the discontinuities, as well as the behaviour predicted by the two models. The elastic analysis suggests a tensile stress over most of the excavation. When reversed sliding is accompanied by contraction, this stretching effect causes the joints to open, and a zero horizontal stress zone develops above the excavation. On the other hand, if reversed sliding is accompanied by further dilation, large horizontal stresses can be generated in the hangingwall layers.

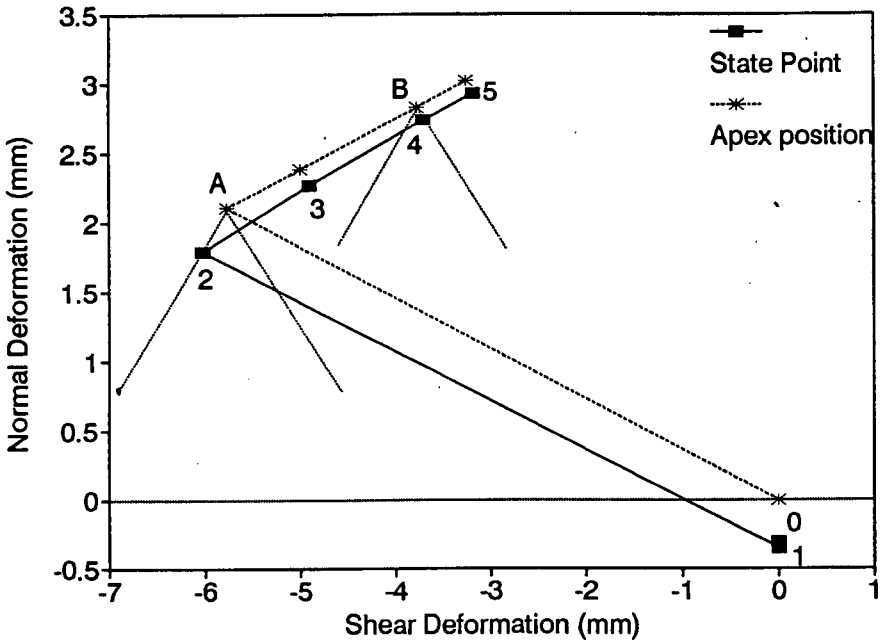



Figure 6.24: Joint history in state space (Coulomb model)

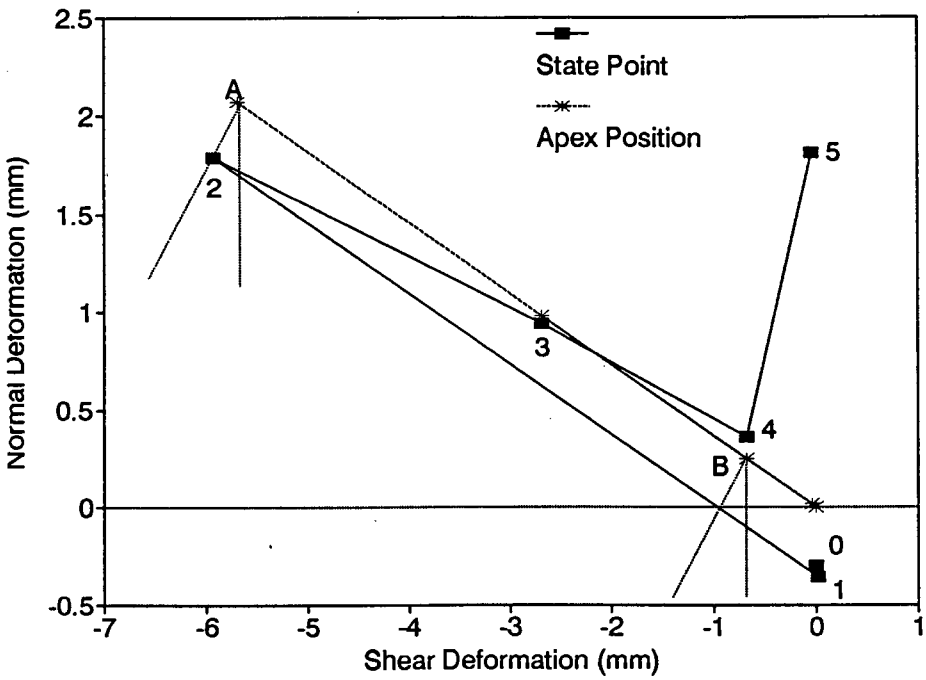



Figure 6.25: Joint history in state space (asperity model)

