On friction in forming : an experimental-numerical method to quantify contact behaviour

ON FRICTION IN FORMING

An experimental-numerical method
to quantify contact behaviour

Frans Starmans

ON FRICTION IN FORMING

An experimental-numerical method
to quantify contact behaviour

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Technische Universiteit Eindhoven,
op gezag van de rector magnificus, prof.ir. M.Tels,
voor een commissie aangewezen door het college van decanen
in het openbaar te verdedigen
op dinsdag 17 januari 1989 om 14.00 uur

door

FRANCISCUS JOZEF MARIE STARMANS

geboren te Schaesberg

Dit proefschrift is goedgekeurd door de promotoren

prof.dr.ir. J.D.Janssen
prof.Dipl.-Ing. J.A.G.Kals

adviseur

ir. W.A.M.Brekelmans

voor mijn ouders

Contents

Summary

In many forming processes contact phenomena play an essential role with respect to needed process forces, induced mechanical stresses in tools and workpiece, resulting workpiece geometry and surface condition. For a dimensioning of tools and choice of lubricants in the design stage, it is of importance to have the disposal of models which describe the mechanical interaction with enough accuracy. Well-known phenomenological models for this interaction are the friction model of Coulomb, the Von Mises model and the elasto-plastic analogies developed the last decade. This thesis considers a number of tools to evaluate the applicability of such contact models and to quantify contact parameters appearing in these models.

After an introduction and a description of the field of investigation in chapter 1, a thermodynamical continuum framework for the contact behaviour in forming processes is presented. The above mentioned mechanical contact models fit within this framework under quite general conditions. Therefore experiments are needed to specify useful contact models in more detail.

Chapter 3 considers possible experimental set-ups and measuring methods, resulting in the choice to investigate *in situ* the upsetting process of axisymmetric workpieces. Chapter 4 presents an experimental-numerical method to analyse forming processes taking measured boundary conditions into account, besides the influence of the measuring accuracy upon the calculated results is studied. Using this method the upsetting process with different sets of boundary conditions is analysed. It turns out that measurement of the normal stress and, if possible, the shear stress in the contact area between tool and workpiece is necessary to obtain reliable values for contact quantities suitable to evaluate the contact behaviour. The accuracy of the contact quantities can be increased by taking also the displacements of the unloaded surface into account, provided these are measured accurately enough.

In chapter 5 a contactless photogrammetric displacement measuring method is presented and evaluated. Experimental research learns that the theoretically predicted accuracy is not reached, due to several deficiencies in the modelling. The influence of some of these can be reduced by an improved set-up. When severe surface roughening occurs it will be doubtful whether the needed accuracy can be established. Measurement of the displacements remains of importance for determining the dimensions of the contact area, which is an input value for the method to quantify the contact stresses according to chapter 6. That chapter describes the design and analysis of a contact stress measuring tool for upsetting

experiments. Measurement of the elastic strains within the tool combined with a numerical model, delivers the values of the contact stresses between tool and workpiece. The strains are measured with thin film strain transducers. A model is obtained for such transducers. The experimental evaluation of this model is started and an experiment is indicated to evaluate the predicted behaviour of the measuring tool.

Finally chapter 7 presents a number of conclusions with respect to the present study and recommendations for a continuation.

Notation

Quantities

a	scalar
$\vec{a}$	vector
$\boldsymbol{a}, {}^3\boldsymbol{a}, {}^4\boldsymbol{a}$	second, third, fourth order tensor
$\underset{\sim}{a}$	column
$\underline{a}$	matrix

Operations and functions

$\underset{\sim}{a}^T, \underline{a}^T$	transposition
$\boldsymbol{a}^{-1}, \underline{a}^{-1}$	inversion
$\boldsymbol{a}^c$	conjugation
${}^3\boldsymbol{a}^{rc}, {}^4\boldsymbol{a}^{lc}, {}^4\boldsymbol{a}^c$	right, left, central conjugation
$\boldsymbol{a}^d$	deviatoric part
$\dot{a}$	material time derivative
$a^{\bullet}$	surface time derivative
$\overset{\triangledown}{a}$	objective time derivative
$\vec{a}\vec{b}, \boldsymbol{a}\vec{b}, \boldsymbol{a}\boldsymbol{b}$	dyadic product
$\vec{a}.\vec{b}, \boldsymbol{a}.\vec{b}, \boldsymbol{a}.\boldsymbol{b}$	inner (dot) product
$\vec{a}\times\vec{b}$	outer (cross) product
$\boldsymbol{a}:\boldsymbol{b}$	double inner (double dot) product
$\mathrm{tr}(\underline{a}), \mathrm{tr}(\boldsymbol{a})$	trace
$\det(\underline{a}), \det(\boldsymbol{a})$	determinant
$\vec{\nabla}$	gradient operator

$E(a)$	expected value
$p(a)$	probability density function
$p(a:b)$	likelihood function
$\exp(a)$	exponential function
$^{10}\log(a)$	logarithm to the base 10
$\min(a,b)$	minimum value
$\lvert a \rvert$	absolute value
$\lVert \vec{a} \rVert$	magnitude

Miscellaneous

a_0	quantity belonging to the initial or reference state
$\hat{a}$	invariant quantity or estimate
$<a>_A$	surface average
$\underline{0}$, $\underline{\underline{0}}$, $\mathbf{0}$	zero column, zero matrix, zero tensor
$\underline{I}$, I, ^4I	unit matrix, unit second order, unit fourth order tensor

1 Introduction

This thesis presents an investigation of the mechanical contact behaviour, occurring between die and workpiece in forming processes, in particular the contact in an upsetting configuration. The research is a continuation of the work by Baaijens (1987), which dealt with the numerical simulation of mechanical contact problems. For such simulations constitutive equations for the contact interaction are needed. In this thesis constitutive principles for the contact stresses are considered and an experimental method to quantify the constitutive equations is developed. Although a restricted class of contact problems is focussed on here, extension of the present work to other cases is possible.

In the past, for calculations on forming processes, mainly the easy to use friction models called after Coulomb and Von Mises were applied to describe the mechanical interaction between deforming material and die or punch. These models cover experimental results to a certain extent only. However, more sophisticated models were hardly significant, because of the absence of appropriate numerical techniques to insert them on the one hand, and a lack of measuring techniques to quantify model parameters on the other hand. Advanced numerical tools for complex contact phenomena are available nowadays, referred is to developments by, among others, Fredriksson (1976), Klarbring (1985) and Baaijens (1987). Together with the numerical tools, extended friction models have been proposed, thus increasing the need for techniques to assess the applicability of such models and to quantify them. After theoretical considerations with respect to friction models, an experimental–numerical technique is described here. Its potentiality is partly investigated and results attained thus far are reported. Before further exposing this work and its purpose and scope in paragraph 1.2, the class of contact problems *a priori* aimed at is described in the next paragraph.

1.1 The restricted class of contact problems

The class of contact problems investigated is defined by a set of restrictions, mentioned pointwise in the sequel, to be characterized as quasi–static mild cold forming conditions.
- Quasi–static and isothermal processes at room temperature. Deformation velocities and contact slip velocities are small. Inertial effects and velocity dependencies are negligible. Absence of high flash temperatures and large

temperature changes due to dissipative processes or external sources.

— Moderate relative displacements between the contact partners. Slip distances in the order of the dimensions of these partners are considered. Effects at large run distances are not dealt with, thus focussing on the behaviour during running in.

— No wear or fretting. Wear particles or conditions of grave material transference between the contact partners are avoided, resulting in only two body contact situations.

— Moderate forming contact pressures. Only one of the contact partners shows large, irreversible deformations.

— Non–agressive physical and chemical surrounding conditions.

— Initial homogeneity at the contact surface. The conditions of lubrication and roughness are equal at all contact points before loading is applied. Dynamic lubrication is irrelevant.

1.2 Purpose and scope of the present research

In many forming processes contact occurs between different physical bodies, interacting mechanically and thermally. These interactions partly define the process energy needed and the resulting state of the product, with respect to shape, surface condition and internal stresses. Important factors influencing the interactions are for instance artificial surface layers, lubricant, initial condition and preparation of the surfaces, temperature, chemical and physical condition of the surroundings such as air composition and humidity. Experimental results may easily be misinterpreted because of the amount of possible interfering factors. Furthermore the success of combining results, obtained under strongly differing conditions, by one particular model for some chosen material combination, is very doubtful. For that reason the restrictions, as given in the previous paragraph, are presupposed. For that very restricted class of contact situations it might be expected that they could be described by one phenomenological model but for the time being this is not the case. As stated in the previous paragraph, the Coulomb and Von Mises model nevertheless are, despite known deficiencies, commonly used.

In the last decade mechanical contact models, analogous to models for elasto–plastic material behaviour, have been formulated by different authors. Those models can describe both the Coulomb and the Von Mises behaviour. They deserve a nearer investigation to examine the applicability for the class of contact problems indicated. Before the methods for this purpose are described, some explanation

about the general strategy is needed.

In tribology research different classifications can be made. Here only the distinction into research on a local and on a global level is made. Investigations from a local point of view concentrate on microscopic phenomena such as contact at microscopic or even molecular small contact spots, in order to define micromechanical contact models. A disadvantage of applying such models for calculations of forming processes is the need to translate them, introducing continuum quantities, however insight is gained in the actual contact interactions. Research on a global scale directly aims at models formulated in continuum quantities. Such models are applicable for calculations, without translation. This global point of view is also taken in this thesis.

A theoretical requirement for continuum contact models is their thermodynamical validity. This is discussed in chapter 2 where continuum contact quantities, balance laws, dependent and independent variables and constitutive equations for the dependent variables are considered. It is shown that the mentioned elasto–plastic friction laws are consistent with the thermodynamic theory described. In order to quantify contact parameters and functions of the elasto–plastic friction model, but of other friction models as well, experiments are essential. Chapter 3 reviews possible experimental set-ups and techniques. The upsetting test is selected for further investigation. A hybrid experimental-computational procedure, which enables the determination of all contact quantities of interest, is described in chapter 4. This procedure is based on a numerical method for simulating the forming process using measured boundary data, taking the accuracy of these data into account. Chapter 5 presents a photogrammetric method to obtain the current positions of material contour points of the workpiece, thus enabling *in situ* measurement. This method is evaluated by upsetting experiments. As the knowledge of contour displacements only is not sufficient for reliable results for the contact quantities, chapter 6 proposes a technique to measure the contact stresses. In fact these stresses could be calculated from the elastic strains of the upsetting punch, *in situ* measured by thin film strain transducers. Use of this stress data, and possibly the contour data, in the simulation of the forming process should offer a useful method to quantify the contact behaviour. Future research is necessary to verify this statement. Finally chapter 7 concludes with comments on the strategy of the present research and the results obtained.

2 Constitutive equations for contact behaviour

For the theoretical validation of constitutive equations for contact behaviour a thermodynamical continuum theory for contact regions is used. Not only the formulation of balance laws is needed but also a consideration of independent variables and constitutive principles. In literature these aspects are not often treated simultaneously, see e.g. the elaborate work of Zmitrowicz (1987).

In this chapter the thermodynamical validity of the Maxwell friction models as developed by Fredriksson (1976), Petersson (1977), Michalowski & Mróz (1978), Cheng & Kikuchi (1985), Baaijens et al. (1986) is shown. For this purpose at first the modelling of contact regions within continuum theory is discussed and balance laws are formulated. The complexity of these laws is reduced by the introduction of some simplifying assumptions and by the specialization to the class of contact processes, restricted to in chapter 1. Distinction between independent and dependent variables and application of some constitutive principles lead to a set of constitutive contact equations. In order to derive the Maxwell friction laws internal variables and evolution laws for these variables are introduced in such a way that thermo-dynamical requirements are obeyed. Experiments are needed to investigate their applicability in practice, an aspect the rest of this thesis is focussed on.

2.1 The description of contact regions in continuum theory

In literature the description of real contact phenomena by means of two smooth body surfaces, possibly via an intermediate surface or layer, is mostly introduced more or less intuitively (Oden & Martins 1985, Ruina 1985). In this paragraph the concept of averaging volumes is used to support this intuitive process mathematically. So better understanding is obtained with respect to the validity of the particular modelling of contact regions.

Considered are two contacting bodies, see figure 2.1.1, mechanically and thermally loaded. Different phenomena occur internally, between the bodies and the surroundings and the bodies mutual. It is assumed that the insides of the bodies may be modelled as continua. In the transition region between the bodies capricious contact results from the roughness of both bodies. In this region chemical and physical different components are present such as oxide layers, lubricant and wear particles (Rabinowicz 1965). The resulting macroscopic behaviour, due to the interaction on a microscopic scale, can be ascribed to a fictitious smooth contact

layer between the two bodies. Phenomena such as slip, contact stiffness and frictional heat can be described using such a layer (Oden & Martins 1985, Zmitrowicz 1987).
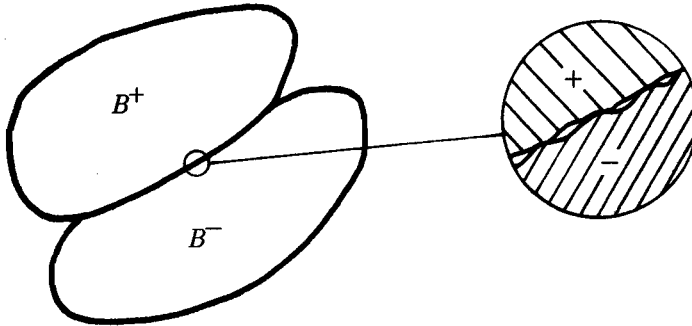
Figure 2.1.1 Two contacting bodies B^+ and B^-

It is noticed that the roughness of the bodies is smoothened but the influence of it will be represented in the layer's behaviour. A comparable statement can be made for lubricational effects.

Although studying the microscopic behaviour is not necessary, definition of the contact layer and the ascription of continuum quantities to it, starts at this scale. Instantaneously a smooth non–unique midplane S, situated in the transition region between the bodies, is appointed and elementary averaging volumes V_E are introduced (Whitaker 1969, Huyghe 1986). The smoothness is defined with respect to local i.e. microscopic fluctuations. This will be quantified in the sequel.
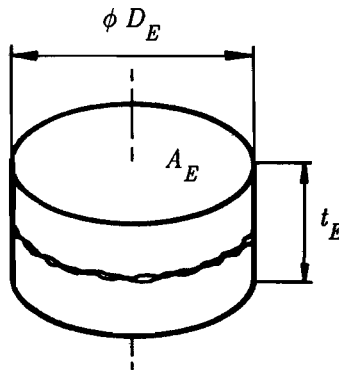
Figure 2.1.2 Elementary averaging volume with a part of the contact region

The points of the midplane S are uniquely identified by two locational parameters η_1 and η_2, composing a column $\underset{\sim}{\eta}$. For every point of S, denoted by $\underset{\sim}{\eta}$, a cylindrical averaging volume V_E with diameter D_E and thickness t_E is defined, see figure 2.1.2. This volume is oriented such that the centre of V_E coincides with the point $\underset{\sim}{\eta}$ of S and the axis of V_E is perpendicular to S at that point. The contact layer is defined as the inside of the envelope of the averaging volumes for all points $\underset{\sim}{\eta}$. The curvature of S inside of V_E is neglected, demanding that D_E is much smaller than a characteristic radius of curvature of S. The thickness of the contact layer equals t_E by approximation this way.

Considered is a quantity φ which is defined for each material component in the transition region as a function of the position (and time). The surface average $<\varphi>_A$ is introduced by

$$<\varphi>_A = \frac{1}{A_E} \int_{V_E} \varphi dV \quad , \quad A_E = \frac{\pi}{4} D_E^2 \tag{2.1.1}$$

and is defined at each point $\underset{\sim}{\eta}$. At some time t, $<\varphi>_A$ can be considered as a function of $\underset{\sim}{\eta}$, D_E and t_E. To quantify the requirements for D_E and t_E, the characteristic behaviour of $<\varphi>_A$ for some $\underset{\sim}{\eta}$ and t, as a function of D_E and t_E, is visualized in figure 2.1.3.
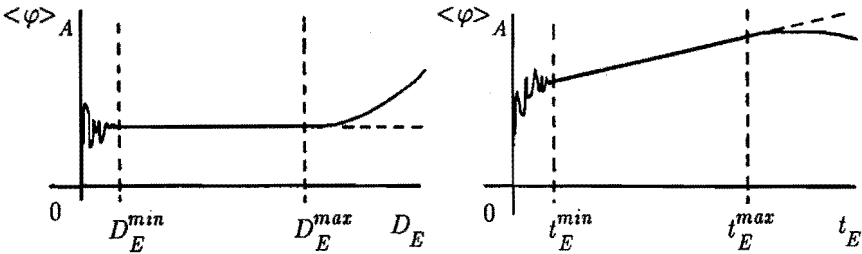
Figure 2.1.3 The averages as functions of the diameter and thickness of the
averaging volume

Strong fluctuations occur below some D_E^{min} and t_E^{min}. These fluctuations result from the roughness of both bodies and the distribution of components with strongly differing values of φ through the transition region. In that lower range the averages are useless as continuum quantities. Above some D_E^{max} and t_E^{max} the averages are

2 Constitutive equations for contact behaviour

smooth but diverge from the values, which would occur for homogeneous thermomechanical loading conditions. This results from variation of φ of one or more components influenced by global conditions. In order not to filter out this global behaviour also the upper range for the averaging volume dimensions is not useful. The averages based on volumes V_E with dimensions in the range between D_E^{min}, t_E^{min} and D_E^{max}, t_E^{max} can be used as local continuum quantities. If such a range cannot be indicated a continuum approach is not possible. For engineering contact problems however generally such a range can be appointed. Suitable choices follow from

$$d << D_E \quad , \quad t_E << L \tag{2.1.2}$$

where d is a characteristic distance for fluctuations in the local contact areas and L a characteristic distance for global variations.

Before deriving balance laws for the contact layer in the next paragraph, some kinematic relations are treated first. In figure 2.1.4 a cross section of a part of the contact layer is sketched. The points momentarily coinciding with the boundary surfaces S^+ and S^- are identified uniquely by surface coordinates ζ^+ and ζ^-.
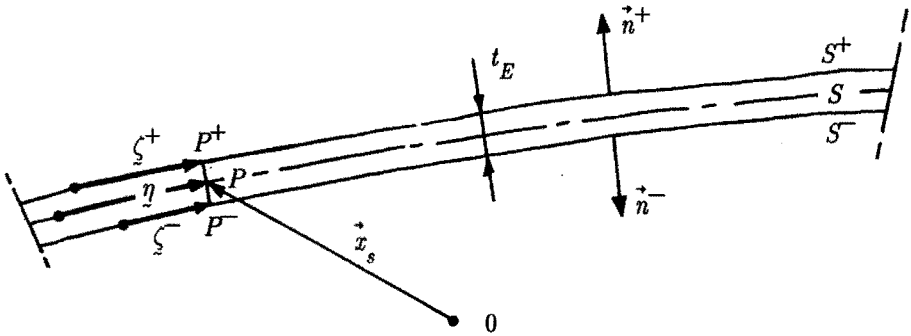
Figure 2.1.4 The contact layer with definitions of quantities

A point P of S with coordinates η is associated with the points P^+ and P^- of S^+ and S^- respectively. These points are located on a line perpendicular to S at P. In this way coordinates ζ^+ and ζ^- are instantaneously related to each η. The position in space of point η of S at time t is indicated by the vector $\vec{x}_s = \vec{x}_s(\eta, t)$ or equivalently

$\vec{x}_s = \vec{x}_s(\zeta^+,t)$ or $\vec{x}_s = \vec{x}_s(\zeta^-,t)$. In the sequel only choices for the coordinates η, ζ^+ and ζ^- are considered, such that $\vec{x}_s$ is continuous in η and t, ζ^+ and t respectively ζ^- and t. The natural and reciprocal surface base vectors $\vec{c}_s$ and $\vec{\gamma}_s$ follow from

$$\vec{c}_s = \frac{\partial \vec{x}_s}{\partial \eta} \quad , \quad \vec{n}_s = \frac{(\vec{c}_s)_1 \times (\vec{c}_s)_2}{\|(\vec{c}_s)_1 \times (\vec{c}_s)_2\|} \quad , \quad \vec{c}_s \cdot \vec{\gamma}_s^T = I \quad , \quad \vec{\gamma}_s \cdot \vec{n}_s = 0 \tag{2.1.3}$$

Since S is a smooth surface the unit normal vector $\vec{n}_s$ to S is a continuous and differentiable function of η. The surface gradient operator $\vec{\nabla}_s$ is defined by

$$\vec{\nabla}_s = \vec{\gamma}_s^T \frac{\partial}{\partial \eta} \tag{2.1.4}$$

The surface velocity $\vec{u}_s$ is the velocity of point η through space

$$\vec{u}_s = \vec{x}_s^{\bullet} \tag{2.1.5}$$

where the dot $\bullet$ denotes the time derivative operator $\left.\frac{\partial}{\partial t}\right|_\eta$.