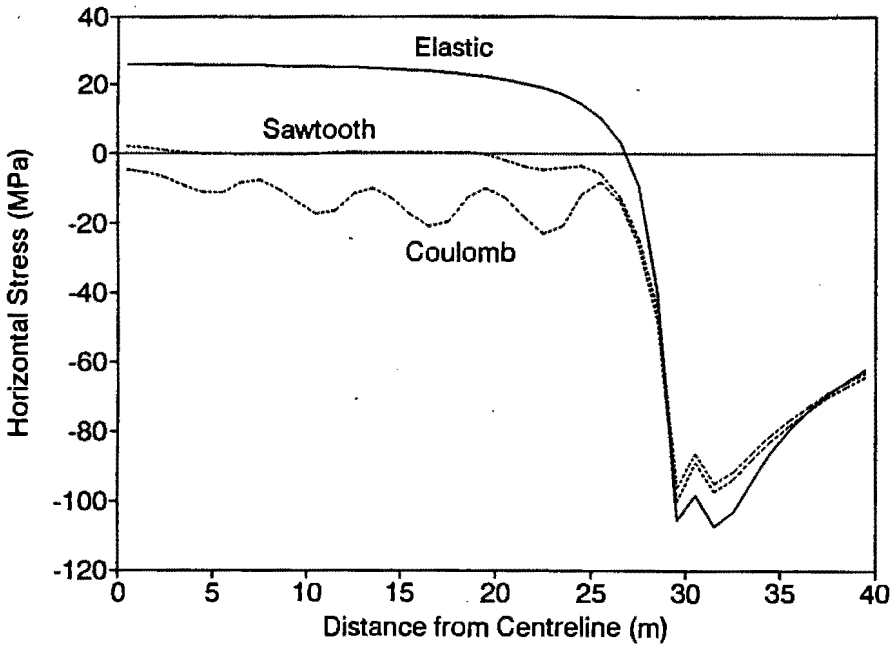



Figure 6.26: Horizontal stress 1 m above excavation

Figure 6.28 shows a contour plot of the horizontal stress distribution near the excavation for the case without backfill and the stope face 24 m from the centreline. The hangingwall beam separates from the overlaying material and becomes unstable at a halfspan of 24 m. Further mining is then impossible unless the beam is supported vertically. The arching effect in the beam can be seen clearly. The displaced shape of the hangingwall and footwall beams at different excavation steps are given in Figures 6.29 and 6.30 respectively. Figure 6.31 shows the vertical displacement at the hangingwall and footwall at the centreline as the mining face advances. Note the large deformation in the hangingwall beam during the last excavation as the beam becomes unstable. The footwall beam, on the other hand, separates from the footwall material and heaves to a maximum vertical displacement at midspan when the face is approximately 20 m from the centreline. The heaving is due to large horizontal compressive stresses in the mined region. The self weight of the beam overcomes this mechanism and causes the beam to displace downwards as the face advances further. Contact with the footwall material will eventually be re-established, with the closure point following the advance of the face. The horizontal stresses in the hangingwall (Fig 6.32) is compressive, and increases with an increasing stope span until the beam reaches the unstable point.

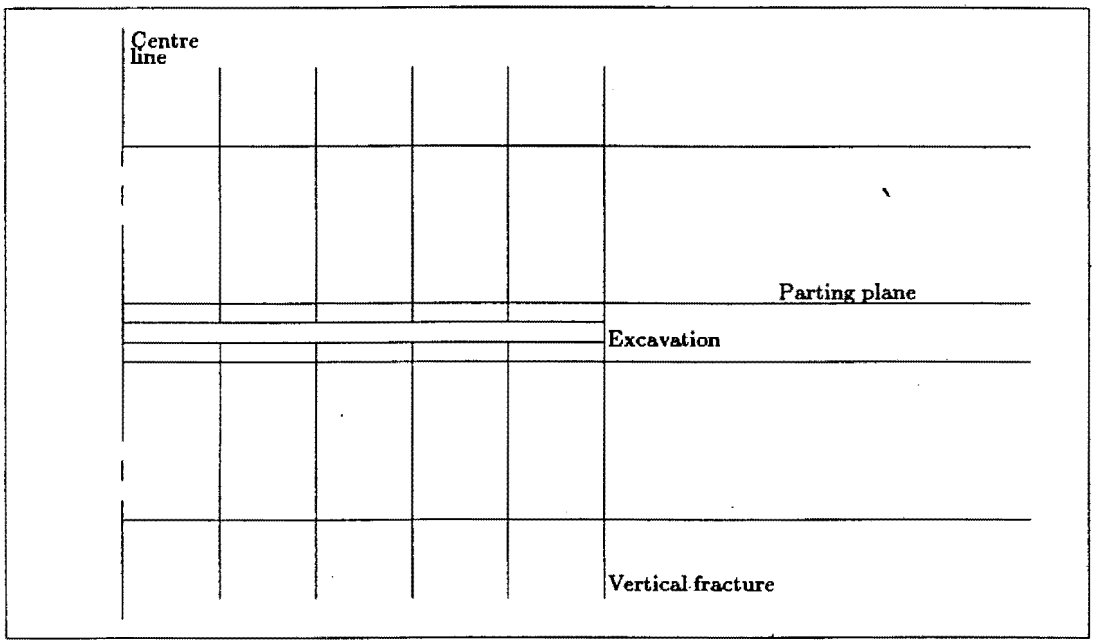



Figure 6.27: Rock mass problem

## 6.5 SUMMARY

We illustrated various aspects of the formulation in this chapter by means of a series of rock mechanics applications. These examples emphasised the fact that rock discontinuities can experience compex behaviour during their loading history, and that any constitutive model should cater for all possible histories of opening, sticking and sliding states, if the behaviour of the rockmass were to be captured adequitely.

The interface element behaves relatively well numerically provided that care is taken with the choice of especially the normal stiffness $K_t$. The magnitude of the normal stiffness in all the examples is of the same order, but larger than the Young's modulus of the rock mass. Such a choice will in most instances provide results of sufficient accuracy as illustrated in Section 6.1. However, it must be emphasised that this choice, as well as the associate accuracy, are also problem dependent.