2.2 General balance law and specific formulations

In this paragraph the balance laws for mass, momentum, moment of momentum, energy and entropy are considered for the class of contacts as defined in chapter 1. These laws are special cases of the general balance law for some (mass associated) quantity φ. For notational convenience φ is supposed to be a scalar quantity. To derive this general formulation for φ, the quantity Q, defined for the averaging volume V_E belonging to some point η, is considered

$$Q = \int\limits_{V_E} \rho\varphi dV = A_E <\rho\varphi>_A \tag{2.2.1}$$

where ρ is the mass density. The quantity Q may be identified physically as e.g. the total mass of the material within V_E, based on the choice $\varphi = 1$. In that case the

resulting balance law will be the law of mass conservation. Simplifications, valid for the restricted class of contacts are indicated in the sequel. Essential assumptions, to simplify the balance laws further, are discussed in the next paragraph.

The time derivative $Q^{\bullet}$ of Q is composed of contributions, caused by material flow (convection), immaterial flow (flux) and sources, subsequently to be reviewed.

1) Material flow through the border ∂V_E of V_E.

   The local material velocity is indicated by $\vec{v}$. The decrease of Q per unit of time due to convection then equals

$$\int_{\partial V_E} \rho\varphi(\vec{v}-\vec{u}_s).\vec{n}dA \tag{2.2.2}$$

   where $\vec{n}$ denotes the unit outward normal vector. For the whole boundary ∂V_E a uniform velocity $\vec{u}_s$ is assumed. Formally this is incorrect in case the orientation of V_E in space changes in time. Because of (2.1.2) these rotational effects may be neglected.

2) Flux through the boundary ∂V_E.

   The local flux is indicated by $\vec{\psi}$. The decrease of Q resulting from fluxes is

$$\int_{\partial V_E} \vec{n}.\vec{\psi}dA \tag{2.2.3}$$

3) Sources within V_E.

   Only the case of a distributed mass associated φ–production Φ is considered. The increase of Q equals

$$\int_{V_E} \rho\Phi dV \tag{2.2.4}$$

Combination and elaboration of these three contributions result in the following general balance law, see appendix 2.2.1

$$<\rho\varphi>_A^{\bullet}+(\vec{\nabla}_s.\vec{u}_s)<\rho\varphi>_A+\vec{\nabla}_s.[(I-\vec{n}_s\vec{n}_s).\left((<\vec{v}\rho\varphi>_A-\vec{u}_s<\rho\varphi>_A)+<\vec{\psi}>_A\right)]-<\rho\Phi>_A=$$

$$= -\vec{n}^+.(\dot{\vec{\psi}}^+ + (\vec{v}^+ - \vec{u}_s)\rho^+\varphi^+) - \vec{n}^-.(\dot{\vec{\psi}}^- + (\vec{v}^- - \vec{u}_s)\rho^-\varphi^-) \qquad (2.2.5)$$

The superscripts $+$ and $-$ indicate bulk continuum quantities at the $+$ and $-$ side of the contact layer. Equation (2.2.5) is only valid if the conditions (2.1.2) for the dimensions of the averaging volume are satisfied.

Some of the terms in (2.2.5) may be neglected or reformulated for the restricted class of contacts. Firstly all sources may be neglected. The mechanical sources are negligibly small whereas the thermal ones are not present. Secondly all fluxes in the plane of the layer are omitted. These inplane fluxes are dominated by the fluxes perpendicular to the layer. Finally the exclusion of severe wear induces only small mass flow from the bulk surroundings into the contact layer. The corresponding convection terms vanish from (2.2.5) by adopting S^+ and S^- as material bulk surfaces, with material coordinates ζ^+ and ζ^-, and surface averages, defined with a changing thickness t_E, locally defined as the distance between the two material bulk surfaces. It shows easily that (2.2.5) then becomes

$$<\rho\varphi>\!\!\dot{}_A + (\vec{\nabla}_s.\vec{u}_s)<\rho\varphi>_A + \vec{\nabla}_s.[(I - \vec{n}_s\vec{n}_s).\big((<\vec{v}\rho\varphi>_A - \vec{u}_s<\rho\varphi>_A) + <\vec{\psi}>_A\big)] - <\rho\Phi>_A =$$

$$= -\vec{n}^+.\dot{\vec{\psi}}^+ - \vec{n}^-.\dot{\vec{\psi}}^- \qquad (2.2.6)$$

while the surface averages should be interpreted as averaged quantities using a variable thickness. The change of the unit normal vectors $\vec{n}^+$ and $\vec{n}^-$ is a second order effect and therefore these vectors may be taken perpendicular to S.

Special choices for φ and $\vec{\psi}$ in (2.2.6) lead to the specific balance laws for the mass, momentum, moment of momentum, energy or entropy of V_E.

Mass: with substitution of $\varphi = 1$, the quantity Q represents mass

$$<\rho>\!\!\dot{}_A + (\vec{\nabla}_s.\vec{u}_s)<\rho>_A + \vec{\nabla}_s.[(I - \vec{n}_s\vec{n}_s).(<\vec{v}\rho>_A - \vec{u}_s<\rho>_A)] = 0 \qquad (2.2.7)$$

Momentum: φ is replaced by the material velocity $\vec{v}$, and $\vec{\psi}$ by the opposite of the conjugate σ^c of the Cauchy stress tensor σ. Only non–polar media are considered so $\sigma^c = \sigma$. Neglecting the momentum, being only interested in quasi–static processes,

the static equilibrium equation results in

$$\vec{n}^+.\sigma^+ + \vec{n}^-.\sigma^- = \vec{0} \tag{2.2.8}$$

This equation represents the law of action and reaction between the two bodies.

Moment of momentum: this yields for non–polar media the static moment equilibrium equation

$$\vec{x}^+ \times (\sigma^+.\vec{n}^+) + \vec{x}^- \times (\sigma^-.\vec{n}^-) = \vec{0} \tag{2.2.9}$$

Use of (2.2.8) in (2.2.9) offers a not trivially fulfilled equation. Because $t_E \ll L$ this remaining equation describes a second order effect. As the first order effect is automatically satisfied by the static equilibrium equation, the static moment equation is not considered further.

Energy: the relevant material coupled energy is the internal energy e, the kinetic energy is negligible for quasi–static processes. The flux consists of a mechanical part $-\vec{v}.\sigma$ and a thermal part $\vec{q}$, equal to the heat flux in the case considered. A balance equation for the internal energy is obtained

$$<\rho e>_A^{\bullet} + (\vec{\nabla}_s.\vec{u}_s)<\rho e>_A + \vec{\nabla}_s.[(I - \vec{n}_s \vec{n}_s).(<\rho \vec{v} e>_A - \vec{u}_s <\rho e>_A)] =$$

$$= -\vec{n}^+.(\vec{q}^+ - \vec{v}^+.\sigma^+) - \vec{n}^-.(\vec{q}^- - \vec{v}^-.\sigma^-) \tag{2.2.10}$$

The bulk of the bodies exchanges heat with the contact layer, but also does work on it, which is used for deformation in thickness and shear direction. The latter results in macroscopic slip between the bodies, the work needed for this will be dissipated almost completely as frictional heat in the contact layer.

Entropy: with the general balance law formulation a Clausius–Duhem equation for the entropy production s is derived

$$<\rho s>_A = <\rho h>_A^{\bullet} + (\vec{\nabla}_s.\vec{u}_s)<\rho h>_A +$$

$$+ \vec{\nabla}_s.[(I - \vec{n}_s \vec{n}_s).(<\vec{v} \rho h>_A - \vec{u}_s <\rho h>_A)] + \vec{n}^+.\vec{w}^+ + \vec{n}^-.\vec{w}^- \tag{2.2.11}$$

where h is the entropy and $\vec{w}$ the entropy flux. The second law of thermodynamics postulates $<\rho s>_A = 0$ for all allowable reversible contact processes and $<\rho s>_A > 0$ for all thermodynamical allowable irreversible contact processes. Thermodynamical allowable indicates that the processes do not violate other thermodynamic concepts such as e.g. the balance of internal energy.

2.3 Simplifying assumptions

For some quantities in the equations, as formulated in the previous paragraph, assumptions are introduced, to simplify the balance laws further.

First the flow terms are considered. Opposed to the flux terms in the previous paragraph, the flow terms in the plane of the contact cannot simply be neglected. Choosing the surface coordinates η such that $(I-\vec{n}_s\vec{n}_s).(<\vec{v}\rho>_A -\vec{u}_s<\rho>_A) = \vec{0}$ the term $\vec{\nabla}_s.[(I-\vec{n}_s\vec{n}_s).(<\vec{v}\rho>_A -\vec{u}_s<\rho>_A)]$ disappears from the balance of mass. The similar terms in the balances of energy and entropy vanish with the assumption $(I-\vec{n}_s\vec{n}_s).(<\vec{v}\rho e>_A -\vec{u}_s<\rho e>_A) \approx (I-\vec{n}_s\vec{n}_s).(<\vec{v}\rho h>_A -\vec{u}_s<\rho h>_A) \approx \vec{0}$.

A second assumption concerns the entropy fluxes $\vec{w}^+$ and $\vec{w}^-$. These fluxes are taken equal to $(\vec{q}/T)^+$ respectively $(\vec{q}/T)^-$, a well–known result from bulk thermodynamics (Müller 1985).

The remaining contact layer quantities $<\rho>_A$, $<\rho e>_A$, $<\rho h>_A$ and $<\rho s>_A$ in the balances of mass, internal energy and entropy provoke the definition of mass weighted contact layer quantities

$$\rho_s = <\rho>_A \quad , \quad e_s = \rho_s^{-1}<\rho e>_A \quad , \quad h_s = \rho_s^{-1}<\rho h>_A \quad , \quad s_s = \rho_s^{-1}<\rho s>_A \qquad (2.3.1)$$

i.e. the mass density, internal energy, entropy and entropy production of the contact layer respectively. With the (non–mass) weighted contact layer temperature $T_s = <T>_A$ the contact layer free energy $f_s = e_s - T_s h_s$ is defined. Then the balances of mass, momentum, energy and entropy result in

$$\rho_s^\bullet + \rho_s \vec{\nabla}_s.\vec{u}_s = 0 \qquad (2.3.2)$$

$$\sigma^+.\vec{n}^+ + \sigma^-.\vec{n}^- = \vec{0} \qquad (2.3.3)$$

$$\rho_s \overset{\bullet}{e_s} = -\vec{n}^+.(\vec{q}^+ - \vec{v}^+.\sigma^+) - \vec{n}^-.(\vec{q}^- - \vec{v}^-.\sigma^-) \tag{2.3.4}$$

$$-\rho_s(\overset{\bullet}{f_s} + T_s\overset{\bullet}{h_s}) + \vec{n}^+.(\vec{q}^+(\frac{T_s - T^+}{T^+}) + \vec{v}^+.\sigma^+) + \vec{n}^-.(\vec{q}^-(\frac{T_s - T^-}{T^-}) + \vec{v}^-.\sigma^-) \geq 0 \tag{2.3.5}$$

2.4 Independent variables and constitutive equations

In this paragraph constitutive quantities, constitutive relations for these quantities and the independent variables appearing in these relationships are considered. Constitutive principles induce restrictions with respect to the initially general constitutive relations (Hunter 1976, Müller 1985).

In continuum thermodynamics the temperature and position vector, as functions of material coordinates and time, are generally accepted as independent variables in constitutive equations. Characteristic for the independent variables is that, if they are known as functions of location and time, all other quantities of interest can be calculated. For the contact layer it is assumed that the following variables may be chosen as the set of independent variables, being functions of $\underset{\sim}{\eta}$ and t

$$\vec{x}_s, \vec{x}^+, \vec{x}^-, T_s, T^+, T^- \tag{2.4.1}$$

The contact layer mass density ρ_s is omitted as independent variable because it can be expressed in $\vec{x}_s$ using the mass balance.

In the remaining thermodynamical balances a set of quantities is present for which constitutive equations are needed. These dependent variables are

$$f_s, h_s, q_n^+, q_n^-, \vec{\sigma}^+, \vec{\sigma}^- \tag{2.4.2}$$

where the normal heat fluxes q_n^+, q_n^- are defined by $q_n^+ = \vec{n}^+.\vec{q}^+$ respectively $q_n^- = \vec{n}^-.\vec{q}^-$, and the stress vectors $\vec{\sigma}^+$, $\vec{\sigma}^-$ by $\vec{\sigma}^+ = \vec{n}^+.\sigma^+$ respectively $\vec{\sigma}^- = \vec{n}^-.\sigma^-$. The normal heat fluxes are related to each other by the energy balance, the stress vectors by the action and reaction law. Therefore a constitutive equation is needed only for $\vec{\sigma}^+$ and q_n^+, while $\vec{\sigma}^-$ and q_n^- are dictated by the mentioned balance laws.

To arrive at constitutive equations, the principles of determination, equipresence

and local action are applied. These result in the statement that each of the variables of (2.4.2) can be expressed in the history of the deformational quantities F_s, F^+, F^-, $\vec{\delta}_s$, $\vec{\delta}^+$ and $\vec{\delta}^-$, and the thermal quantities $\vec{g}_s$, $\vec{g}^+$, $\vec{g}^-$, T_s, T^+ and T^-. The deformation tensors F_s, F^+ and F^-, defined by

$$F_s = (\frac{\partial \vec{x}_s}{\partial \underline{\eta}})^T \vec{\gamma}_{s0} \quad , \quad F^+ = (\frac{\partial \vec{x}^+}{\partial \underline{\zeta}^+})^T \vec{\gamma}_0^+ \quad , \quad F^- = (\frac{\partial \vec{x}^-}{\partial \underline{\zeta}^-})^T \vec{\gamma}_0^- \tag{2.4.3}$$

take the non-material inplane deformation of the contact layer and the material deformation of the bulk surfaces into account. Here $\vec{\gamma}_{s0}$, $\vec{\gamma}_0^+$ and $\vec{\gamma}_0^-$ are the reciprocal vector bases in some reference state, belonging to the surfaces S, S^+ and S^- respectively. The deformation vectors $\vec{\delta}_s$, $\vec{\delta}^+$ and $\vec{\delta}^-$, defined by

$$\vec{\delta}_s = \vec{\delta}_{s0} + \int_{t_0}^t \left(\vec{v}^+(\underline{\eta},\tau) - \vec{v}^-(\underline{\eta},\tau)\right) d\tau \quad , \quad \vec{\delta}^+ = \vec{\delta}_0^+ + \int_{t_0}^t \left(\vec{v}^+(\underline{\zeta}^+,\tau) - \vec{v}^-(\underline{\zeta}^+,\tau)\right) d\tau \quad ,$$

$$\vec{\delta}^- = \vec{\delta}_0^- + \int_{t_0}^t \left(\vec{v}^+(\underline{\zeta}^-,\tau) - \vec{v}^-(\underline{\zeta}^-,\tau)\right) d\tau \tag{2.4.4}$$

account for thickness deformations and slip, which may be considered as extreme shear deformation of the contact layer. The time integrations start at time t_0 assigned to the reference state and end at the current time t. Finally the temperature gradients $\vec{g}_s$, $\vec{g}^+$ and $\vec{g}^-$ are defined by

$$\vec{g}_s = \vec{\nabla}_s T_s \quad , \quad \vec{g}^+ = \vec{\nabla}_s T^+ \quad , \quad \vec{g}^- = \vec{\nabla}_s T^- \tag{2.4.5}$$

History variables, also called hidden variables, are introduced to take the history of the independent variables into account. The time derivatives of these variables are given by evolution laws, in which only the current values of the independent variables and the hidden variables appear, based on the principle of equipresence. In analogy with the foregoing, history variables $\underline{H}_s(\underline{\eta},t)$ of the contact layer and history variables $\underline{H}^+(\underline{\zeta}^+,t)$, $\underline{H}^-(\underline{\zeta}^-,t)$ of the surfaces S^+ and S^- respectively, are introduced.

Resuming, each constitutive quantity $C(\eta,t)$ of (2.4.2) is assumed to be a function of the actual values of the independent variables (2.4.1), belonging to contact layer point η, augmented with the actual values of the history variables belonging to η (to which always a ζ^+ and ζ^- are associated, see paragraph 2.1)

$$C = C(F_s, F^+, F^-, \vec{\delta}_s, \vec{\delta}^+, \vec{\delta}^-, \vec{g}_s, \vec{g}^+, \vec{g}^-, T_s, T^+, T^-, \underset{\sim}{H}_s, \underset{\sim}{H}^+, \underset{\sim}{H}^-) \tag{2.4.6}$$

For the history variables' evolution laws, the following types of equations will apply

$$\overset{\bullet}{\underset{\sim}{H}}_s = \overset{\bullet}{\underset{\sim}{H}}_s(F_s, ..., \underset{\sim}{H}^-) \quad , \quad \overset{\bullet}{\underset{\sim}{H}}{}^+ = \left.\frac{\partial \underset{\sim}{H}^+}{\partial t}\right|_{\zeta^+} = \overset{\bullet}{\underset{\sim}{H}}{}^+(F_s, ..., \underset{\sim}{H}^-) \quad ,$$

$$\overset{\bullet}{\underset{\sim}{H}}{}^- = \left.\frac{\partial \underset{\sim}{H}^-}{\partial t}\right|_{\zeta^-} = \overset{\bullet}{\underset{\sim}{H}}{}^-(F_s, ..., \underset{\sim}{H}^-) \tag{2.4.7}$$

For an analysis of contact behaviour in the restricted class, it seems unnecessary to cope with all independent variables in (2.4.6) and (2.4.7). It is assumed that only independent variables of the contact layer together with T^+ and T^- are relevant to construct constitutive equations. It is also assumed that for the layer's coordinate system η, the coordinate system of the surface of the irreversible deforming body may be chosen without conflicting with paragraph 2.3. So the equations (2.4.6) and (2.4.7), choosing B^+ arbitrarily as the deforming body, simplify to

$$C(\zeta^+,t) = C(F_s, \vec{\delta}_s, \vec{g}_s, T_s, T^+, T^-, \underset{\sim}{H}_s) \quad , \quad C \in \{f_s, h_s, q_n^+, \vec{\sigma}^+\} \tag{2.4.8}$$

$$\overset{\bullet}{\underset{\sim}{H}}_s(\zeta^+,t) = \left.\frac{\partial \underset{\sim}{H}_s}{\partial t}\right|_{\zeta^+} = \overset{\bullet}{\underset{\sim}{H}}_s(F_s, \vec{\delta}_s, \vec{g}_s, T_s, T^+, T^-, \underset{\sim}{H}_s) \tag{2.4.9}$$

Before the implications of the above on the entropy inequality are studied in the next paragraph, a last constitutive principle is applied, i.e. the principle of objectivity. It states that constitutive equations should be invariant for observer transformations. Part of the principle is the assumption that mass density, stress vector, heat flux vector, internal energy, entropy, temperature and heat sources are objective quantities. Elaboration of the principle may be done by the introduction of

two different observers with the same notion of time and distance (see e.g. Müller 1985 or appendix 2.4.1). To obey objectivity for the constitutive equations, invariant or corotational quantities (Van Wijngaarden 1988) can be used. In appendix 2.4.1 a non-singular second order tensor A, only depending on the kinematics of S^+, is introduced to transform objective quantities into invariant quantities. Also the choice for A is considered there. It turns out to be convenient to introduce the invariant stress vector $\hat{\vec{\sigma}}^+$ and deformation vector $\hat{\vec{\delta}}_s$ by

$$\hat{\vec{\sigma}}^+ = \frac{\rho_{s0}}{\rho_s}A^{-1}.\vec{\sigma}^+ \quad , \quad \dot{\hat{\vec{\delta}}}_s = A^c.(\vec{v}^+ - \vec{v}^-) \quad , \quad \hat{\vec{\delta}}_s(t_0) = \vec{0} \tag{2.4.10}$$

Instead of F_s the invariant contact layer right Cauchy-Green tensor $C_s = F_s^c.F_s$ can be used in the constitutive equations, besides suitable invariant history variables $\underset{\sim}{\hat{H}}$ and $\hat{\vec{g}}_s = A^c.\vec{g}_s$. This results in

$$\hat{C} = \hat{C}(C_s, \hat{\vec{\delta}}_s, \hat{\vec{g}}_s, T_s, T^+, T^-, \underset{\sim}{\hat{H}}_s) \quad , \quad \hat{C} \in \{f_s, h_s, q_n^+, \hat{\vec{\sigma}}^+\} \tag{2.4.11}$$

$$\dot{\underset{\sim}{\hat{H}}}_s = \dot{\underset{\sim}{\hat{H}}}_s(C_s, \hat{\vec{\delta}}_s, \hat{\vec{g}}_s, T_s, T^+, T^-, \underset{\sim}{\hat{H}}_s) \tag{2.4.12}$$

2.5 The entropy inequality

In this paragraph the implications of the entropy inequality, using the results of the previous paragraph, are studied. With the invariant quantities it follows from equation (2.3.5)

$$-\rho_{s0}(\dot{f}_s + \dot{T}_s h_s) + \frac{\rho_{s0}}{\rho_s}(q_n^+ \frac{T_s - T^+}{T^+} + q_n^- \frac{T_s - T^-}{T^-}) + \hat{\vec{\sigma}}^+.\dot{\hat{\vec{\delta}}}_s \geq 0 \tag{2.5.1}$$

As usual it is assumed that a decomposition into the inequalities

$$-\rho_{s0}(\dot{f}_s + \dot{T}_s h_s) + \hat{\vec{\sigma}}^+.\dot{\hat{\vec{\delta}}}_s \geq 0 \tag{2.5.2}$$

$$q_n^+ \frac{T_s - T^+}{T^+} + q_n^- \frac{T_s - T^-}{T^-} \geq 0 \qquad (2.5.3)$$

is allowed. The second inequality restricts the possible heat fluxes perpendicular to the contact layer. Simple constitutive expressions for q_n^+ and q_n^- are given by

$$q_n^+ = \alpha^+ (T_s - T^+) \quad , \quad \alpha^+ > 0 \qquad (2.5.4)$$

$$q_n^- = \alpha^- (T_s - T^-) \quad , \quad \alpha^- > 0 \qquad (2.5.5)$$

Use of (2.4.11) and (2.4.12) in (2.5.2) leads to the restrictions

$$\frac{\partial f_s}{\partial C_s} = 0 \quad , \quad \frac{\partial f_s}{\partial \overset{\scriptstyle 3}{\delta}_s} = \overset{\scriptstyle 3}{\sigma}{}^+ \quad , \quad \frac{\partial f_s}{\partial \overset{\scriptstyle 3}{g}_s} = \vec{0} \quad , \quad \frac{\partial f_s}{\partial T_s} = -h_s \quad , \quad \left(\frac{\partial f_s}{\partial \underset{\sim}{\hat{H}}_s}\right)^T \otimes \dot{\underset{\sim}{\hat{H}}}_s \leq 0 \qquad (2.5.6)$$

where $\otimes$ is an operator determined by the character of the different constituents of $\underset{\sim}{\hat{H}}{}^+$. The result (2.5.6) states that f_s is independent of the inplane deformations of the contact layer and consequently that the contact stresses are not influenced by these deformations. This is felt as a lack in the presented theory. Deformations of the surface of the formed body cause roughness changes, affecting the stiffness of the contact layer and the frictional behaviour. In order to deal with such influences inplane contact layer stresses (a flux, neglected in paragraph 2.2) unequal zero are needed, a case which is not elaborated here. The result (2.5.6) for the contact stresses $\overset{\scriptstyle 3}{\sigma}{}^+$ is valid only if $\dot{\underset{\sim}{\hat{H}}}_s$ is independent of $\overset{\scriptstyle 3}{\dot{\delta}}_s$. This will be the case in the sequel. The history parameters are restricted further by the remaining inequality of (2.5.6).

2.6 Maxwell friction models

In this paragraph a special choice for the history variables leads to a constitutive equation for the contact stress vector resembling the constitutive equation for elasto–plastic material behaviour. This method was described for bulk material behaviour by Van Wijngaarden (1988). The resulting equations are analogous to the

equations given by e.g. Cheng & Kikuchi (1985) and Baaijens (1987).

The contact layer's free energy is given by

$$f_s = f_s(\vec{\delta}_s - \vec{\delta}_s^{ir}, T_s, H_t, H_n) \tag{2.6.1}$$

in which H_t and H_n are scalar history variables. The vector quantity $\vec{\delta}_s^{ir}$ is the irreversible contact layer deformation vector, a first order history variable. In (2.6.1) it is assumed that only the reversible part $\vec{\delta}_s^r = \vec{\delta}_s - \vec{\delta}_s^{ir}$ of $\vec{\delta}_s$ is of interest for the free energy. This assumption is similar to the elastic (reversible) and plastic (irreversible) deformation decomposition in bulk mechanics (Lee 1969, Sidoroff 1973). Note that in normal and in inplane (or slip) direction reversible and irreversible deformation of the contact layer is possible. Evolution laws for $\dot{\vec{\delta}}_s^{ir}$, $\dot{H}_t$ and $\dot{H}_n$ as function of $\vec{\delta}_s - \vec{\delta}_s^{ir}$, T_s, H_t and H_n are needed. For $\dot{\vec{\delta}}_s^{ir}$ it is assumed that a potential $\varphi = \varphi(\vec{\delta}_s - \vec{\delta}_s^{ir}, T_s, H_t, H_n)$ or equivalently a potential $\varphi = \varphi(\vec{\sigma}^+, T_s, H_t, H_n)$ exists such that

$$\dot{\vec{\delta}}_t^{ir} = \frac{\partial \varphi}{\partial \vec{\tau}} \quad , \quad \dot{\vec{\delta}}_n^{ir} = \frac{\partial \varphi}{\partial \vec{p}} \tag{2.6.2}$$

where the invariant tangential or frictional stress vector $\vec{\tau}$, normal or compressive stress vector $\vec{p}$, irreversible tangential displacement or slip vector $\dot{\vec{\delta}}_t^{ir}$ and irreversible normal displacement or compression vector $\dot{\vec{\delta}}_n^{ir}$ are defined by

$$\vec{\tau} = (I - \vec{n}_\sigma \vec{n}_\delta) . \vec{\sigma}^+ \quad , \quad \vec{p} = \vec{\sigma}^+ - \vec{\tau} \quad , \quad \dot{\vec{\delta}}_t^{ir} = (I - \vec{n}_\delta \vec{n}_\sigma) . \dot{\vec{\delta}}_s^{ir} \quad , \quad \dot{\vec{\delta}}_n^{ir} = \dot{\vec{\delta}}_s^{ir} - \dot{\vec{\delta}}_t^{ir} \quad ,$$

$$\vec{n}_\sigma = A^{-1} . \vec{n}^+ \quad , \quad \vec{n}_\delta = A^c . \vec{n}^+ \tag{2.6.3}$$

In an isotropic case (2.6.2) can be simplified to

$$\dot{\vec{\delta}}_t^{ir} = \frac{\vec{\tau}}{\tau}\frac{\partial\varphi}{\partial\tau} \quad , \quad \dot{\vec{\delta}}_n^{ir} = \frac{\vec{p}}{p}\frac{\partial\varphi}{\partial p} \quad , \quad \varphi = \varphi(\tau,p,T_s,H_t,H_n) \tag{2.6.4}$$

with $\tau = \|\vec{\tau}\|$ and $p = \|\vec{p}\|$. It is supposed that f_s is given by

$$f_s = \frac{1}{2}(\vec{\delta}_t^r.S_t.\vec{\delta}_t^r + \vec{\delta}_n^r.\vec{\delta}_n^r s_n) + \psi(T_s,H_t,H_n) \tag{2.6.5}$$

where S_t and s_n denote a constant stiffness tensor and scalar respectively. Combination of (2.5.6) and the above equations delivers a constitutive Maxwell relationship for the invariant tangential and normal stress vector

$$\frac{1}{\rho_{s0}}\dot{\vec{\tau}} + S_t.\frac{\vec{\tau}}{\tau}\frac{\partial\varphi}{\partial\tau} = S_t.\dot{\vec{\delta}}_t \tag{2.6.6}$$

$$\frac{1}{\rho_{s0}}\dot{\vec{p}} + s_n\frac{\vec{p}}{p}\frac{\partial\varphi}{\partial p} = s_n\dot{\vec{\delta}}_n \tag{2.6.7}$$

If A is a rotation tensor and $S_t = s_t I$ the above relationships can be formulated as

$$\frac{1}{\rho_{s0}}(\dot{\vec{\tau}} + A.(\dot{A}^{-1}).\vec{\tau} - \frac{\dot{\rho}_s}{\rho_s}\vec{\tau}) + s_t\frac{\vec{\tau}}{\tau}\frac{\partial\varphi}{\partial\tau} = (\frac{\rho_{s0}}{\rho_s})^{-1}s_t\dot{\vec{\delta}}_t \tag{2.6.8}$$

$$\frac{1}{\rho_{s0}}(\dot{\vec{p}} + A.(\dot{A}^{-1}).\vec{p} - \frac{\dot{\rho}_s}{\rho_s}\vec{p}) + s_n\frac{\vec{p}}{p}\frac{\partial\varphi}{\partial p} = (\frac{\rho_{s0}}{\rho_s})^{-1}s_n\dot{\vec{\delta}}_n \tag{2.6.9}$$

with $\vec{\tau} = (I - \vec{n}^+\vec{n}^+).\vec{\sigma}^+$, $\vec{p} = \vec{\sigma}^+ - \vec{\tau}$, $\dot{\vec{\delta}}_t = (I - \vec{n}^+\vec{n}^+).\dot{\vec{\delta}}_s$ and $\dot{\vec{\delta}}_n = \dot{\vec{\delta}}_s - \dot{\vec{\delta}}_t$. With the mass balance, $\dot{\rho}_s/\rho_s$ can be replaced by the surface magnification rate $\vec{\nabla}_s.\vec{v}^+$ while ρ_{s0}/ρ_s can be indentified as the surface magnification factor with respect to the reference state. The combinations $\dot{\vec{\tau}} + A.(\dot{A}^{-1}).\vec{\tau}$ and $\dot{\vec{p}} + A.(\dot{A}^{-1}).\vec{p}$ are objective rates of $\vec{\tau}$ and

$\vec{p}$ respectively. Identification with such rates in literature is possible, see appendix 2.6.1.

As for constitutive equations for metals, the history variables H_t and H_n introduce hardening (or softening). Irreversible work hardening is considered as an example. In that case H_t is identified as the irreversible frictional work and H_n as the irreversible compressive work

$$H_t = \int_{t_0}^{t} \vec{\tau} \cdot \dot{\vec{\delta}}_t^{ir} d\tau \quad , \quad H_n = \int_{t_0}^{t} \vec{p} \cdot \dot{\vec{\delta}}_n^{ir} d\tau \tag{2.6.10}$$

For their evolution laws it follows with (2.6.2)

$$\dot{H}_t = \tau \frac{\partial \varphi}{\partial \tau} \quad , \quad \dot{H}_n = p \frac{\partial \varphi}{\partial p} \tag{2.6.11}$$

The restriction of (2.5.6) results in

$$(\frac{\partial f_s}{\partial H_t} - 1)\tau \frac{\partial \varphi}{\partial \tau} + (\frac{\partial f_s}{\partial H_p} - 1)p \frac{\partial \varphi}{\partial p} \leq 0 \tag{2.6.12}$$

If the deformation of the contact layer in thickness direction is unimportant or negligible, the normal part of the deformation vector disappears as a variable in the constitutive equations. It may be replaced by the normal stress vector $\hat{\vec{p}}$ for which the constitutive equation then cancels. Consideration of the entropy inequality results in the statement that f_s has to be independent of $\hat{\vec{p}}$. The only influence of $\hat{\vec{p}}$ upon $\hat{\vec{\tau}}$ then remaining is through φ, which is not restricted in any way. Hence, reversible shear behaviour cannot be influenced by the normal stress and the same remarks as with respect to (2.5.6) in paragraph 2.5 can be made.

2.7 Concluding remarks

In this chapter contact behaviour in a continuum formulation is discussed. Mechanical constitutive contact equations are derived, which are consistent with the formulated thermodynamical theory. Coulomb and Von Mises like friction laws are not rejected, as they are special cases of these constitutive equations (Baaijens 1987). To obtain more severe restrictions for the constitutive equations, the different thermodynamical quantities have to be specified in more detail. For such a specification, not nearer investigated in this thesis, results from structural modelling of the real contacts (e.g. Francis 1977, Landheer et al. 1980, Tangena 1988) may be used. As for such an approach contact models on a microscopic scale are needed, also results from research on the microscopic physical behaviour of contacting surfaces are indispensable to obtain more restrictive statements.

The constitutive equations given are derived after a lot of simplifications and assumptions. Whether such equations can describe real contact behaviour should be evaluated by experiments. For that reason the rest of this thesis deals with experimental procedures.

3 Experimental set–up and quantification

In this chapter the experimental set–up and methods for quantifying contact models are considered. With respect to the type of experiment, the upsetting of cylindrical test specimen is judged to be useful for the class of contacts as described in chapter 1. An experimental–numerical method is proposed which enables the determination of the *in situ* value of all contact quantities needed.

In the first paragraph the choice for the upsetting process is motivated. Thereafter it is elucidated that, for the contact processes considered, determination of the *in situ* contact situation is desirable. Of course the measurement procedure may not influence the contact processes. A hybrid experimental-numerical method is presented accommodating these demands. It combines *in situ* measurements with a numerical simulation of the forming process in order to determine quantities which are not measurable *in situ*. Different aspects of this method are indicated and elaborated further in subsequent chapters. Finally some approximate calculations serve to dimension the set–up and to quantify process parameters, in order to fulfil the requirements of the class of contacts considered.

3.1 Choice of the experimental set–up

In the past many experimental methods are developed for the investigation of contact behaviour in forming. Examples are the upsetting of ring-shaped test pieces and the drawing of sheet material (Gräbener 1981). In these tests only global quantities, such as external loads, reaction forces, process velocities and dimension changes, are measured. Assumptions about the contact behaviour are used to translate these measured values into one contact characterizing parameter value i.c. the coefficient of friction of the Coulomb model or the interface friction factor of the Von Mises model. This translation is performed with calibration curves of an output quantity as function of a chosen input quantity, calculated for various parameter values. If experimental data fits with such a curve, it is decided that for the given case the assumed friction model applies, with a parameter value corresponding to the curve. If none of the curves is actually fitting, the friction model is rejected, or the parameter is supposed to be non-constant. Such conclusions are not satisfactory. In the first case the question arises whether there may be other frictional models resulting in the same curve. For the second situation, one would like to quantify the change of the parameter as a function of contact quantities. To overcome these

problems it is necessary to appoint (intuitively) the contact quantities of importance and to determine the local values. An extra complication is that only very few experimental methods are available to perform direct measurements in the contact regions occurring in forming processes. A short review of these methods is given in the next paragraph. It is concluded that a direct evaluation of the contact stress distribution and the relative displacements in the contact region is still an unsolved problem. To evade from this problem, in the next paragraph an indirect measuring method is described and elaborated in subsequent chapters for the case of the upsetting of cylindrical test pieces. The upsetting test is chosen because of its simple geometrical properties, thus excluding complications of geometrical source. As the method may be extended to other types of tests, the small range of contact conditions covered by the upsetting test, is not essential.

3.2 A hybrid method

Before the hybrid procedure, combining the capabilities of numerical and experimental research to quantify the contact behaviour, is described, some available experimental techniques for obtaining directly *in situ* data at contact regions are reviewed.

Experiments collect *in situ* data if results are obtained without influencing the process the object of interest is subjected to. For the contact processes considered here, such *in situ* techniques are preferable. The running in behaviour is known to show sometimes bad reproducibility (Schey 1983), mainly due to the difficulty to generate equal initial contact conditions for different experiments. For measurements during one intermitted experiment, it is impossible to restore the contact conditions after releasing the forming product, while besides effects of relaxation will interfere. To circumvent these problems here, only *in situ* techniques are considered, delivering data more or less continuously during one experiment. Reproducibility of the contact behaviour then becomes a phenomenon which can be studied afterwards, instead of being a disturbing factor. This however will not be treated further in this thesis.

Being interested in stresses and displacements at the contact region between tool and workpiece, only techniques for measuring these quantities are focussed on here. With respect to stresses some available techniques may be mentioned. Examples are techniques using the stress dependent fluoresence wavelength shift of ruby (Brouha et al. 1979) or light interference such as caustics (Theocaris & Razem 1979). They

require special tool material, combined with an appropriate design. Besides this disadvantage, the first method is unable to deliver the normal and the friction stress independently. Other stress measuring techniques such as stress sensitive film (Timothy et al. 1987) or diffraction techniques like neutron diffraction (Stacey et al. 1985) are not suitable for *in situ* application. The only known *in situ* technique applied in forming, delivering both normal and friction stress, employs pin pressure sensors (Van Rooyen & Backofen 1960, Vater & Nebe 1965, Tuncer & Dean 1987) with high sensitivity and reliability. Unfortunately the contact behaviour is influenced in an unknown way at the tool surface, resulting from the discontinuities in surface and stiffness. Further only a limited set of pins can be applied because of their size. With respect to displacement measurements even less possibilities are available, all of them only applicable in a limited way. Examples are measurements with grids (Lange 1985), not applicable *in situ*, or with rubies as used for stress measurements (Brouha et al. 1979), needing special tool materials and suffering from an unknown influence of the rubies on the local contact behaviour.

As proceeds from the previous, no satisfactory methods are available for a direct measurement of the contact stresses and displacements. A method for the determination of these quantities is therefore searched for in an inverse analysis of the mechanical behaviour of the formed body. The term inverse analysis is used here to indicate an analysis opposite to a direct analysis, solving a field problem with knowledge (Kubo 1988) of the
− boundary, enclosing the domain,
− field equations,
− boundary conditions for the whole boundary,
− initial conditions for time dependent problems,
− material properties,
− distributed sources in the domain.
For a direct analysis of static isothermal problems in continuum mechanics the above can be roughly translated to the
− boundary of the body under consideration,
− equilibrium equations and constitutive equations for the stresses,
− three independent boundary conditions for stresses or displacements,
− initial stress state,
− mechanical material properties occurring in the constitutive equations, e.g. elastic moduli and flow stress,
− distributed forces acting on the body.

Lack of one or more of the data required leads to an inverse problem. For the upsetting process this inverse problem concerns the identification of the unknown stresses and displacements at the contact surfaces of the formed body. To perform this identification, extra information is needed to compensate for the lacking boundary conditions.