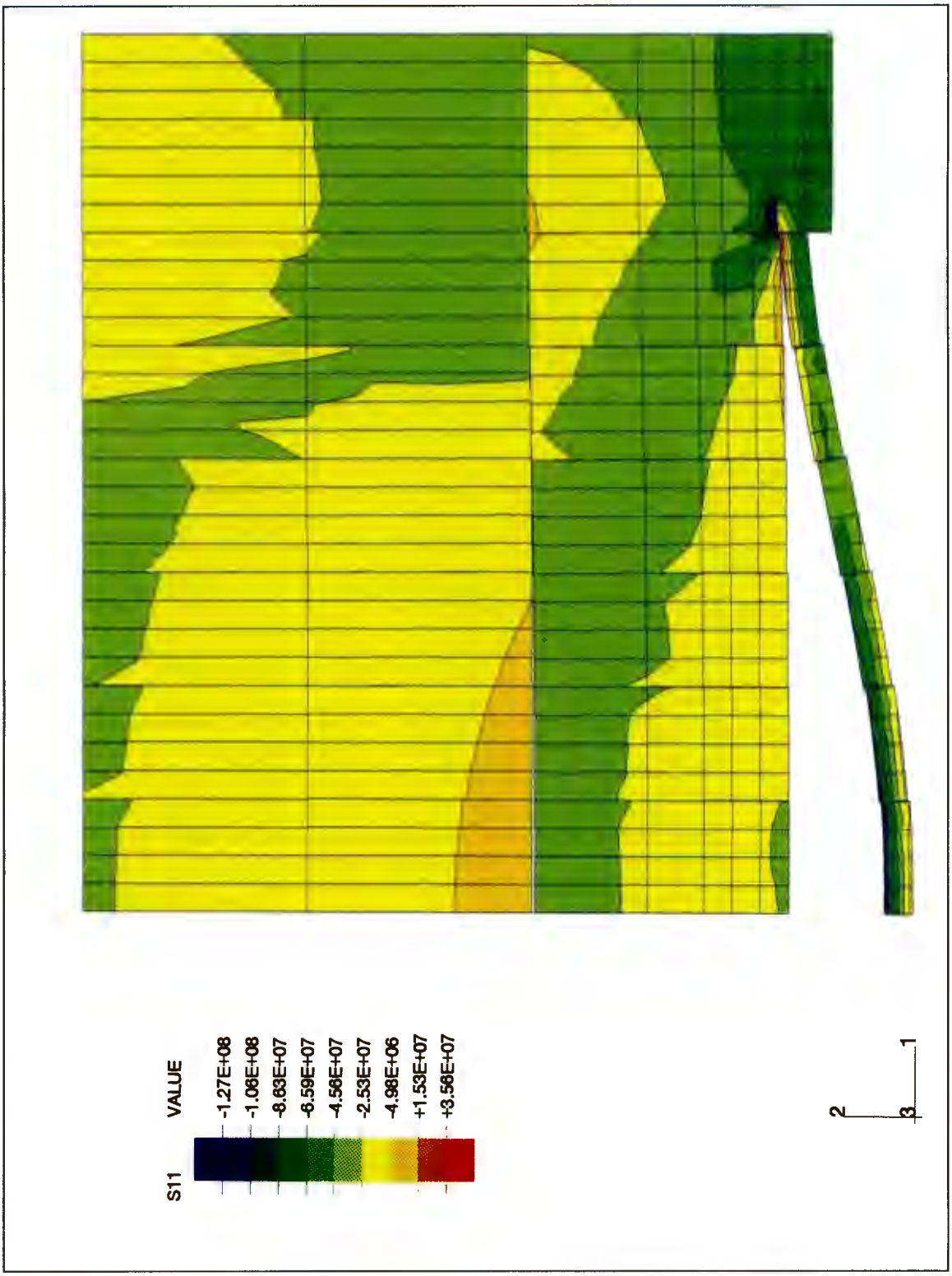



Figure 6.27: Horizontal stress distribution

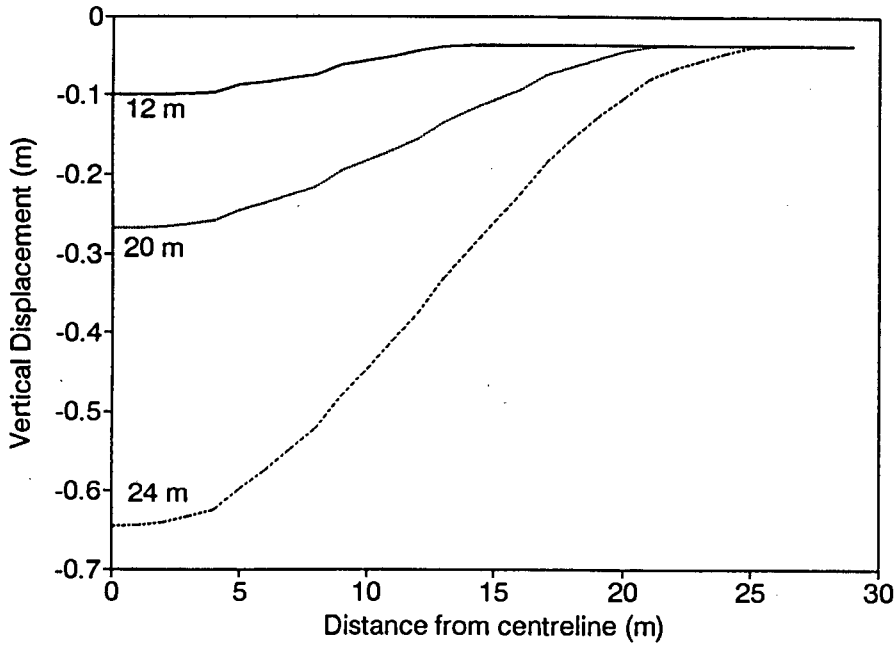



Figure 6.29: Displaced shape of hangingwall beam

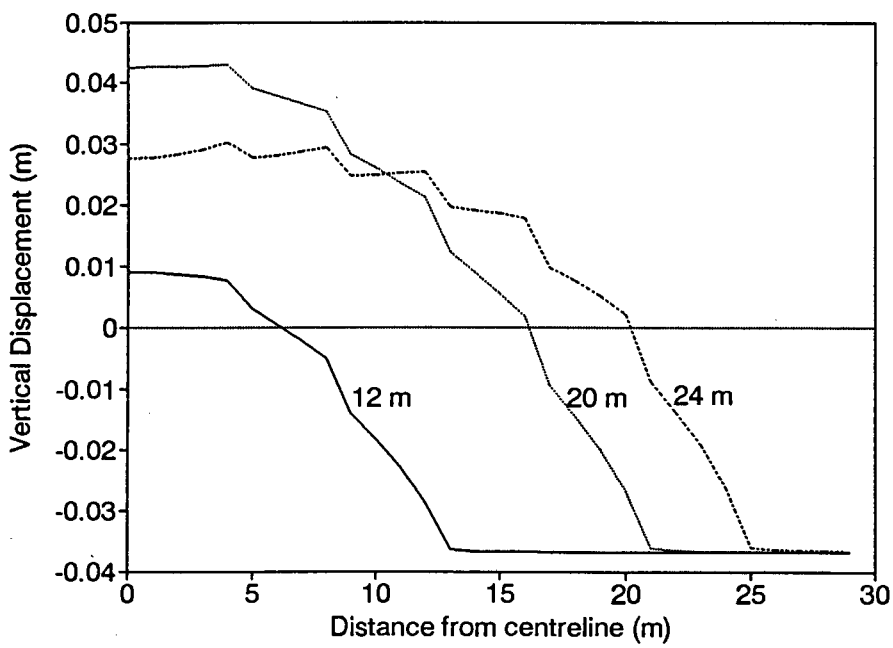



Figure 6.30: Displaced shape of footwall beam

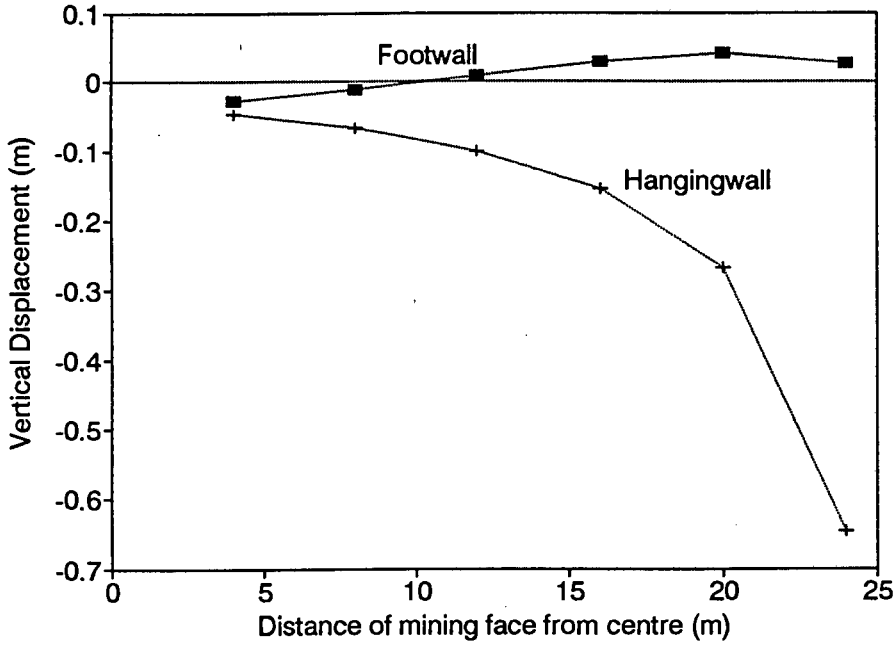



Figure 6.31: Vertical Displacement at midspan with stope advance

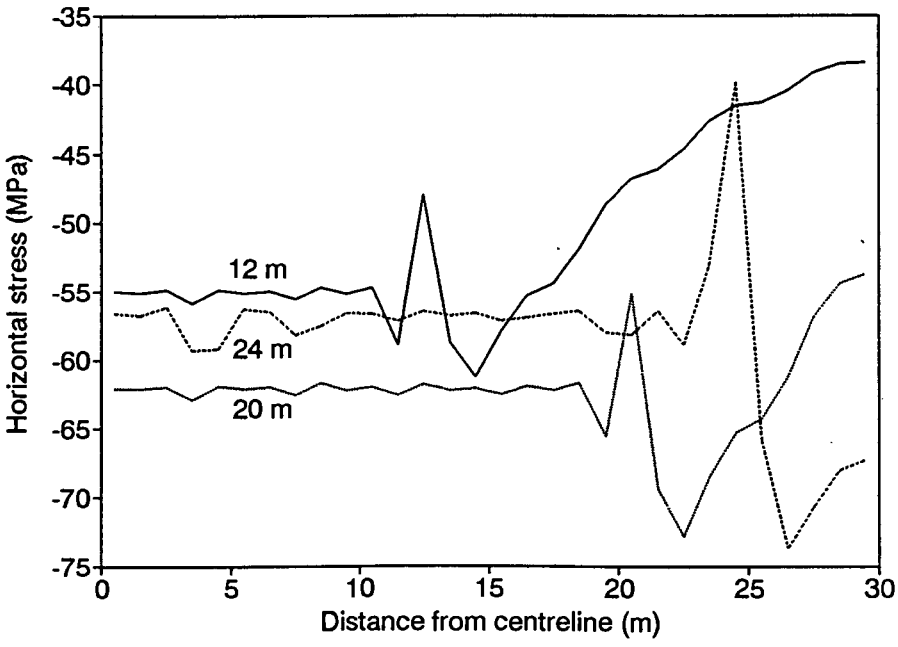



Figure 6.32: Horizontal stress in hangingwall beam

# CHAPTER 7
# CONCLUSIONS

The thesis considered the formulation of an interface element suitable for implementation into a standard non-linear finite element code. An isoparametric joint element has been implemented as a user element into the finite element program ABAQUS. The element is fully compatible with other plane elements in the code.

Particular attention was paid to ensure that the constitutive models cater for all possible histories of opening, closing and sliding (accompanied by dilation or contraction) in any direction, in a consistent