It is assumed that the cylindrical test pieces remain axisymmetric and symmetric with respect to their midplane. In figure 3.2.1 a quarter of the cross section of an upsetting test piece in some deformed state is sketched. Indicated are the domain B of the body, contact boundary ∂B_c, the visible outer boundary ∂B_o, the symmetry plane boundary ∂B_s and the one-dimensional centre line boundary ∂B_{cl}.
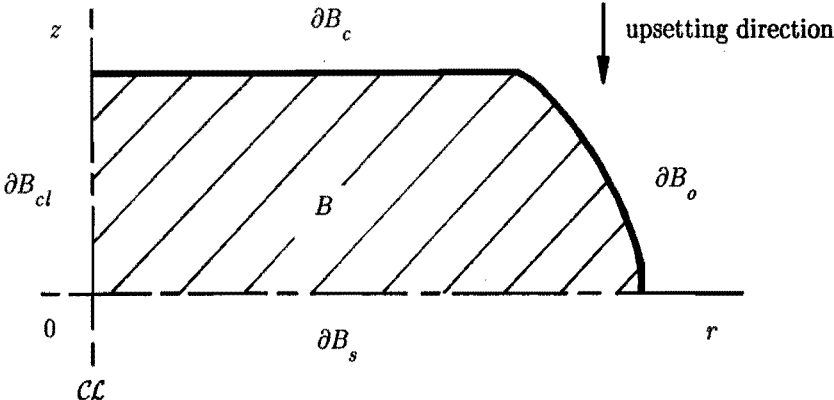
Figure 3.2.1 Right upper quarter of the cross section of an upsetting test piece

The boundary conditions on the parts ∂B_s and ∂B_{cl} are completely prescribed. The conditions on contact part ∂B_c are incompletely defined as only the axial displacement is known. To compensate the incompleteness, the displacements of ∂B_o are measured during upsetting. Therefore the conditions on ∂B_o become overdefined as then normal and shear stress are known, as well as axial and radial displacement. The photogrammetric method used for the displacement measurement is described in chapter 5.

Inverse problems of this kind are ill-posed. So the solution for contact stresses and radial displacements on ∂B_c may be ill-conditioned. Simulations in the next chapter, where the procedure is described to solve the inverse problem, indeed reveal this fact. To overcome this problem a contact stress measuring tool is designed which is indirectly used to measure the normal and shear stress at the contact

boundary. The stresses follow from the solution of an inverse problem with respect to the tool, comparable to the above posed problem. A detailed exposition is given in chapter 6. The stresses measured may be introduced as boundary conditions at the contact boundary ∂B_c, which causes the condition on this boundary to be overdefined and transforms the inverse problem into an overdetermined direct problem. The solution procedure presented in the next chapter also enables to deal with this type of problems, taking into account the accuracy of the different sets of measured boundary conditions. This method is of a hybrid character, as it combines numerical techniques with experimental data. It is therefore further termed an experimental–numerical method.

3.3 Estimations for the process conditions

The process conditions for cylindrical test pieces of commercially pure aluminium with an initial diameter D_0 and initial height h_0 of 60 mm are estimated. An approximate value for the maximum allowed forming velocity, in order to satisfy the isothermal conditions as indicated in chapter 1, is resulting. With respect to the chosen diameter it is noticed that the minimum contact area of the test pieces for which reasonable stress data can be obtained is limited by the size of the sensors of the stress measuring tool. Because the robustness of the strain sensors in the application of the measuring tool is *a priori* unknown, aluminium is chosen as a first test material because of its low yield stress and despite sometimes awkward behaviour in forming. In a later stadium other materials can be used.
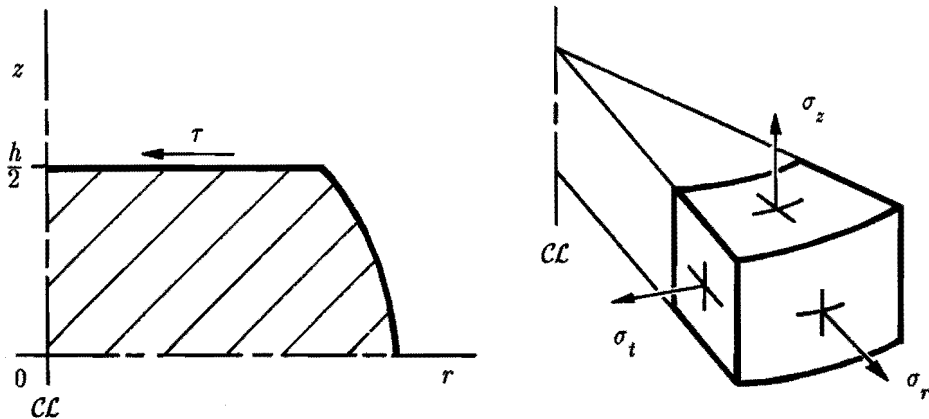
Figure 3.3.1 Some definitions of mechanical quantities

An elementary plasticity analysis of axisymmetric upsetting offers an estimate of the mechanical conditions (e.g. Lange 1985). Stresses and displacements are assumed to be independent of the axial coordinate z. As usual in this kind of analyses the influence of the friction stress τ at both sides of the upsetting cylinder is taken into account by a volumetric radial force $-2\tau/h$. The only non-trivial equilibrium equation is given by

$$\frac{d\sigma_r}{dr} + \frac{\sigma_r - \sigma_t}{r} = -\frac{2\tau}{h} \tag{3.3.1}$$

Applying the Levy–Von Mises flow rule and the proportionality of the radial velocity with r, the radial stress σ_r and tangential stress σ_t turn out to be equal. The Von Mises yield criterion relates σ_z and σ_r as

$$\sigma_r - \sigma_z = Y \tag{3.3.2}$$

where Y denotes the relevant yield stress. Assuming a Von Mises friction model with a constant friction factor m ($0 \leq m \leq 1$) implies

$$|\tau| = m\frac{Y}{\sqrt{3}} \tag{3.3.3}$$

and supposing the yield stress to be independent of r, it follows that

$$\sigma_z = -Y(1 + \frac{m}{\sqrt{3}}\frac{D-2r}{h}) \quad , \quad F_p = \frac{\pi}{4}D^2 Y(1 + \frac{m}{3\sqrt{3}}\frac{D}{h}) \tag{3.3.4}$$

with F_p the upsetting load. With $Y = 90$ N/mm^2 and $m = 0.1$, a maximum absolute value of the axial stress of about 100 N/mm^2 and an upsetting load of about 0.5 MN are found.

   To approximate the temperature rise to be expected, the upsetting cylinder and the press are modelled one–dimensionally. Because of symmetry with respect to the cylinder midplane, only the part $z \geq 0$ has to be considered. For $0 \leq z \leq h/2$ a homogeneous heat source H, equal to the sum of the local deformation power and the frictional power averaged over the volume, is assumed

$$H = (\sigma_z - |\tau|\frac{D}{h})\frac{\dot{h}}{h}$$ (3.3.5)

For $z > h/2$ (in the upsetting tool, tool holder and press) the cross section area $A(z)$ increases for increasing z. Neglecting the contact heat resistance between the domain parts, it follows as the steady state solution from Fourier's law

$$\Delta T = \frac{Hh^2}{8}(\lambda_c^{-1} + 2B\lambda_p^{-1}) \quad , \quad B = \frac{A(h/2)}{h/2}(\frac{dA}{dz})^{-1}$$ (3.3.6)

with ΔT the temperature rise at the origin and λ_c, λ_p the heat conductivities of the cylinder and press material respectively. Combining (3.3.5), (3.3.6) and assuming $B = 1, \frac{\dot{h}}{h}$ should obey

$$\frac{\dot{h}}{h} \geq \frac{8}{\sigma_z h^2 - |\tau| Dh} \frac{\lambda_c \lambda_p}{2\lambda_c + \lambda_p} \Delta T_{max}$$ (3.3.7)

to achieve $\Delta T \leq \Delta T_{max}$ with ΔT_{max} the maximum allowable temperature rise. An acceptable value for ΔT_{max} is 25 K, as the change of yield stress for aluminium is then about 5% ($\frac{1}{Y}\frac{\partial Y}{\partial T} \approx 0.002$ K^{-1}, derived from data of Brandes 1983). Taking $\lambda_c = 200$ W/mK, $\lambda_p = 40$ W/mK, $\sigma_z = -100$ N/mm^2 results in

$$\frac{\dot{h}}{h} \geq -0.01 \text{ s}^{-1}$$ (3.3.8)

So for forming velocities smaller than 0.6 mm/s the upsetting process may be supposed to be isothermal.

4 An experimental–numerical method

A method is described for the solution of history dependent non-linear mechanical continuum problems with exactly prescribed as well as not exactly prescribed or measured boundary conditions. An elaboration is presented for the axisymmetric case and rigid plastic material behaviour, in order to analyse the contact problem between die and upsetting workpiece as described in the previous chapter.

In the first paragraph a detailed description of the continuum problem with exact and measured boundary conditions is given. A distinction is made between the formulation of the mechanical behaviour and the incorporation of stochastic boundary conditions. The last necessitates for an estimation strategy. For the analysis of the mechanical behaviour the incremental finite element method is chosen. As estimation strategy the discrete maximum likelihood method is applied. An estimate according to this method maximizes the so-called likelihood function. Because of structural and geometrical non-linearity, this maximization has to be performed iteratively. Some aspects of the numerical implementation of the experimental-numerical method are discussed. The last paragraph deals with simulations of the upsetting process. These show that use of geometrical data of the visible surface of the workpiece does not suffice to achieve reliable results for stress and relative displacements at the contact surface, which supports observations as reported in literature (Herbertz & Wiegels 1981). Accounting for the measured upsetting force as extra input improves the results insignificantly. Measured contact stress data however introduce a tremendous improvement. If it is not possible to measure the contact stresses accurately or only partly, the displacements of the unloaded surface can improve the calculated contact quantities, provided they are measured with a high accuracy.

4.1 Introduction

Considered is an irreversible behaving body B enclosed by a boundary ∂B, with a known initial state at some reference time t_0. Both the body and its boundary consist of an invariant set of material points. Due to mechanical loads the displacements of the material points are large, which causes geometrical non-linearity. Thermal effects, inertial effects and mass distributed loads are supposed to be so small that they may be neglected.

The deformation of the body is considered during a time interval $t_0 \leq \tau \leq t$. At

any time τ and at any internal point of the body, local equilibrium requires for the Cauchy stress tensor σ to satisfy

$$\vec{\nabla}.\sigma = \vec{0} \tag{4.1.1}$$

The initial value of the stress tensor is denoted by σ_0, the actual value of σ at time t depends on the history $(t_0 \leq \tau \leq t)$ of the right Cauchy–Green tensor C, expressible by means of a constitutive functional S

$$\sigma = \sigma_0 + \mathop{S}_{\tau=t_0}^{t} [C(\tau)] \tag{4.1.2}$$

Mechanical properties figuring in S should be known for the material under consideration. With three independent boundary conditions at each point of ∂B, explicitly described during $t_0 \leq \tau \leq t$, equations (4.1.1) and (4.1.2) define a direct problem as indicated in the previous chapter. The solution can be approximated by a straightforward analysis using a numerical scheme such as the finite element method. In physical reality however sometimes boundary loads or positions are implicitly described, their values are not known or can only be measured with limited accuracy. To solve the problem for such cases with some wanted reliability, sufficient information of other quantities at the boundary or at interior points has to be supplied and taken into account in the calculation.

For the continuum problem a discretized approach is adopted. It is much easier to indicate a well-posed problem in a discretized approach than in a continuum approach. Extension of the set of discretization points in time and space increases the accuracy of the approximation. Measurements however are only possible in a limited number of material points, thus restricting the calculational capabilities.

In paragraph 4.2 the discretized strategy with respect to the solution of the equations (4.1.1) and (4.1.2) is discussed. An incremental finite element formulation is chosen. The material behaviour is assumed to be rigid plastic. In paragraph 4.3 the discretized estimation strategy for the incremental position field of the body with partly uncertain boundary conditions is elaborated. Information with respect to the uncertainty of the position field and derivable quantities can be obtained. The estimation strategy used is the maximum likelihood method which takes the stochastic behaviour of the different boundary conditions into account.

Many research has already been performed to combine numerical techniques for field problems such as the finite element method, and stochastic theories. A rough distinction is made into developments mainly supported by probability analyses and those supported by estimation theory. The first category focusses on the prediction of statistical properties, resulting from variations in applied loads, material properties and so on (e.g. Handa & Andersson 1981, Drewniak & Pawicki 1985, Liu et al. 1987). Results are useful for reliability studies and risk analyses. The second category involves the prediction of the behaviour of a structure or of uncertain parameters, given a set of observations (e.g. Contreras 1980, Kunisch & White 1986, Mura et al. 1986). An application to contact problems, combining the non-linear finite element method and estimation theory is not known to the author, however for the linear case some examples are available (e.g. Oda & Shinada 1987). Because of the resemblence with the linear analysis of the stress measuring tool in chapter 6, a detailed discussion is postponed to that chapter.

4.2 Strategy for the non–linear behaviour

A well-known strategy to determine approximate solutions of differential equations is based on the method of weighted residuals. The equilibrium equation (4.1.1), weighted over the domain B by a weighting vector field $\vec{w}$, in the weak formulation reads

$$\int_B (\vec{\nabla}\vec{w})^c : \sigma dV = \int_{\partial B} \vec{w}.\vec{b} dA \qquad (4.2.1)$$

with $\vec{b}$ the external boundary load vector. Substitution of the constitutive equation (4.1.2) delivers an integral equation for the position field of the material points of the body for any τ, $t_0 \leq \tau \leq t$. Elaboration of that equation depends on the nature of the constitutive equation. Here rigid plastic material behaviour, obeying the Levy–Von Mises flow rule presupposing isotropic hardening, is chosen

$$\sigma = -pI + \frac{2}{3}\frac{Y}{\dot{\bar{\epsilon}}}D \quad , \quad p = -\frac{1}{3}\mathrm{tr}(\sigma) \quad , \quad \dot{\bar{\epsilon}} = (\frac{2}{3}D:D)^{\frac{1}{2}} \qquad (4.2.2)$$

D is the deformation rate tensor and $\bar{\epsilon}$ is the equivalent plastic strain. Equation

(4.2.2) supposes D to be unequal to $\mathbf{0}$. The stress tensor σ satisfies the Von Mises flow criterion

$$\frac{3}{2}\sigma^d : \sigma^d = Y^2 \quad , \quad \sigma^d = \sigma + pI = \sigma - \frac{1}{3}\mathrm{tr}(\sigma)I \qquad (4.2.3)$$

The yield stress Y for many metals under cold forming conditions in non-cyclic processes only depends on the history parameter $\bar{\epsilon}$. At elevated temperature and for complex deformation processes, more history parameters and modified expressions for (4.2.2) and (4.2.3) are needed (e.g. Frost & Ashby 1982, Meyer-Nolkemper 1982, Chan et al. 1988). Elastic effects are not modelled by (4.2.2), as they may be neglected for large plastic strains. This is quite convenient as will be seen in the next paragraph. In the constitutive model the pressure p occurs as an extra unknown. This pressure should result from the requirement of incompressibility, $\mathrm{tr}(D) = 0$. This requirement is taken into account, weighted over B with a scalar weighting field q

$$\int_B \mathrm{tr}(D)q\,dV = 0 \qquad (4.2.4)$$

The time domain $t_0 \leq \tau \leq t$ is discretized in finite steps called increments. The deformation rate tensor D is assumed to be constant during an increment Δt. For small incremental displacements $\vec{u}$, the tensor D can be approximated by

$$D = \frac{A}{\Delta t} \quad , \quad A = \frac{1}{2}\big(\vec{\nabla}\vec{u} + (\vec{\nabla}\vec{u})^c\big) \qquad (4.2.5)$$

where the gradient operator $\vec{\nabla}$ is defined with respect to the configuration at the end of the increment. From (4.2.2) and (4.2.5) it follows

$$\sigma = -pI + \frac{2}{3}\frac{Y(\bar{\epsilon}_b + \Delta\bar{\epsilon})}{\Delta\bar{\epsilon}}A \quad , \quad \Delta\bar{\epsilon} = \big(\tfrac{2}{3}A{:}A\big)^{\frac{1}{2}} \qquad (4.2.6)$$

with $\bar{\epsilon}_b$ the equivalent plastic strain at the beginning of the increment and provided $A \neq \mathbf{0}$, while instead of (4.2.4) can be required

$$\int_B \text{tr}(A)q\,dV = 0 \tag{4.2.7}$$

The equations (4.2.1) and (4.2.7) should hold at the end of each increment for arbitrary allowable weighting fields, $\vec{w}$ and q respectively. Besides, (4.2.6) should be satisfied. Substitution of (4.2.6) in (4.2.1) results in

$$\int_B (\vec{\nabla}\vec{w})^c : (-pI + \frac{2}{3}\frac{Y(\overline{\epsilon}_b + \Delta\overline{\epsilon})}{\Delta\overline{\epsilon}}A)\,dV = \int_{\partial B} \vec{w}.\vec{b}\,dA \tag{4.2.8}$$

The domain B is divided into finite subdomains, called elements. A Lagrangian approach is adopted here, so the elements are material elements. In the elements the displacements, the pressure and the weighting fields are interpolated between the nodal points. In case of a Galerkin approach for the element assembly holds

$$\vec{u} = \sum_{i=1}^{n} \varphi^i \vec{u}^i \quad , \quad \vec{w} = \sum_{i=1}^{n} \varphi^i \vec{w}^i \quad , \quad p = \sum_{j=1}^{m} \psi^j p^j \quad , \quad q = \sum_{j=1}^{m} \psi^j q^j \tag{4.2.9}$$

with n, m the number of interpolation functions, for the displacements and pressure respectively. Applying the principle of weighted residuals, i.e. requiring satisfaction of (4.2.8) and (4.2.7) for all allowable weighting fields, results in a set of vector equations (related to equilibrium) and scalar equations (related to volume invariance) for each increment. With respect to some vector base these equations formally can be written as

$$\begin{bmatrix} \underaccent{\tilde}{f}(\underaccent{\tilde}{u},\underaccent{\tilde}{p}) \\ \underaccent{\tilde}{v}(\underaccent{\tilde}{u}) \end{bmatrix} = \begin{bmatrix} \underaccent{\tilde}{0} \\ \underaccent{\tilde}{0} \end{bmatrix} \tag{4.2.10}$$

where $\underaccent{\tilde}{u}$ and $\underaccent{\tilde}{p}$ are composed of the nodal values of the displacements and pressure. As the set (4.2.10) is non-linear in the displacements and pressures, necessitating an iterative solution procedure, the iterative changes

$$\begin{bmatrix} \delta\underaccent{\tilde}{f} \\ \delta\underaccent{\tilde}{v} \end{bmatrix} = \begin{bmatrix} \underaccent{\tilde}{A} & \underaccent{\tilde}{B} \\ \underaccent{\tilde}{C} & \underaccent{\tilde}{0} \end{bmatrix} \begin{bmatrix} \delta\underaccent{\tilde}{u} \\ \delta\underaccent{\tilde}{p} \end{bmatrix} \tag{4.2.11}$$

are of interest. This relation will be applied for the numerical elaboration.

When displacement and pressure fields are known, the stresses at an arbitrary point of B can be calculated using (4.2.6) and the interpolations per element. It is noticed that the equivalent plastic strain must be interpolated also, as it is only known in a discrete number of points per element at the start of a particular increment. The stresses calculated this way are generally discontinuous over the element boundaries. Because in reality these fields are continuous, some smoothing technique is usually applied. Here a straightforward unweighted averaging of stresses in the nodal points at the element boundaries is used. In the linear case of chapter 6, an alternative formulation is applied to reach continuity. If stress data occur as boundary conditions, the iterative stress may be needed for the solution procedure. The iterative column $\delta\underset{\sim}{\sigma}$ with stress components at some point of B can be written as

$$\delta\underset{\sim}{\sigma} = [\underline{D}\ \underline{E}] \begin{bmatrix} \delta\underset{\sim}{u} \\ \delta\underset{\sim}{p} \end{bmatrix} \tag{4.2.12}$$

The matrices $\underline{D}$ and $\underline{E}$ depend on $\underset{\sim}{u}$ and $\underset{\sim}{p}$. For spatial fixed boundary points they differ from the matrices for material fixed boundary points because of the influence of convective terms.

4.3 Strategy for the state estimation

In the previous paragraph the behaviour of a body of rigid plastic material is approximated with a discretized incremental formulation. Assuming the initial conditions and material behaviour to be known, the mechanical condition of the body after the first increment is described by $n+m$ nodal unknowns, composed of n displacement degrees of freedom and m pressure degrees of freedom, in the previous denoted by $\underset{\sim}{u}$ and $\underset{\sim}{p}$ respectively. The n displacements arbitrarily can be replaced by n nodal locations in the sequel. For some k–th increment the condition of the body at the end of the increment depends on n location degrees of freedom and m pressure degrees of freedom more than its condition at the begin of the increment. These nodal degrees of freedom are gathered in the column $\underset{\sim}{s}(k)$, further called the incremental state of the k–th increment. After k increments, the total of $(n+m)k$ unknowns is gathered in the accumulated state $\underset{\sim}{s}_K$. In this way the accumulated state

$$\underset{\sim}{s}_K^T = [\underset{\sim}{s}(1)^T\ \underset{\sim}{s}(2)^T...\underset{\sim}{s}(k)^T] \tag{4.3.1}$$

is the combination of the incremental states $\underset{\sim}{s}(1)$ to $\underset{\sim}{s}(k)$.

To quantify the accumulated state $\underset{\sim}{s}_N$, belonging to a deformation process over a time period divided into N increments, at least $(n+m)N$ values of relevant quantities with respect to the condition of the body during the time period have to be known. A part of these quantities will be exactly known, the rest of them measurable with only a limited accuracy. As relevant quantities not only the nodal degrees of freedom have to be thought of, but also quantities like stresses, strains and forces. The relevant quantities $\underset{\sim}{h}_e(k)$ and $\underset{\sim}{h}_m(k)$ of the k-th increment can uniquely be expressed in the accumulated state $\underset{\sim}{s}_K$, so

$$\underset{\sim}{h}_e(k) = \underset{\sim}{h}_e(\underset{\sim}{s}_K) \quad , \quad \underset{\sim}{h}_m(k) = \underset{\sim}{h}_m(\underset{\sim}{s}_K) \tag{4.3.2}$$

where e and m indicate that the values of these quantities are exactly known respectively measurable. The quantities $\underset{\sim}{h}_e(k)$ are called the exact state dependent quantities, $\underset{\sim}{h}_m(k)$ the measurable state dependent quantities. The actual data values for $\underset{\sim}{h}_e(k)$ compose the column $\underset{\sim}{z}_e(k)$, the exact observation data, and the actual data values for $\underset{\sim}{h}_m(k)$ the column $\underset{\sim}{z}_m(k)$, the measured observation data. Per increment, the incremental exact observation data $\underset{\sim}{z}_e(k)$ are equal to $\underset{\sim}{h}_e(k)$, whereas the incremental measured observation data $\underset{\sim}{z}_m(k)$ are supposed equal to $\underset{\sim}{h}_m(k)$ to which an unknown stochastic column $\underset{\sim}{v}(k)$ with incremental observation errors is added

$$\underset{\sim}{z}_e(k) = \underset{\sim}{h}_e(\underset{\sim}{s}_K) \quad , \quad \underset{\sim}{z}_m(k) = \underset{\sim}{h}_m(\underset{\sim}{s}_K) + \underset{\sim}{v}(k) \tag{4.3.3}$$

With the same convention as for the state, accumulated columns $\underset{\sim}{h}_{e,K}$, $\underset{\sim}{h}_{m,K}$, $\underset{\sim}{z}_{e,K}$ and $\underset{\sim}{z}_{m,K}$ are defined using the incremental columns $\underset{\sim}{h}_e(k)$, $\underset{\sim}{h}_m(k)$, $\underset{\sim}{z}_e(k)$ and $\underset{\sim}{z}_m(k)$. So it can be stated for the accumulated exact and measurable state dependent quantities

$$\underset{\sim}{h}_{e,K} = \underset{\sim}{h}_{e,K}(\underset{\sim}{s}_K) \quad , \quad \underset{\sim}{h}_{m,K} = \underset{\sim}{h}_{m,K}(\underset{\sim}{s}_K) \tag{4.3.4}$$

The incremental observation errors of k increments combine into the accumulated observation error column $\underset{\sim}{v}_K$.

The problem to be solved for the deforming body, over a time period divided into a total of N increments, can now be stated as searching that solution $\hat{\underset{\sim}{s}}_N$ for $\underset{\sim}{s}_N$ which obeys

$$z_{e,N} = h_{e,N}(\hat{s}_N) \tag{4.3.5}$$

and satisfies some criterion with respect to

$$z_{m,N}, h_{m,N}, v_N, \hat{s}_N \tag{4.3.6}$$

The above formulation constitutes a static non-linear estimation problem with constraints. The solution $\hat{s}_N$, hopefully close to the unknown true accumulated state, is called the estimate and the algorithm or expression leading to it the estimator. For many estimation problems estimators that are optimal in the sense of some criterion are known (see e.g. Schweppe 1973, Norton 1986, Jansen 1987[1]). Before the above stated constrained problem is elaborated, some estimators for static non-linear estimation problems without constraints are reviewed.

In this intermediate consideration all measured observation data are gathered in z, measurable state dependent quantities in h, observation errors in v and the state in s. These columns are related by

$$z = h(s) + v \tag{4.3.7}$$

If no extra information about the observation errors is available, a least squares method can be applied. This method proposes as estimate for the state s, that solution $\hat{s}$, which minimizes the quadratic form $(z-h)^T(z-h)$. A weighted least squares method takes knowledge about relative accuracy of the measurements into account. The expression $(z-h)^T W(z-h)$ with a positive definite weighting matrix W is then minimized, to reach that more accurate data influence the solution stronger than less accurate data. Statistical properties as the mean and covariance are used by minimal covariance estimators, also called Markov estimators. To define these statistical properties, the expected value operator E is of importance. The operator E provides the expected value of the operand $g(v)$, which is some function of the random variables v. The random variables v are supposed to have a (multivariate) probability density function $p(v)$. It holds by definition

$$E(g(v)) = \int \dots \int g(v) p(v) \, dv_1 \dots dv_l \tag{4.3.8}$$

with l the number of random variables i.e. the length of $\underset{\sim}{v}$. The integrations in (4.3.8) are over the entire domain of $\underset{\sim}{v}$. The mean column $\underset{\sim}{\mu}$ and covariance matrix $\underset{\sim}{\Sigma}$ of $\underset{\sim}{v}$ are defined as

$$\underset{\sim}{\mu} = E(\underset{\sim}{v}) \quad , \quad \underset{\sim}{\Sigma} = E\big((\underset{\sim}{v}{-}\underset{\sim}{\mu})(\underset{\sim}{v}{-}\underset{\sim}{\mu})^T\big) \tag{4.3.9}$$

where the expected values are assigned per component of the operand. The mentioned minimal covariance estimators generally suppose the mean and covariance matrix of the errors as known, according to

$$\underset{\sim}{\mu} = \underset{\sim}{0} \quad , \quad \underset{\sim}{\Sigma} = \underline{R} \tag{4.3.10}$$

The estimate $\hat{\underset{\sim}{s}}$ follows from requiring a minimum for some norm of the covariance matrix of $\hat{\underset{\sim}{s}}$. For the linear case such an estimator is given in paragraph 6.2.3. The supposition of zero mean in (4.3.10) reflects the absence of systematic errors with respect to the observation data and of model errors for the state dependent quantities. If not only (4.3.10) is known but moreover the probability density function $p(\underset{\sim}{v})$, a maximum likelihood estimate for $\underset{\sim}{s}$ is possible. It maximizes for the actual observation data $\underset{\sim}{z}$ the likelihood function with respect to $\underset{\sim}{s}$. The likelihood function is defined as the probability density function $p(\underset{\sim}{z}:\underset{\sim}{s})$ of the observation data $\underset{\sim}{z}$ given some state $\underset{\sim}{s}$. For non-linear estimation problems this estimate is asymptotic unbiased and asymptotic efficient under fairly general conditions. An explanation of these properties is given in appendix 4.3.1.

Returning to the estimation problem defined for the deforming body a reliable estimate $\hat{\underset{\sim}{s}}_N$ can be obtained with a maximum likelihood estimator. For measuring errors generally a Gaussian probability applies. For $\underset{\sim}{v}_N$ with zero mean, the probability density function then reads

$$p(\underset{\sim}{v}_N) = \big((2\pi)^{l_N}\det(\underline{R}_N)\big)^{-\frac{1}{2}} \exp(-\tfrac{1}{2}\underset{\sim}{v}_N^T\underline{R}_N^{-1}\underset{\sim}{v}_N) \tag{4.3.11}$$

with l_N the length of $\underset{\sim}{v}_N$ and $\underline{R}_N$ its covariance matrix. Replacing $\underset{\sim}{v}_N$ by $\underset{\sim}{z}_{m,N}{-}\underset{\sim}{h}_{m,N}(\underset{\sim}{s}_N)$ in (4.3.11) yields the likelihood function $p(\underset{\sim}{z}_N:\underset{\sim}{s}_N)$. Maximizing this function for the actual observation data $\underset{\sim}{z}_{m,N}$ under the constraints (4.3.5) delivers the wanted estimate $\hat{\underset{\sim}{s}}_N$.

A disadvantage of this method is of computational origin. To obtain a reasonable approximation of the non-linear behaviour of body B, the incremental state columns with each $n+m$ nodal degrees of freedom has to be relatively large, as well as the total number of increments N. This results in an accumulated state column $\underset{\sim}{s}_N$ of length $(n+m)N$. To deal with this column in a maximization process of the function $p(\underset{\sim}{z}_N : \underset{\sim}{s}_N)$, depending on $\underset{\sim}{s}_N$ in a complex way, would need large computation time and memory. For this reason, among others, in estimation theory for dynamic systems filtering techniques are developed to determine estimates for one time step instead of simultaneously for the total set of time steps. The estimation of the incremental state $\underset{\sim}{s}(k)$ for the k-th time step is then based on momentary observation data and some estimate $\hat{\underset{\sim}{s}}_{K-1}$ for the accumulated state $\underset{\sim}{s}_{K-1}$. For the estimation of $\underset{\sim}{s}(k)$ in the considered case a simple filter is proposed in the sequel, based on the previous incremental state estimate $\hat{\underset{\sim}{s}}(k-1)$ and the incremental observation data $\underset{\sim}{z}(k)$.

In the foregoing $\underset{\sim}{h}_e(k)$ and $\underset{\sim}{h}_m(k)$ are expressed as functions of $\underset{\sim}{s}_K$, see (4.3.2), thus of $\underset{\sim}{s}_{K-1}$ and $\underset{\sim}{s}(k)$. Noting the particular history dependent material behaviour, as described in paragraph 4.2, these variables may be replaced by $\underset{\sim}{H}(k)$, $\underset{\sim}{s}(k)$ and $\Delta\underset{\sim}{s}(k)$, with $\Delta\underset{\sim}{s}(k)$ the incremental state change defined by

$$\Delta\underset{\sim}{s}(k) = \underset{\sim}{s}(k) - \underset{\sim}{s}(k-1) \tag{4.3.12}$$

and $\underset{\sim}{H}(k)$ a column with the history variables resulting after the k-th increment i.c. the equivalent plastic strains at a discrete number of material points. The choice of the material behaviour enables that they can be expressed additively as

$$\underset{\sim}{H}(k) = \underset{\sim}{H}(k-1) + \Delta\underset{\sim}{H}\big(\underset{\sim}{s}(k), \Delta\underset{\sim}{s}(k)\big) \tag{4.3.13}$$

with $\Delta\underset{\sim}{H}$ a non-linear function of $\underset{\sim}{s}(k)$ and $\Delta\underset{\sim}{s}(k)$. It is noticed that the incremental state change $\Delta\underset{\sim}{s}(k)$ contains the nodal displacements and the changes of the pressure degrees of freedom of the k-th increment. If an estimate $\hat{\underset{\sim}{s}}(k-1)$ for the incremental state $\underset{\sim}{s}(k-1)$ is available, $\underset{\sim}{s}(k-1)$ may be written as

$$\underset{\sim}{s}(k-1) = \hat{\underset{\sim}{s}}(k-1) + \delta\underset{\sim}{s}(k-1) \tag{4.3.14}$$

with errors $\delta\underset{\sim}{s}(k-1)$. Also errors $\delta\underset{\sim}{H}(k-1)$ will apply for $\underset{\sim}{H}(k-1)$. Assuming these

errors relatively small, for the quantities $\underset{\sim}{h}_e(k)$ holds

$$\underset{\sim}{h}_e(k) = \underset{\sim}{h}_e\Big(\hat{H}(k{-}1){+}\Delta\underset{\sim}{H}\big(\underset{\sim}{s}(k),\underset{\sim}{s}(k){-}\hat{\underset{\sim}{s}}(k{-}1)\big),\underset{\sim}{s}(k),\underset{\sim}{s}(k){-}\hat{\underset{\sim}{s}}(k{-}1)\Big)+$$

$$-\Big(\frac{\partial}{\partial\Delta\underset{\sim}{s}}(\underset{\sim}{h}_e^T)\Big)^T\delta\underset{\sim}{s}(k{-}1)+\Big(\frac{\partial}{\partial H}(\underset{\sim}{h}_e^T)\Big)^T\Big(\delta\underset{\sim}{H}(k{-}1){-}\Big(\frac{\partial}{\partial\Delta\underset{\sim}{s}}(\Delta\underset{\sim}{H}^T)\Big)^T\delta\underset{\sim}{s}(k{-}1)\Big) \quad (4.3.15)$$

where quadratic and higher order terms in the errors are neglected. A similar relationship for $\underset{\sim}{h}_m(k)$ can be stated. The only unknown in the zero-th order term of (4.3.15) and the zero-th order term for $\underset{\sim}{h}_m(k)$, is the incremental state $\underset{\sim}{s}(k)$. If the first order terms in the errors are negligible, a maximum likelihood estimate $\hat{\underset{\sim}{s}}(k)$ for $\underset{\sim}{s}(k)$ can be obtained in a straightforward way. If the first order terms are not negligible, the estimate $\hat{\underset{\sim}{s}}(k)$ solely based on the zero-th order term will contain an extra bias error, besides the bias error caused by the estimation method. The bias error $\underset{\sim}{b}(k)$ of the estimate $\hat{\underset{\sim}{s}}(k)$ is defined by

$$\underset{\sim}{b}(k) = E\big(\hat{\underset{\sim}{s}}(k)\big){-}\underset{\sim}{s}_t(k) \quad (4.3.16)$$

with $\underset{\sim}{s}_t(k)$ the true incremental state, see also appendix 4.3.1.

For the analysis of the upsetting experiment, for which the bias errors are expected to remain small, the maximum likelihood estimator using only the zero–th order terms is chosen. A part of the exact and measurable state dependent quantities consists of positions of boundary nodes, linear in $\underset{\sim}{s}(k)$. This contributes to a bounding of the bias terms for two reasons. Firstly the maximum likelihood method delivers bias free estimates for models with only linear state dependent quantities.

Secondly such state dependent quantities force the estimator to choose $\hat{\underset{\sim}{s}}(k)$ close to the true incremental state $\underset{\sim}{s}_t(k)$ despite possible bias errors in $\hat{H}(k{-}1)$, caused by the maximum likelihood method. In that case the expression for the error covariance matrix, as given in appendix 4.3.1, offers a lower bound for the reliability of the resulting incremental state estimate. The estimation problem for the one-dimensional tension experiment on elasto-plastic material, as investigated by Jansen (1987[2]), indeed showed the bounding of bias errors.

Another aspect, highlighted in Jansen's work, is the consequence of sudden changes in constitutive behaviour. In elasto-plasticity such changes occur when, due

to local unloading, the material behaviour changes from elasto-plastic to purely elastic. For the case considered by Jansen, a satisfactory solution is found with a hypothesis testing method, see appendix 4.3.2. In a multidimensional situation, an implementation of this method would be highly combinatorial and hence computation time consuming. Here this implementation is avoided by assuming the material behaviour purely plastic according to paragraph 4.2. This assumption will lead to erroneous results or even no solution when a forming process is analysed in which large elastic zones occur. In the upsetting configuration under low or mild frictional conditions and with not too slender workpieces, elasticity exclusively occurs at the very beginning of the loading. For doubtful cases, the results can be checked by using them as input for a standard elasto-plastic analysis program (MARC 1988, DIANA 1988).

4.4 Numerical implementation

In this paragraph some aspects of the numerical implementation of the incremental state estimation are explained. Attention is paid to the chosen finite element discretization, the optimization method and the boundary conditions.

Because of the promising numerical results for incompressible flow (e.g. Van de Vosse 1987), a triangular isoparametric $P_2^+-P_1$ Crouzeix-Raviart element is applied for the numerical simulations. This element, with an extended quadratic interpolation for shape and position field and a linear pressure field discontinuous over the element boundaries, provides accurate results for quite coarse meshes. It further shows the capability of numerical smoothing in the neighbourhood of singular points. Such singularities can occur under sticking conditions at the outer edge of the contact boundary (Van Wijngaarden 1988). To save computation time, in (Eulerian) flow calculations the internal node velocities and pressure derivatives can be eliminated on element level (Griffiths 1979). In the (Lagrangian) approach adapted here the internal node positions and pressure gradient could be eliminated on element level using the absence of mass distributed loads and the local constraint of volume invariance. The optimization per iteration then can be performed on the remaining unknowns. However to maintain the local constraints during this global optimization, calculations on element level are continuously needed, resulting in only little or no savings on computation time. Elimination of constraints on element level is therefore not performed, resulting in a larger set of global unknowns, but a simpler optimization algorithm.

Various solution methods exist to tackle optimization problems (see e.g. Gill et al. 1981, Schoofs 1987), many of them are commercially available in numerical libraries (IMSL 1987, NAG 1987). For the optimization considered here, the solution method has to minimize the following function

$$\left(z_m - h_m(s)\right)^T R_m^{-1}\left(z_m - h_m(s)\right) \tag{4.4.1}$$

with respect to the unknowns s, with the non-linear equality constraints

$$z_e = h_e(s) \tag{4.4.2}$$

The penalty function approach is a very elegant and comprehensible method to deal with the constraints (4.4.2). This method transforms the constrained minimization of (4.4.1) into the unconstrained minimization of

$$F(s) = \left(z_m - h_m(s)\right)^T R_m^{-1}\left(z_m - h_m(s)\right) + \left(z_e - h_e(s)\right)^T R_e^{-1}\left(z_e - h_e(s)\right) \tag{4.4.3}$$

The penalty matrix R_e^{-1}, not necessarily diagonal, may be interpreted as a fictitious inverse covariance matrix of the exact data z_e. This interpretation is helpful to choose reasonable values for R_e in relation to R_m. Because of its robustness, the Newton method is used to perform the unconstrained minimization in an iterative manner. This search direction method uses per iteration step the direction

$$-H(s)^{-1}g(s) \tag{4.4.4}$$

to determine a new approximation $\hat{s}$ for s, such that along the search path $\hat{s}$ minimizes $F(s)$. In (4.4.4), g and H are the gradient column and Hessian matrix of $F(s)$, respectively defined as

$$g = \frac{\partial F}{\partial s} \quad , \quad H = \frac{\partial}{\partial s}\left(\left(\frac{\partial F}{\partial s}\right)^T\right) \tag{4.4.5}$$

For the function $F(s) = f(s)^T P f(s)$ considered here, with P composed of R_m^{-1} and R_e^{-1}, neglecting the second derivative of g with respect to s yields

$$\underline{H}(\underline{s}) = 2\underline{J}(\underline{s})^T \underline{R}_t \underline{J}(\underline{s}) \quad , \quad \underline{J}(\underline{s})^T = \frac{\partial}{\partial\underline{s}}(\underline{f}^T)$$ (4.4.6)

where $\underline{J}$ is the Jacobian matrix of the functions $\underline{f}$ with respect to the variables $\underline{s}$. For the one-dimensional optimization along the search direction, a golden section search is employed, needing only the evaluation of function values.