manner. The formulation of these non-linear incremental equations was carried out on the basis of a backward difference discretisation in time. The advantage of this approach is that no qualitative distinction need be made between opening and closing, on the one hand, and sliding on the other. Further, a convenient formulation of the constitutive equations is facilitated by representing the different contact conditions in relative displacement (or state) space. Since the state diagram in relative displacement space changes from one time-step to the next, evolution equations for the updating are required. These concepts were illustrated for three rock joint models: a dilatant Coulomb model, a sawtooth asperity model and a logarithmic spiral model. The models are based on a penalty formulation to enforce the contact constraints, and explicit equations for the tangent stiffness matrix and for the corrector step of the Newton-Raphson iterative algorithm, were provided.

The predictions of the formulation, the differences between the three constitutive models, and the effectiveness of the model, were illustrated by means of a number of rock mechanics examples. The interface behaves relatively well numerically, but represents highly idealised rock joint behaviour. It is possible to incorporate more representative rock joint models within the framework of the current formulation, however, most extensions will result in increasing computational effort. The major consideration here is whether the improved constitutive model would justify the corresponding computational cost.

# REFERENCES


1. B Amadei and S Saeb, "Constitutive models of rock joints", (Edited by N Barton and O Stephansson), *Proc. Int. Symp. Rock Joints*, Loen, Norway, 581–594, June 1990.

2. G Archambault, M Fortin, D E Gill, M Aubertin, and B Ladanyi, "Experimental investigations for an algorithm simulating the effect of variable normal stiffness on discontinuities shear strength", (Edited by N Barton and O Stephansson), *Proc. Int. Symp. Rock Joints*, Loen, Norway, 141–148, June 1990.