In the upsetting experiment, different sets of boundary conditions can be distinguished. These will be kinematical, i.e. positions of boundary nodes, or dynamical, i.e. stresses or forces. These quantities sometimes can be considered as exact boundary conditions. In the interior of the body there are constraints of zero nodal forces and equations following from volume invariance. The exact boundary conditions (except exact nodal point positions as these can be prescribed easily) and all constraints are included in the function $\underline{f}$ mentioned above. With use of the iterative equations of paragraph 4.2 and the quite trivial iterative equations for nodal point positions, the Jacobian matrix is analytically known. Therefore advantage can be taken of the structure of $\underline{J}$ and $\underline{H}$, resulting in savings in computation time and memory.

Finally it is noticed that in the simulations of the next paragraph, differences in axial positions of contact points of the upsetting tool are neglected. With a maximum contact pressure of the order of 100 N/mm^2, a maximum absolute value of the axial elastic strain near the centre line of about $5*10^{-4}$ (elastic modulus E of the tool taken $2.1*10^5$ N/mm^2) may occur. Over a distance of 100 mm this would cause a displacement difference of 0.05 mm, which is an estimation for the displacement difference between the contact points of the upsetting tool near the centre line and the points at the outer edge of the contact area. With respect to the total dimensions and deformations of the test pieces, this is so small that the tool is assumed to remain flat.

4.5 Simulations of the upsetting experiment

In this paragraph results of calculations for the upsetting test, based on the earlier described numerical procedure, are presented. The aim is to gain insight in the accuracy of the contact quantities which can be expected for different sets of boundary conditions, i.e. different experimental set-ups. From this paragraph it is concluded that the quantities at the contact boundary of the upsetting test piece,

cannot be calculated accurately enough from the displacements of the unloaded outer surface. Measurement of the contact stresses (or the radial displacements of points at the contact boundary) is necessary for a reliable quantification of the contact behaviour. However when it is not possible to measure these contact stresses accurately, the displacements of the outer surface measured with a high accuracy can improve the calculated contact quantities.

The calculations are executed on a configuration as sketched in figure 4.5.1.
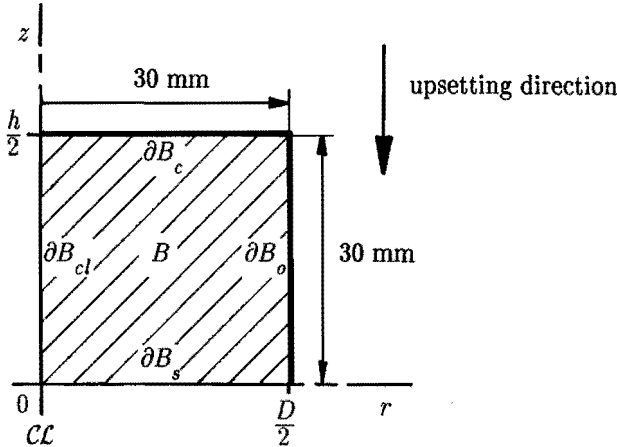
Figure 4.5.1 The upsetting configuration in the reference state

The yield stress Y as function of the equivalent plastic strain is taken as

$$Y = 32 + 120\bar{\epsilon} - 60\bar{\epsilon}^2 + 99\bar{\epsilon}^3 \quad (\text{N/mm}^2) \tag{4.5.1}$$

which results from a fit through data of a Rastegaev test on pure aluminium of Reiss & Pöhlandt (1985). In 10 increments of 3 mm upsetting each, the workpiece is upsetted to a remaining height of 30 mm. The non-trivial contact quantities of interest at boundary part ∂B_c are the radial displacements with respect to the reference configuration, the normal stress and the shear stress. For the boundary parts ∂B_c and ∂B_o the boundary conditions, used for the calculations, are varied as indicated in table 4.5.1. Results from case 1 will serve as input for the other calculations, which in fact should result in the same solution. The covariance estimates according to appendix 4.3.1 are used to study for each case the reliability of this solution at the contact boundary. Case 3 serves to study the influence of the measured upsetting force as extra boundary condition with respect to case 2.

boundary condition		case 1	2	3	4	5
∂B_c	measured axial displacement	×	×	×	×	×
	Von Mises friction model	×				
	measured normal contact stress				×	×
	measured upsetting force			×		
∂B_o	stress free	×	×	×	×	×
	measured displacements		×	×		×

Table 4.5.1 The boundary conditions of the 5 cases considered

Case 5 aims to investigate whether measured displacements of the outer unloaded surface contribute to an improvement of the contact quantities accuracy as resulting from case 4. For all cases the standard deviation of the axial displacement at the contact boundary is supposed to be 0.01 mm, which is quite accurate. All measured quantities are supposed to be mutual independent. The experiences obtained from the different problem formulations are discussed sequentially.

<u>Case 1</u> The boundary conditions as specified in table 4.5.1, together with the trivial boundary conditions along ∂B_{cl} and ∂B_s stated in chapter 3, constitute a direct problem. For the constant friction factor m of the Von Mises model, a value of 0.1 is applied. The standard deviation of the friction stress calculated with this model is assumed to be 0.1 N/mm^2.
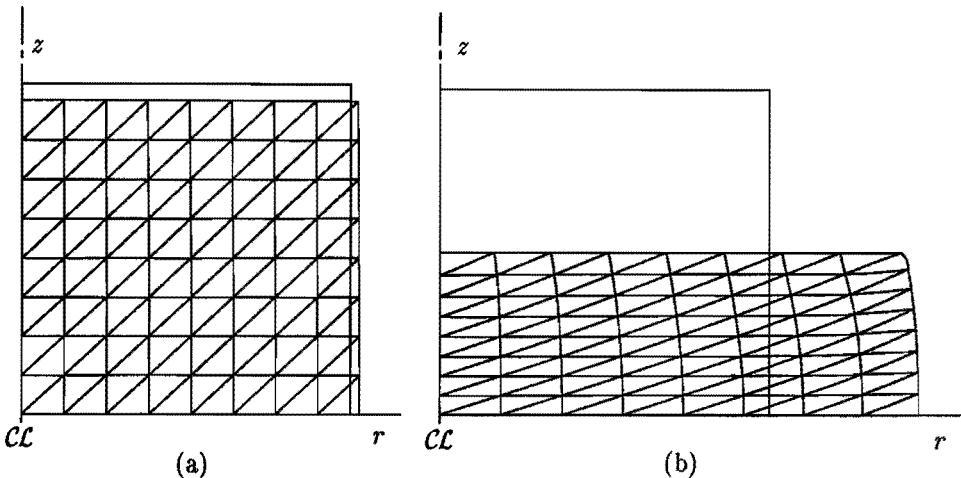
Figure 4.5.2 Deformed meshes after 3 mm (a) and 30 mm (b) upsetting

4 An experimental—numerical method

In figure 4.5.2 the deformed meshes after 3 mm upsetting (a) and 30 mm upsetting
(b) are shown. In the sequel the results will be presented always for these two
upsetting stages. The distribution of the contact quantities at these stages, the
radial displacement $\Delta r = r-r_0$, the normal stress σ_n and the frictional shear stress τ
are given in the figures 4.5.3, 4.5.4, and 4.5.5 respectively. The displacement is
considered as a function of the reference radial position, the stresses as a function of
the actual radial position.
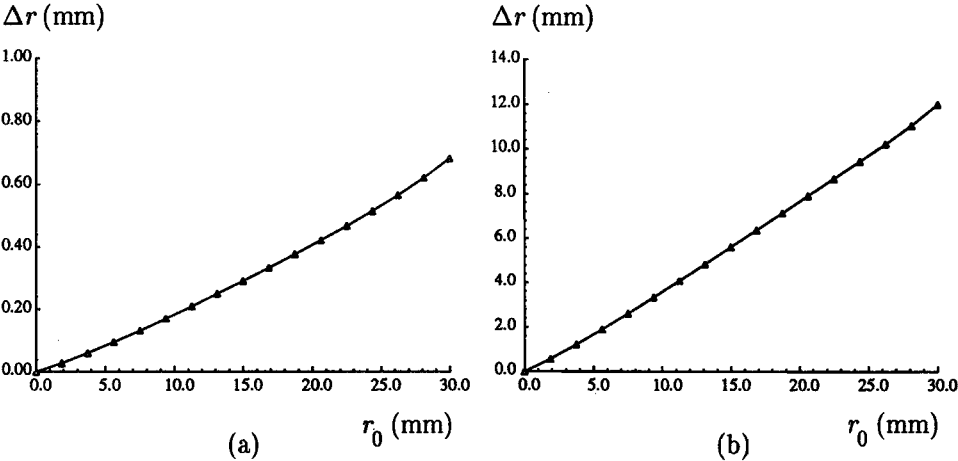
Figure 4.5.3 The radial displacement Δr at the contact surface
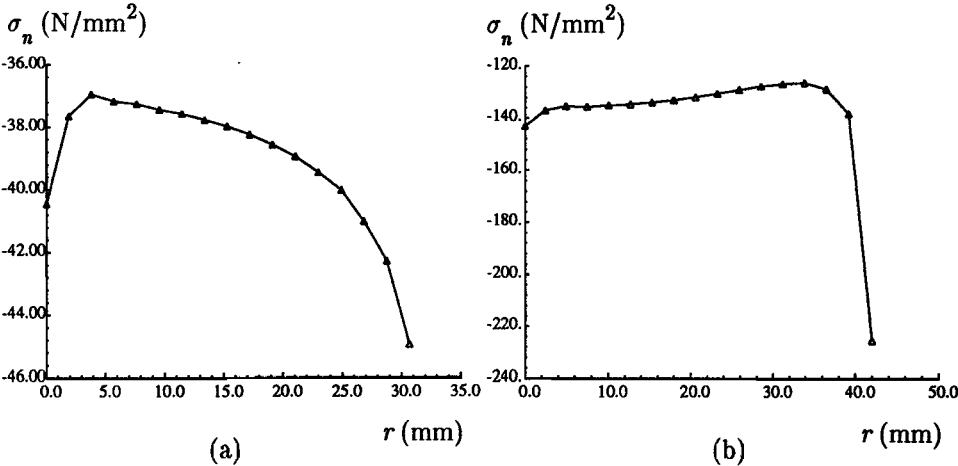
Figure 4.5.4 The normal stress σ_n at the contact surface
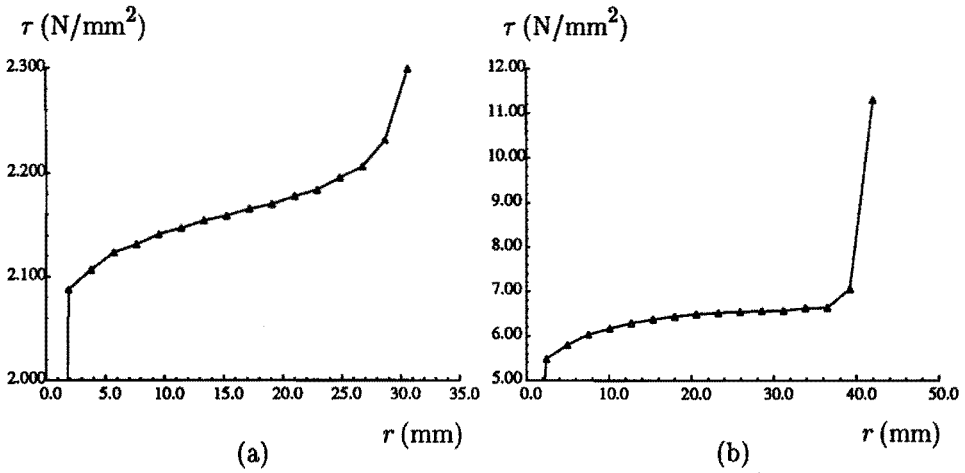
Figure 4.5.5 The friction stress τ at the contact surface.

The irregular behaviour of the normal stress near the centre line is partly caused by the applied stress interpolation. Further it is noticed that the height of the friction hill in the high upsetting range is in fair accordance with the value predicted by equation (3.3.4). As a measure for the reliability of the contact quantities a dimensionless standard deviation s_d is introduced, which is comparable with the coefficient of variation from statistics.
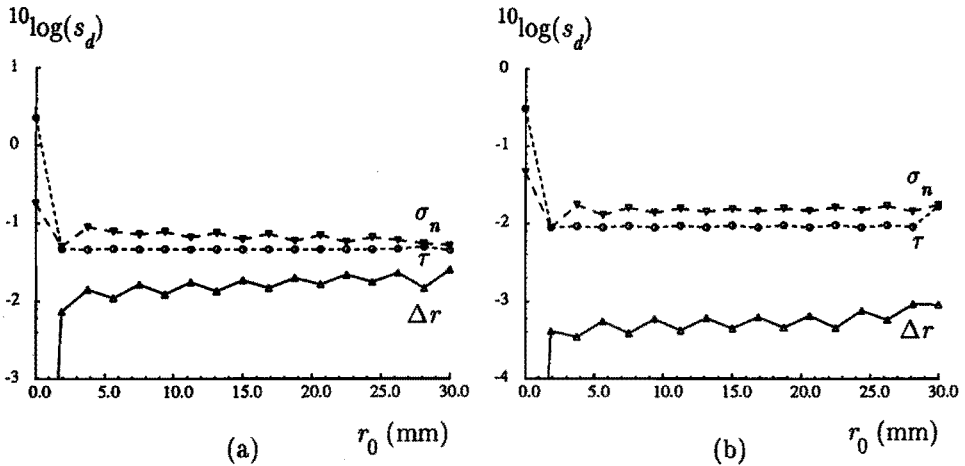
Figure 4.5.6 The dimensionless standard deviations for case 1

4 An experimental–numerical method

For each contact quantity s_d is defined as a function of the position at the contact surface by

$$s_d = \frac{\hat{s}}{\hat{\mu}_m} \qquad (4.5.2)$$

with $\hat{\mu}_m$ the maximum of the absolute estimation of the quantity along the contact surface, and $\hat{s}$ the locally estimated standard deviation. It can be interpreted as a relative error. The advantage of this definition is that the accuracy of the contact quantities, with mutually very different ranges, can be compared. The calculated values of s_d for the radial displacement, normal stress and friction stress after 3 mm and 30 mm upsetting are given in figure 4.5.6 as a function of the reference radial position.

Case 2   In the second case, the knowledge of the contact model is not taken into account. For compensation the measured axial and radial displacement of the outer surface with a standard deviation of 0.01 mm are used as boundary conditions. It is noticed that for each increment the measured displacements are referenced to the original state, so no accumulation of measuring errors occurs. This second problem shows to be very ill-conditioned. Due to this, lack of convergence occurred in the calculation when it was started for each increment from a first guess for the solution too far from the actual solution.
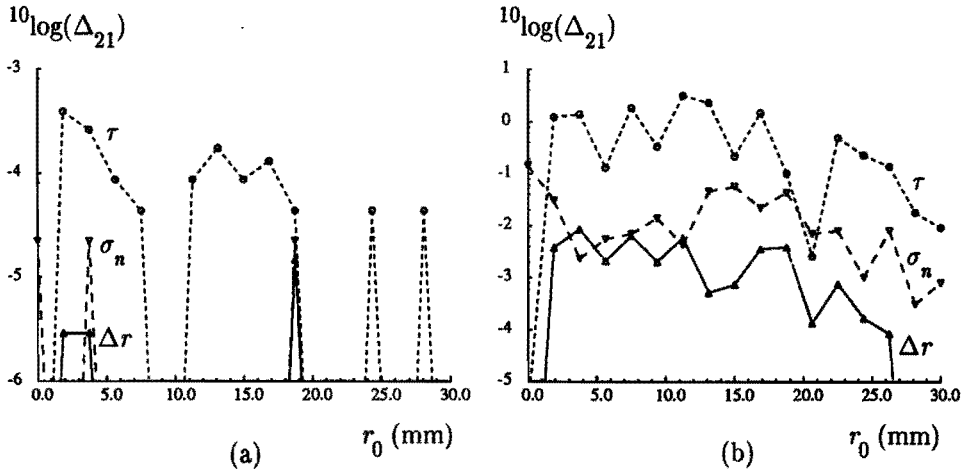
Figure 4.5.7 Dimensionless absolute differences of case 2 with respect to case 1

Therefore the incremental first guesses are taken equal to the incremental solutions of the direct problem of case 1. The solution then calculated shows differences (increasing per increment) from the solution of the direct problem. To illustrate this, figure 4.5.7 shows the dimensionless absolute differences Δ_{21} between the solution of the current case and the solution of case 1. The method to make the differences dimensionless is equal to the one used in equation (4.5.2). The differences are caused by the truncation of the displacements of the outer surface calculated for case 1, before they are used as input boundary condition in the current case. In fact these truncation errors act as very small measuring errors. The ill-conditioning is also illustrated by the resulting dimensionless standard deviations of the contact quantities, given in figure 4.5.8.
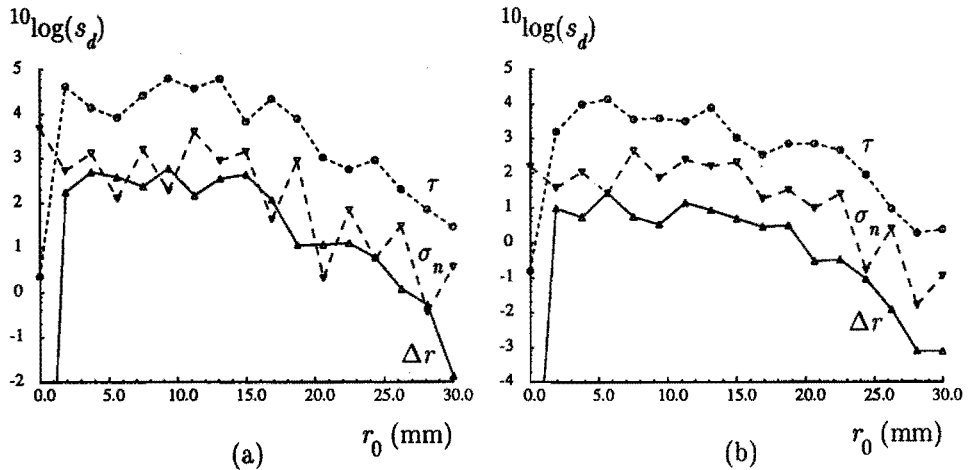
Figure 4.5.8 The dimensionless standard deviations for case 2

The accuracy of the measured displacements at ∂B_o should be increased beyond physical possibilities to obtain acceptable accurate contact quantities.

Case 3 With respect to case 2, the extra input of a measured upsetting force is used. It is supposed that this force (in the order of 0.1 MN) is very accurately known with a standard deviation of 100 N. To achieve convergence of the solution process the same procedure as for the previous case was necessary. The dimensionless standard deviations become smaller, see figure 4.5.9, however not enough to reach usefulness.
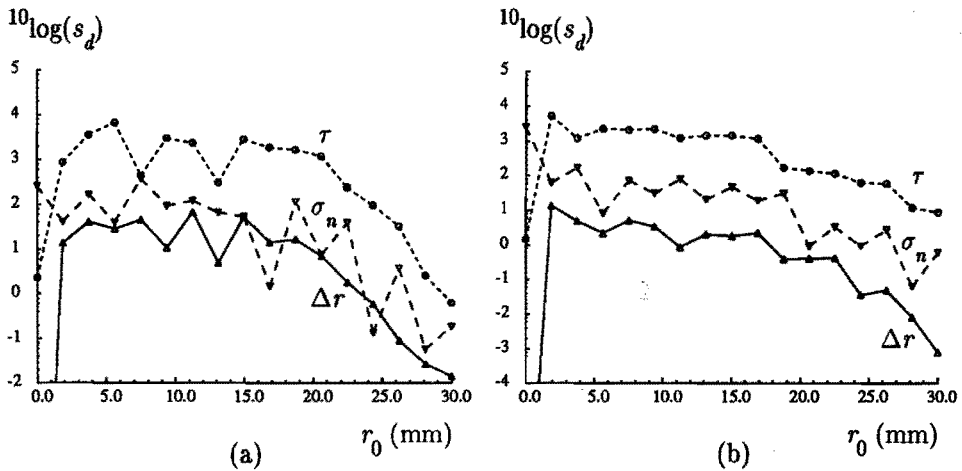
Figure 4.5.9 The dimensionless standard deviations for case 3

Case 4 Instead of the knowledge of the contact model of case 1, the measured normal stresses serve as boundary conditions at the contact surface. The standard deviation of this stress is very conservatively supposed to be 10 N/mm^2. The resulting dimensionless standard deviations are given in figure 4.5.10. The large value of s_d for the friction stress τ is caused by the small absolute values of τ.
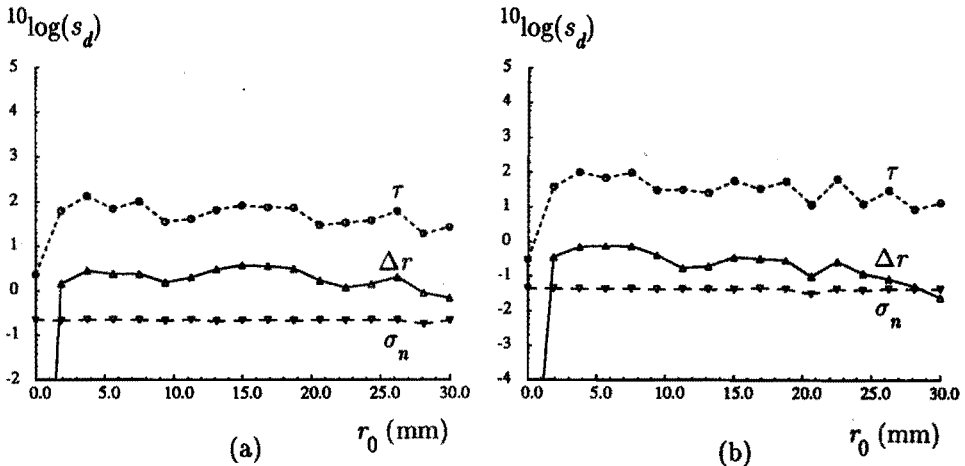
Figure 4.5.10 The dimensionless standard deviations for case 4

<u>Case 5</u>   In this case it is considered whether the estimates of case 4 can be improved by measured contour data. As in the cases 2 and 3, a standard deviation of 0.01 mm is supposed for the displacements of the contour points. In figure 4.5.11 improvements can be established, especially for the high upsetting range and near the outer radius.
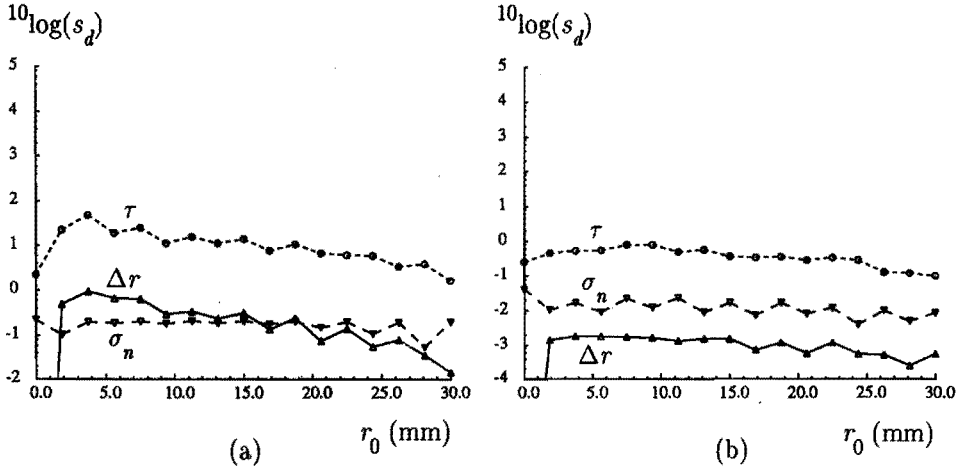
Figure 4.5.11 The dimensionless standard deviations for case 5

In contrast to the reliability of the radial displacement and normal stress, the reliability of the friction stress is unsatisfactory. This has to be improved before calculated results for the friction stress can be used for quantifying contact models. Three ways are mentioned to reach this improvement. At first the standard deviation of the normal stress is taken quite large in the above calculation. Measurement of the normal stress with a higher accuracy will reduce the standard deviation of the friction stress. Secondly use can be made of the smoothness of the contact stress pattern. Decrease of the number of contact stress unknowns will result in a higher accuracy for the contact quantities. This strategy is also applied for calculations with respect to the stress measuring tool in chapter 6. In that chapter it turns out that the friction stress can be measured together with the normal stress, thus offering a third way for raising the reliability of the friction stress. The fitting quality of the, with a reduced number of contact stress unknowns, estimated contact stress pattern can be judged by using the reached minimum value of $F(\underline{s})$ according to equation (4.4.3). A statistical test, resulting a quantitative measure for the fitting quality, is the χ^2-test, which is nearer described in paragraph 5.4.

From the previous it is concluded that the use of measured displacements of the contour, the upsetting displacement and the upsetting force only cannot lead to a successful estimate of the contact quantities. Measurement of the stresses at the contact area is necessary. Improvement of the results calculated from contact stress data may be obtained if additively quite accurate contour displacements are available. This is one of the reasons to consider displacement measurements in more detail. There is another reason why contour displacements remain of interest. An often used strategy to quantify the constant friction factor of the Von Mises model or the Coulomb friction factor, already indicated in chapter 3, is based on fitting the calculated upsetting force to measured values. In this way no indication is available whether the fitted model is indeed valid. This indication can be obtained by comparing the calculated contour displacements to the measured ones. Furthermore they give a possibility to extend this method of parameter quantification to models with more than one parameter. Nevertheless the main conclusion of this chapter is that a method to measure the contact stresses *in situ* in the contact region is of great importance. As an impulse to solve this problem a design of a contact stress measuring tool is presented and analysed in chapter 6. An important input variable for the application of this tool is the dimension of the contact area, which also results from the displacement measurement described in the next chapter.

5 Photogrammetric contour measurement

In this chapter the photogrammetric method, applied to measure the contour of workpieces during upsetting *in situ*, is described. For this purpose the workpieces are provided with marks at their cylindrical surface before upsetting. During upsetting single images are recorded of a workpiece together with a set of calibration marks. Image processing delivers the image coordinates of the centres of the calibration and workpiece marks. Calibration of the relationship between object space coordinates and image coordinates is executed for each image separately, using the calibration mark data. Although only single images are recorded, the three-dimensional space coordinates of the workpiece marks can be calculated supposing an axisymmetric contour shape.

In paragraph 5.1 the choice for a photogrammetric technique is motivated. The procedure roughly indicated above is outlined in more detail. The relationship between three-dimensional object space coordinates and two-dimensional image coordinates is subject of paragraph 5.2. In this relationship parameters appear whose values depend on the optical system. The method to calculate these parameters for each image, given the calibration mark data, is indicated. Paragraph 5.3 describes the reconstruction of the space coordinates of the workpiece marks from the image coordinates using the axisymmetry of the workpiece. Thereafter in paragraph 5.4 some aspects of the possible accuracy of the calibration and the three-dimensional coordinates reconstruction are indicated. A short description of the numerical implementation of the calibration and reconstruction algorithms is presented in paragraph 5.5. Results of simulations, also presented in paragraph 5.5, establish the usefulness of these algorithms. An evaluation of the method is performed in paragraph 5.6 by a set of upsetting experiments. It shows that the accuracy obtained is mainly limited by the violation of assumptions about the physical behaviour of the workpiece.

5.1 Introduction

One of the aims, formulated in chapter 3, is to determine during upsetting *in situ* the spatial positions of material points on the initially cylindrical surface of a workpiece, here shortly described as determining the contour of that workpiece. Some contour measurement techniques are mentioned here. Although with shape measuring techniques material surface points are not followed, some of these

methods are also reviewed in the sequel as far as they can easily be adopted for contour measurement. Direct and indirect methods for contour measurement can be distinguished. Direct methods make use of mechanical sensors, shadow images or photogrammetric position determination of surface marks. Indirect methods primarily concentrate on surface strain measurements. An example of a direct method using mechanical sensors is the measurement of the diameter at different locations on an axisymmetric upsetting test piece, as performed by Rasmussen et al. (1984). To follow material contour points an adaptation of the sensing elements is needed. A disadvantage is that only a very limited amount of data can be gathered this way, due to the voluminous mechanical construction. The determination of the radius in the necking zone of a tensile bar, as reported by Galenkamp & Van Wijngaarden (1985), is an example of a shadow image technique. Use of grooved workpieces enables point tracking. Both methods presuppose axisymmetry. A photogrammetric method was for example recently developed by Peters (1987) for the measurement of displacement fields in soft tissue. It appeared to be suited for large strains and displacements, characteristic also for the case considered here. With respect to the indirect measurement procedures, examples are Moiré, speckle methods (Kobayashi 1987) and use of strain gauges. Assuming axisymmetry, the displacements of the contour points may be calculated. A major drawback is the restriction to small strain cases.

Based on the previous considerations it is decided to modify Peters' strain measurement tool into a version able to determine the three-dimensional positions of marks on the workpiece. These positions are reconstructed from images recorded during upsetting. In general such a reconstruction is possible if at least two simultaneously recorded images from different positions are available or if the three-dimensional coordinates of the material points obey some known relationship. As it is supposed that an axisymmetric workpiece remains axisymmetric during upsetting, simultaneous recording of two or more images is not necessary. Therefore only single images are used to reconstruct the contour of the workpiece.

Before upsetting, the cylindrical surface of a test piece is provided with marks. During upsetting, images of the deforming test piece together with a calibration frame are recorded. On the calibration frame marks are located whose three-dimensional positions are accurately known with respect to some coordinate system. With the calibration marks the actual relationship between three-dimensional object coordinates and two-dimensional coordinates of the resulting image is determined. The reason to perform this calibration for each image separately is twofold. At first

the medium used for recording and analysing is photographic film. For several reasons, the image on the photo to be analysed may differ, partly in a stochastic manner, from the ideal image. Therefore for each photo a quantification of the distortion is needed. Secondly the camera is, because of practical reasons, not rigidly connected with the calibration frame which is attached to one of the upsetting tools. To avoid a laborious procedure in order to obtain a well-known camera position, the calibration mark data are also used to determine this position afterwards.

The photos are digitized by a scanner and the output data are processed by Peters' data analysing program, resulting in the two-dimensional image coordinates of the centres of calibration and workpiece marks. With the calibration mark image coordinates the camera position and image distortion are calculated for each photo. The workpiece mark data of the first photo of an upsetting experiment, with the workpiece in undeformed state, are used to identify material points on the cylindrical surface. For all subsequent photos the three-dimensional coordinates of the material points identified are reconstructed from the workpiece mark data, using axisymmetry.

5.2 Central projection, distortion model and calibration

In this paragraph the equations describing the transformation of object space coordinates to image coordinates are considered. At first this transformation is modelled by the theory of the central projection. Modifications are needed to account for various systematic errors. Finally the calibration procedure with the resulting equations is discussed.
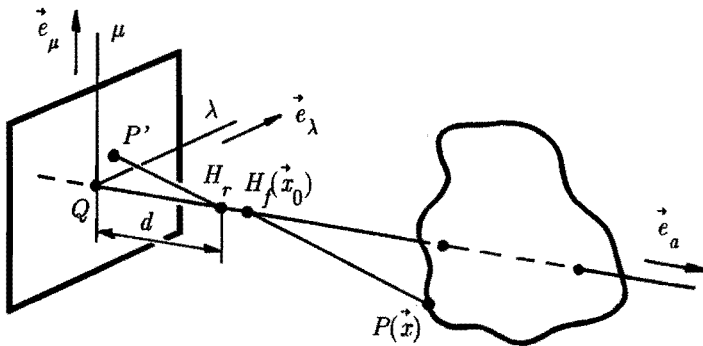
Figure 5.2.1 The transformation of the object space to the film image plane

Using a camera with a set of linear lenses, the transformation of a point $P(\vec{x})$ in the three-dimensional object space into a point P' in the two-dimensional film image plane is visualized in figure 5.2.1 (Arnold et al. 1971). It is assumed that the film image plane is flat and perpendicular to the so-called optical axis with unity direction vector $\vec{e}_a$. On the optical axis the principal nodal points H_f and H_r are situated, H_f the front lens node with position vector $\vec{x}_0$ and H_r the rear lens node. The connection lines PH_f and H_rP' are parallel. Within the film image plane an orthonormal two-dimensional vector base $\{\vec{e}_\lambda, \vec{e}_\mu\}$ and distance measuring coordinates (λ, μ) are introduced. The centre of projection or principal focal point Q, i.e. the intersection point of the film image plane and the optical axis, is arbitrarily chosen as the origin of the film plane coordinate system. The coordinates (λ, μ) of P' are related to the position vector $\vec{x}$ of P by

$$\lambda = -d\frac{\vec{e}_\lambda \cdot (\vec{x}-\vec{x}_0)}{\vec{e}_a \cdot (\vec{x}-\vec{x}_0)} \quad , \quad \mu = -d\frac{\vec{e}_\mu \cdot (\vec{x}-\vec{x}_0)}{\vec{e}_a \cdot (\vec{x}-\vec{x}_0)} \tag{5.2.1}$$

with d the principal distance, i.e. the distance of H_r to the film image plane. Prints of the recorded film images are analysed by use of a scanner, which assigns distance measuring coordinates (p,q) to points of the print. The unity vectors $\vec{e}_p$ and $\vec{e}_q$ parallel to the coordinate axes of the scanner may be imagined to coincide with $\vec{e}_\lambda$ and $\vec{e}_\mu$, thus $\vec{e}_\lambda = \vec{e}_p$ and $\vec{e}_\mu = \vec{e}_q$. The coordinates p and q are assumed to be linear in λ and μ according to $p = p_0 + m\lambda/d$ and $q = q_0 + m\mu/d$ where m is a magnification factor and (p_0, q_0) are the print image coordinates of the centre of projection, so

$$p = p_0 - m\frac{\vec{e}_p \cdot (\vec{x}-\vec{x}_0)}{\vec{e}_a \cdot (\vec{x}-\vec{x}_0)} \quad , \quad q = q_0 - m\frac{\vec{e}_q \cdot (\vec{x}-\vec{x}_0)}{\vec{e}_a \cdot (\vec{x}-\vec{x}_0)} \tag{5.2.2}$$

As the vectors $\vec{e}_p$, $\vec{e}_q$ and $\vec{e}_a$ compose an orthonormal base, the relations (5.2.2) contain 9 independent parameters.

In reality the equations (5.2.2) will not be exactly obeyed, due to different systematic and stochastic error sources. Systematic errors originate from lens aberration, film and print deformation and scanner deficiencies. To correct for such

errors a method closely related to the one described by Peters (1987) is applied. It is based on the use of isoparametric coordinates (Zienkiewicz 1977, Bathe 1982). With n calibration marks as nodal points, the coordinates (p,q) are transformed into isoparametric coordinates (ξ,η) using shape functions $\varphi_i(\xi,\eta)$ and the coordinates (p_i,q_i) of the corresponding calibration marks

$$p = \sum_{i=1}^{n} \varphi_i(\xi,\eta)p_i \quad , \quad q = \sum_{i=1}^{n} \varphi_i(\xi,\eta)q_i \tag{5.2.3}$$

In this way for each point (p,q) within the area considered unique isoparametric coordinates (ξ,η) are defined. With respect to the relations (5.2.2) it is assumed that these can be adopted for systematic errors as follows

$$p+\Delta p = p_0 - m\frac{\vec{e}_p \cdot (\vec{x}-\vec{x}_0)}{\vec{e}_a \cdot (\vec{x}-\vec{x}_0)} \quad , \quad q+\Delta q = q_0 - m\frac{\vec{e}_q \cdot (\vec{x}-\vec{x}_0)}{\vec{e}_a \cdot (\vec{x}-\vec{x}_0)} \tag{5.2.4}$$

where the correction functions Δp and Δq obey relationships equivalent to (5.2.3)

$$\Delta p = \sum_{i=1}^{n} \varphi_i(\xi,\eta)\Delta p_i \quad , \quad \Delta q = \sum_{i=1}^{n} \varphi_i(\xi,\eta)\Delta q_i \tag{5.2.5}$$

in which Δp_i and Δq_i are the coordinate corrections for calibration mark i. Because with respect to the image coordinates (p,q) the corrections Δp and Δq can describe arbitrary translation, rotation and magnification, which are already represented by other parameters, four (independent) values for Δp and Δq may be prescribed. Improvement of error corrections can be obtained by increasing the number n of calibration marks.

By introduction of the correction functions, the number of independent parameters in (5.2.4) is increased with respect to (5.2.2) from 9 to $9+2n-4$. Together they compose the parameter column γ. To determine γ for a particular image at least 3 extra calibration marks are needed besides the number (n) to quantify the systematic error correction. Extraction of the parameter values from the calibration mark data is performed by a maximum likelihood estimator. Therefore it is supposed that the print image coordinates (z_p, z_q) of the centre of some (calibration or workpiece) mark, as measured by the scanning device, obey the relationships

$$z_p = p + \Delta p + v_p \quad , \quad z_q = q + \Delta q + v_q \tag{5.2.6}$$

with $p+\Delta p$ and $q+\Delta q$ according to (5.2.4) and (5.2.5), and stochastic errors v_p and v_q. Some of the properties of these stochastic errors, due to the discrete nature of the scanning device, are investigated by Peters (1987). For (5.2.6) it is assumed that the centre of a mark on the print corresponds with the centre of that mark in the object space. Errors introduced by this assumption are subject of paragraph 5.4. The measured scanner coordinates of the centres of the calibration marks compose the column $\underset{\sim}{z}_c$. With the foregoing it can be stated, supposing that the positions of the centres of the calibration marks in the object space are exactly known, that

$$\underset{\sim}{z}_c = \underset{\sim}{h}_c(\gamma) + \underset{\sim}{v}_c \tag{5.2.7}$$

with $\underset{\sim}{v}_c$ a column containing stochastic errors and $\underset{\sim}{h}_c$ a column function in which only the column γ occurs as unknown. The stochastic errors $\underset{\sim}{v}_c$ are supposed to have a (multivariate) Gaussian probability density function p according to

$$p(\underset{\sim}{v}_c) = \left((2\pi)^{2n_c} \det(\underline{R}_c) \right)^{-\frac{1}{2}} \exp(-\frac{1}{2} \underset{\sim}{v}_c^T \underline{R}_c^{-1} \underset{\sim}{v}_c) \tag{5.2.8}$$

with n_c the total number of calibration marks and $E(\underset{\sim}{v}_c) = \underset{\sim}{0}$, $E(\underset{\sim}{v}_c \underset{\sim}{v}_c^T) = \underline{R}_c$ where E is the expected value operator. Replacing $\underset{\sim}{v}_c$ by $\underset{\sim}{z}_c - \underset{\sim}{h}_c(\gamma)$ in (5.2.8), the likelihood function $p(\underset{\sim}{z}_c; \gamma)$ can be obtained. The maximum likelihood estimate $\hat{\gamma}$ is defined as the solution γ, that maximizes $p(\underset{\sim}{z}_c; \gamma)$ for the particular set of measurement data $\underset{\sim}{z}_c$, or equivalently, minimizes the function

$$F(\gamma) = \left(\underset{\sim}{z}_c - \underset{\sim}{h}_c(\gamma) \right)^T \underline{R}_c^{-1} \left(\underset{\sim}{z}_c - \underset{\sim}{h}_c(\gamma) \right) \tag{5.2.9}$$

As $\underset{\sim}{h}_c(\gamma)$ is non-linear in γ, an iterative solution procedure is needed to calculate $\hat{\gamma}$. Some details of the numerical implementation of this procedure are given in paragraph 5.5.

5.3 Reconstruction of the three-dimensional coordinates of the workpiece marks

Subject of this paragraph is the determination of the contour of a workpiece from the image coordinates of the marks. For this purpose first the positions of the marks on the undeformed workpiece have to be established. In some deformed state the three-dimensional positions of these marks are reconstructed, supposing preservation of axisymmetry. An approximation for the position of an arbitrary surface point is obtained by a special interpolation using orthogonal polynomials.