3. S Bandis, A C Lumsden, and N R Barton, "Experimental studies of scale effects on the shear behaviour of rock joints", *Int. J. Rock Mech. Min. Sci. Geomech. Abstr.*, **18**, 1–21, 1981.

4. N R Barton, "Deformation phenomena in jointed rock", *Geotechnique*, **36** (2), 147–167, 1986.

5. N R Barton, "The shear strength of rock and rock joints", *Int. J. Rock Mech. Min. Sci. Geomech. Abstr.*, **13**, 255–279, 1976.

6. N R Barton and S C Bandis, "Review of the predictive capabilities of JRC-JCS model in engineering practice", (Edited by N Barton and O Stephansson), *Proc. Int. Symp. Rock Joints*, Loen, Norway, 603–610, June 1990.

7. N R Barton, A C Lumsden, and S C Bandis, "Fundamentals of rock joint behaviour", *Int. J. Rock Mech. Min. Sci. Geomech. Abstr.*, **20**, 249–268, 1983.

8. K J Bathe and A Chaudhary, "A solution method for planar and axisymmetric contact problems", *Int. J. Num. Meth. Eng.*, **21**, 65–88, 1985.

9. G Beer, "An isoparametric joint/interface element for finite element analysis", *Int. J. Num. Meth. Eng.*, **21**, 585–600, 1985.

10. G Beer, G N Pande, and J R Williams, Numerical Methods in Rock Mechanics, John Wiley and Sons, 40–46, 97–108, 122–131, 1990.

11. W W Bird and J B Martin, "Internal variable descriptions of rock joints with emphasis on solution algorithms", CERECAM Tech. Rep. 154, FRD/UCT Centre for Research in Computational and Applied Mechanics, University of Cape Town, Sept. 1991.

12. A Pinto Da Cunha, *Scale Effects*, Proc. First Int. Workshop of Scale Effects in Rock Masses, A.A. Balkema, Rotterdam, 1990.

13. C S Desai, M M Zaman, J G Lightner, and H J Siriwardane, "Thin-layer elements for interfaces and joints", *Int. J. Num. Anal. Meth. Geomech.*, **8**, 19–43, 1984.

14. A Francavilla and O C Zienkiewicz, "A note on numerical computation of elastic contact problems", *Int. J. Num. Meth. Eng.*, **9**, 913–924, 1975.

15. C Gerrard, "Shear failure of rock joints, appropriate constraints for empirical relations", *Int. J. Rock Mech. Min. Sci. Geomech*, **23**, 421–429, 1986.

16. J Ghaboussi, E L Wilson and J Isenberg, "Finite element for rock joints and interfaces", *J. Soil Mech. Found. Div. ASCE*, **99** (SM10), 833–848, 1973.

17. R E Goodman and J DuBois, "Duplication of dilatency in analysis of jointed rocks", *J. Soil Mech. Found. Div. ASCE*, **98** (SM4), 399–422, 1972.

18. R E Goodman, R L Taylor, and T L Brekke, "A model for the mechanics of jointed rock", *J. Soil Mech. Found. Div. ASCE*, **94** (SM3), 637–659, 1968.

19. D V Griffiths, "Numerical modelling of interfaces using conventional finite elements", In *Proc. of 5th Int. Conf. Num. Meth. Geomech.*, 837–844, 1985.

20. J M Handanyan, E R Danek, R A D'Andrea, and J D Sage, "The role of tension in failure of jointed rocks", (Edited by N Barton and O Stephansson), *Proc. Int. Symp. Rock Joints*, Loen, Norway, 195–202, June 1990.

21. L A Herrmann, "Finite element analysis of contact problems", *J. Eng. Mech. Div. ASCE.*, **104** (EM5), 1043–776, 1978.

22. F E Heuze and T G Barbour, "New models for rock joints and interfaces", *J. Geotec. Eng. Div. ASCE*, **108** (GT5), 757–776, 1982.

23. Hibbitt, Karlsson and Sorensen, Inc., *ABAQUS User's Manual*, Providence R.I., USA, 1990.

24. R W Hutson and C W Dowding, "Joint asperity degradation during cyclic shear", *Int. J. Rock Mech. Min. Sci. Geomech*, **27**, 109–119, 1990

25. M G Katona, "A simple contact-friction element with applications to buried culverts", *Int. J. Num. Anal. Meth. Geomech.*, **7**, 371–384, 1983.

26. N Kikuchi and Y J Song, "Penalty/finite element approximations of a class of unilateral problems in linear elasticity", *Quarterly of Applied Mathematics*, **39**, 1–22, 1981.

27. H K Kutter and G Weissbach, "Der Einflußvon Verformungs- und Belastingsgeschichte auf den Scherwiderstand von Gesteinskluften Unter Besonderer Berucksichtigung der Mylonitbildung", *Final report of DFG research project Ku 361/2/4*, 1980.

28. J B Martin and B D Reddy, "Variational principles and solution algorithms for internal variable formulations in plasticity", *Omaggio a Giulio Ceradini*, Universita di Roma "La Sapienza" (Roma), 465–478, 1988.

29. J B Martin, B D Reddy, T B Griffin and W W Bird, "Applications of mathematical programming concepts to incremental elastic-plastic analysis", *Engineering Structures*, **9**, 171–176, 1987.

30. J T Oden and N Kikuchi, "Finite element methods for constrained problems in elasticity", *Int. J. Num. Meth. Eng.*, **18**, 701–725, 1982.

31. Y Ohnishi and P G R Dharmaratne, "Shear behaviour of physical models of rock joints under constant normal stiffness conditions", (Edited by N Barton and O Stephansson), *Proc. Int. Symp. Rock Joints*, Loen, Norway, 267–274, June 1990.

32. G N Pande and K G Sharma, "On joint/interface elements and associated problems of numerical ill-conditioning", *Int. J. Num. Anal. Meth. Geomech.*, **3**, 293–300, 1979.

33. G N Pande and W Xiong, "An improved multilaminate model of jointed rock masses", (Edited by G N Pande R Dungar and J A Studer), *Numerical Models in Geomechanics*, 218–226, A A Balkema, Rotterdam, 1982.

34. F D Patton, "Multiple modes of shear failure in rock", In *Proc. of 1st Congress ISRM, Lisbon*, 509–513, **1**, 1966.

35. M E Plesha, "Constitutive model and finite element procedure for dilatent contact problems", *J. Eng. Mech. Div. ASCE.*, **115** (12), 2649–2668, 1989.

36. M E Plesha, "Constitutive models for rock discontinuities with dilatency and surface degradation", *Int. J. Num. Anal. Meth. Geomech.*, **11** (4), 345–362, 1987.

37. W E Roberts and H H Einstein, "Comprehensive model for rock discontinuities", *J. Geotec. Eng. Div.*, **104** (GT5), 553–569, 1978.

38. T Rodic and D R J Owen, "A plasticity theory of friction and joint elements", (Edited by E Hinton, D R J Owen and E Onate), *Computational Plasticity*, 1043–1062, 1989.

39. C A Skinas, S C Bandis, and C A Demiris, "Experimental investigations and modelling of rock joint behaviour under constant stiffness", (Edited by N Barton and O Stephansson), *Proc. Int. Symp. Rock Joints*, Loen, Norway, 301–308, June 1990.

40. H van Langen and P A Vermeer, "Interface elements for singular plasticity points", *Int. J. Num. Anal. Meth. Geomech.*, **15**, 301–315, 1991.

41. E L Wilson, "Finite elements for foundations, joints and fluids", (Edited by G Gudehus), Finite Elements in Geomechanics, John Wiley and Sons, 319–375, 1977.

42. P Wriggers, J C Simo, and R L Taylor, "Penalty and augmented Lagrangian formulations for contact problems", In *Proc. NUMETA '85 Conf., Swansea*, 97–106, 1985.

43. S Xu and M H de Freitas, "Kinematic mechanisms of shear deformation and the validity of Barton's shear models", (Edited by N Barton and O Stephansson), *Proc. Int. Symp. Rock Joints*, Loen, Norway, 767–774, June 1990.

44. J S Kuijpers and J A L Napier, "The effect of loading history on stress generation due to inelastic deformations around deep-level tabular stopes" *J. South African Inst. Min. Metall.*, **91**, 183–194, 1991.