For the experiments described in paragraph 5.6, it is chosen to provide the workpiece with marks by hand, using a stencil and felt-tipped pen as marking medium. This method is easy to perform and the marks pattern wanted can be changed in a simple way. The resulting pattern on a workpiece will however show irregularities with respect to the ideal pattern according to the stencil. In order to correct for these irregularities the image of the undeformed state of each workpiece, recorded just before upsetting, is employed. The position of the centre of such a workpiece in a coordinate system attached to the calibration frame (origin 0, orthonormal base $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$) is indicated by the position vector $\vec{s}_0$, see figure 5.3.1.
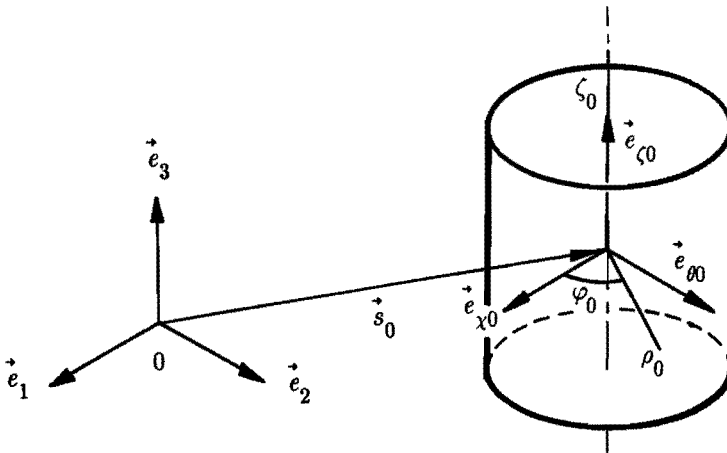
Figure 5.3.1 A workpiece in undeformed state just before upsetting

The orientation of the workpiece is described by a material fixed orthonormal base $\{\vec{e}_{\chi 0}, \vec{e}_{\theta 0}, \vec{e}_{\zeta 0}\}$ with origin at $\vec{s}_0$. Material points can be uniquely identified by the cylindrical coordinates $(\rho_0, \varphi_0, \zeta_0)$. The position of the centre of each mark i on the workpiece is given by $(R_0, \varphi_0^i, \zeta_0^i)$, as for all marks the radial position is equal to the

initial workpiece radius R_0. The coordinates φ_0^i and ζ_0^i for the centre of mark i are considered as a label to identify the mark and are called the material coordinates of mark i. The position vector $\vec{x}_0^i$ of a mark i can be denoted as

$$\vec{x}_0^i = \vec{s}_0 + \zeta_0^i \vec{e}_{\zeta 0} + R_0 \left(\cos(\varphi_0^i) \vec{e}_{\chi 0} + \sin(\varphi_0^i) \vec{e}_{\theta 0} \right) \tag{5.3.1}$$

The intended axial and circumferential coordinates $(z_{\varphi 0}^i, z_{\zeta 0}^i)$ of the marks on the stencil in general differ from the real values (φ_0^i, ζ_0^i) on the object. The relations can be denoted by

$$z_{\varphi 0}^i = \varphi_0^i + v_{\varphi 0}^i \quad , \quad z_{\zeta 0}^i = \zeta_0^i + v_{\zeta 0}^i \tag{5.3.2}$$

with stochastic errors $v_{\varphi 0}^i$ and $v_{\zeta 0}^i$. The equations (5.3.2) for the intended coordinates of all workpiece marks are written as

$$\underset{\sim}{z}_{w0} = \underset{\sim}{\delta}_{w0} + \underset{\sim}{v}_{w0} \tag{5.3.3}$$

with the real coordinates (φ_0^i, ζ_0^i) gathered in the column $\underset{\sim}{\delta}_{w0}$.

The measured image coordinates (z_p^i, z_q^i) of mark i will obey the relationships

$$z_p^i = p^i + \Delta p^i + v_p^i \quad ; \quad z_q^i = q^i + \Delta q^i + v_q^i \tag{5.3.4}$$

with $p^i + \Delta p^i$, $q^i + \Delta q^i$ according to (5.2.4) and (5.2.5), and stochastic errors v_p^i, v_q^i. Applying (5.2.4) in (5.3.4) the position vector $\vec{x}$ has to be replaced by the position vector $\vec{x}_0^i$ of mark i according to (5.3.1). The relations (5.3.4) for the image coordinates of all workpiece marks together are concisely denoted as

$$\underset{\sim}{z}_{i0} = \underset{\sim}{h}_{i0}(\underset{\sim}{\gamma}, \underset{\sim}{\delta}) + \underset{\sim}{v}_{i0} \tag{5.3.5}$$

The functions in the column $\underset{\sim}{h}_{i0}$ are non-linear in the calibration parameters $\underset{\sim}{\gamma}$ and the column $\underset{\sim}{\delta}$, which is composed of the real coordinates $\underset{\sim}{\delta}_{w0}$ and six variables for the position and orientation of the workpiece. Supposing that the calibration parameters

γ are exactly known, with (5.3.3) and (5.3.5) a maximum likelihood estimate $\hat{\underline{\delta}}$ for $\underline{\delta}$ can be derived. Therefore the stochastic errors are assumed to have a Gaussian probability density function

$$p\left(\begin{bmatrix} \underline{v}_{w0} \\ \underline{v}_{i0} \end{bmatrix}\right) = \left((2\pi)^{4n_m}\det(\underline{R}_{w0})\det(\underline{R}_{i0})\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\underline{v}_{w0}^T\underline{R}_{w0}^{-1}\underline{v}_{w0} - \frac{1}{2}\underline{v}_{i0}^T\underline{R}_{i0}^{-1}\underline{v}_{i0}\right)$$ (5.3.6)

with n_m the number of workpiece marks, $E(\underline{v}_{w0}) = E(\underline{v}_{i0}) = \underline{0}$, $E(\underline{v}_{w0}\underline{v}_{w0}^T) = \underline{R}_{w0}$ and $E(\underline{v}_{i0}\underline{v}_{i0}^T) = \underline{R}_{i0}$. It is obvious that the stochastic errors $\underline{v}_{w0}$ and $\underline{v}_{i0}$ are independent, so $E(\underline{v}_{w0}\underline{v}_{i0}^T) = \underline{0}$. The likelihood function $p\left([\underline{z}_{w0}^T \ \underline{z}_{i0}^T]^T:\underline{\delta}\right)$ follows from (5.3.6) by replacing $\underline{v}_{w0}$ and $\underline{v}_{i0}$ by $\underline{z}_{w0}-\underline{\delta}_{w0}(\underline{\delta})$ and $\underline{z}_{i0}-\underline{h}_{i0}(\gamma,\underline{\delta})$ respectively. Maximization of this function is identical to minimizing the function

$$G(\underline{\delta}) = \left(\underline{z}_{w0}-\underline{\delta}_{w0}(\underline{\delta})\right)^T\underline{R}_{w0}^{-1}\left(\underline{z}_{w0}-\underline{\delta}_{w0}(\underline{\delta})\right) + \left(\underline{z}_{i0}-\underline{h}_{i0}(\gamma,\underline{\delta})\right)^T\underline{R}_{i0}^{-1}\left(\underline{z}_{i0}-\underline{h}_{i0}(\gamma,\underline{\delta})\right)$$ (5.3.7)

and delivers the maximum likelihood estimate $\hat{\underline{\delta}}$. An iterative solution procedure is needed to calculate $\hat{\underline{\delta}}$, because of non-linearities.

After evaluation of the calibration parameters and the real coordinates (φ_0^i,ζ_0^i) of the workpiece marks in the undeformed state, the remaining problem is to calculate the object space positions of the centres of the marks for the workpiece in some deformed state. If the deformation of a mark i itself is more or less uniform, its centre may be assigned to the same material surface point with material coordinates (φ_0^i,ζ_0^i) in each state. Thus the actual contour is determined by the object space positions of the centres of all marks. To reconstruct the three-dimensional position for each workpiece mark i the measured image positions (z_p^i,z_q^i) according to (5.3.4) are available. Supposing the calibration parameters and coordinates (φ_0^i,ζ_0^i) of all workpiece marks to be known, there are only $2n_m$ measurement data for $3n_m$ unknowns. Preservation of axisymmetry has to be used to calculate the unknowns. In some deformed state the position and orientation of the workpiece in the object space is described by a vector $\vec{s}$ and the orthonormal material base $\{\vec{e}_\chi,\vec{e}_\theta,\vec{e}_\zeta\}$ with its origin at $\vec{s}$, see figure 5.3.2. The vector $\vec{s}$ is the position vector of the intersection

     5 Photogrammetric contour measurement

point of the centre line of the workpiece with the plane through the surface points with material coordinate $\zeta_0 = 0$. This plane is not assumed to be a symmetry plane a priori. The centre of mark i will have the position $\vec{x}^i$ according to

$$\vec{x}^i = \vec{s} + \zeta^i \vec{e}_\zeta + \rho^i\left(\cos(\varphi^i)\vec{e}_\chi + \sin(\varphi^i)\vec{e}_\theta\right) \tag{5.3.8}$$

with $(\rho^i, \varphi^i, \zeta^i)$ the cylindrical coordinates of mark i.
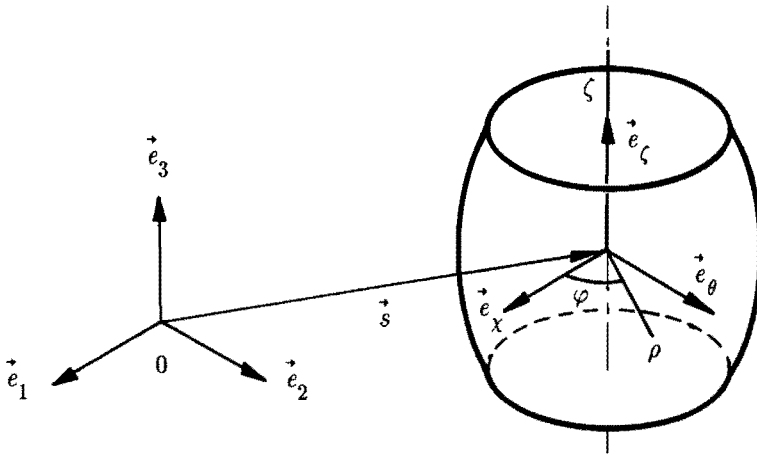
Figure 5.3.2 A workpiece in some deformed state

Because of axisymmetry it may be written for $(\rho^i, \varphi^i, \zeta^i)$

$$\rho^i = R(\zeta_0^i) \quad , \quad \varphi^i = \varphi_0^i \quad , \quad \zeta^i = Z(\zeta_0^i) \tag{5.3.9}$$

with R and Z functions of ζ_0. These functions can be discretized by a set of functions with corresponding parameters. The variables to be determined are the position and orientation of the workpiece and $n_R + n_Z$ parameters for the interpolation of R and Z. These variables however are not independent. The workpiece imaginary shifted in the direction of the optical axis and proportionally changed in its dimensions, results in the same image. Therefore an extra assumption with respect to the position or dimensions of the real workpiece is needed. Here it is assumed that the workpiece is volume invariant which introduces a dependency between the parameters of the approximations for R and Z. A set of $n_R + n_Z - 1$ parameters can easily be introduced

by using Legendre polynomials as is described in appendix 5.3.1. To determine the remaining variables, gathered in the column $\underset{\sim}{\epsilon}$, again the maximum likelihood method is adapted. With (5.3.4) the measured image coordinates of the workpiece marks $\underset{\sim}{z}_m$ can be expressed as

$$\underset{\sim}{z}_m = \underset{\sim}{h}_m(\underset{\sim}{\epsilon}) + \underset{\sim}{v}_m \tag{5.3.10}$$

Supposing a Gaussian probability density function for $\underset{\sim}{v}_m$ with zero mean and covariance matrix $\underline{R}_m$, delivers the maximum likelihood estimate $\hat{\underset{\sim}{\epsilon}}$, i.e. column $\underset{\sim}{\epsilon}$ that minimizes

$$H(\underset{\sim}{\epsilon}) = \left(\underset{\sim}{z}_m - \underset{\sim}{h}_m(\underset{\sim}{\epsilon})\right)^T \underline{R}_m^{-1}\left(\underset{\sim}{z}_m - \underset{\sim}{h}_m(\underset{\sim}{\epsilon})\right) \tag{5.3.11}$$

The solution is calculated iteratively.

5.4 Aspects of accuracy

In this paragraph attention is focussed on the obtainable accuracy of calibration and three-dimensional coordinate reconstruction. The influence of the stochastic errors is studied by numerical simulations, presented in paragraph 5.5. A few statistical quantities used in that paragraph are reviewed here. Thereafter the influence of systematic measurement errors, neglected in the previous, is estimated. They show to be negligibly small. Finally systematic errors due to the assumption of volume invariance are considered.

One of the statistic quantities considered in the next paragraph is the covariance matrix of the resulting estimates. Approximations for the covariance matrix of the maximum likelihood estimates $\hat{\gamma}$, $\hat{\underset{\sim}{\delta}}$ and $\hat{\underset{\sim}{\epsilon}}$ can be given as indicated in chapter 4 and appendix 4.3.1. Further it is known from statistics that (see e.g. Sokal & Rohlf 1969), supposing Gaussian probability density functions for the measured data, the functions $F(\hat{\gamma})$, $G(\hat{\underset{\sim}{\delta}})$ and $H(\hat{\underset{\sim}{\epsilon}})$ are χ^2-distributed. The resulting minimum values can therefore be used to determine the fitting quality of the measured data with respect to the resulting estimated model. Here the fitting quality is defined by the probability P_F that a particular χ^2 value as finally reached, could occur by chance. Values larger than 0.1 make the fit reliable. Values larger than 0.001 may be

acceptable if the errors are non-normal distributed or underestimated, or if the data show outlier points. An extra statistical method used to check the results is the Kolmogorov–Smirnov test to compare the assumed Gaussian distributions with the resulting estimated data. The test delivers the probability P_{KS} that differences between the assumed distribution and the estimated data larger than the observed differences occur by chance. As significance level a value of about 0.01 is reasonable. For a small data set P_{KS} becomes meaningless.

For the positions of the calibration marks, systematic errors, neglected in the previous, may occur due to temperature changes of the calibration frame. The frame used in the experiments is constructed of aluminium with a thermal coefficient of linear expansion of about 24 μm/mK. A typical dimension equals 100 mm. Thus for a temperature change of 1 K, a maximum error is introduced of about 2.4 μm. Although this is quite small, heating of the frame, due to the presence of light sources needed for the image recording, has to be avoided. Intermitted lighting and the connection of the frame to one of the tools contribute to this purpose.

Some of the systematic errors in the three-dimensional positions of the workpiece marks are invoked by oblique orientation with respect to the optical axis, the surface curvature and the change of surface strain over the marks. The magnitude of these errors can be estimated from simple considerations of geometrical nature. Without going into details the results of these considerations will be presented.

For a flat non-deforming workpiece mark with its normal direction parallel to the optical axis, the centre of the image of the mark will coincide with the image of the centre of the mark in the object space, if photogrammetric distortional errors are not present. If the angle between the normal direction of the mark and the optical axis is unequal zero, the centre of the image of the mark will not correspond to the image of the centre in the object space. It can be proved that for the distance deviations Δ_o, referred to the object geometry, holds

$$\Delta_o < 0.2 \frac{d^2}{l} \quad (\frac{d}{l} << 1) \tag{5.4.1}$$

with d a significant dimension of the mark and l the distance of the mark to the front lens node. With e.g. $d = 2$ mm and $l = 1000$ mm the error will be smaller than 1 μm, which is extremely small with respect to the influence of stochastic errors. If besides the oblique orientation the surface is curved with curvature radius $\overline{R}$, an extra shift Δ_c may occur, satisfying

$$\Delta_c < 0.1 \frac{d^2}{\overline{R}} \tag{5.4.2}$$

With $d = 2$ mm and $\overline{R} = 30$ mm, Δ_c becomes 13 μm. As this maximum value only occurs for a few marks or even no mark at all on a particular image, this error can also be neglected. If a non-uniform surface strain with characteristic value g_ϵ for the gradient is present for a flat mark, an error Δ_s

$$\Delta_s < 0.1 \, d^2 g_\epsilon \tag{5.4.3}$$

is introduced. So for $d = 2$ mm and $g_\epsilon < 0.002$ mm^{-1} this error is smaller than 1 μm. In principle these three errors can be corrected for, as the obliqueness, curvature and strain are approximately known during the calculation of the contour. However, because the errors are small, they are simply neglected.

Finally systematic errors related to the use of volume invariance are considered. Due to elastic deformation of the workpiece, the volume will change approximately by an amount $\frac{1}{4}\pi D^2 h \frac{p}{\kappa}$ with D diameter, h height, p the average hydrostatic pressure and κ the compression modulus. The reconstruction algorithm will estimate the workpiece further away from the camera than in reality is the case. From this an overestimate of the height and diameter will result. With $p = 40$ N/mm^2, $D = 85$ mm, $h = 30$ mm and $\kappa = 70000$ N/mm^2 a maximum radius error of about 5 μm occurs. A comparable error is caused by neglecting the elastic indentation of the tools.

5.5 Numerical implementation and simulations

In this paragraph shortly the optimization method, used in the implementation of the calibration and reconstruction estimators, is considered. Numerical simulations establish the usefulness of the estimators and the attainable accuracy. It is shown that this accuracy is not sufficient to improve the results of the calculations according to chapter 4. The next paragraph will show that also the behaviour of the test pieces interferes with the calculation of more reliable results this way.

To perform the optimizations as formulated in the paragraphs 5.2 and 5.3 a routine (E04GBF) from the NAG-library (1987) is used. This routine combines the

Newton method with a quasi–Newton method for the unconstrained minimization of non-linear functions $Q(\underline{s})$ in the unknowns $\underline{s}$ of the type

$$Q(\underline{s}) = \underline{q}(\underline{s})^T \underline{q}(\underline{s}) \qquad (5.5.1)$$

Besides a subroutine for the evaluation of $\underline{q}(\underline{s})$ for some $\underline{s}$, also a subroutine for the calculation of the Jacobian matrix of $\underline{q}(\underline{s})$ is expected by the NAG-routine. The functions to be minimized, presented in the paragraphs mentioned above, can easily be reformulated according to (5.5.1) and the Jacobian matrices can be derived analytically.

To evaluate the estimators a Monte Carlo strategy is adapted, i.e. fictitious values are assumed for the non-stochastic quantities somehow close to the values occurring in an experiment. These quantities are
— calibration mark positions in the object space,
— camera position and orientation,
— magnification factor,
— print image coordinates for the principal focal point,
— distortion errors for all images within one experiment,
— stencil pattern,
— position and orientation of the workpiece,
— contour development.
Next all stochastic errors are simulated by drawing random numbers from appropriate distributions. Thus random Gaussian deviations are generated for
— scanning errors in the calibration mark image coordinates,
— scanning errors in the workpiece mark image coordinates,
— workpiece mark pattern errors.
This results in simulated sets of image coordinate data for the marks. With these data the calibration and reconstruction are performed and the results for the geometry are compared to the known original data. This procedure is executed a number of times. The acquired results are supposed to be representative for the general behaviour.

In figure 5.5.1 the mutual positions of the different parts of the set-up are schematically indicated along the optical axis. Point Q is the principal focal point on the film, H_r and H_f the rear and front lens nodes. The principal distance d has a value of about 216 mm. The calibration frame has calibration marks in two parallel planes with a mutual distance d_{fr} of 20 mm.
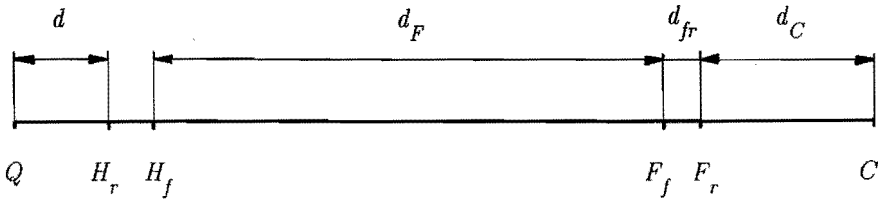
Figure 5.5.1 Some positions and distances of importance

The 16 marks of the front plane at position F_f are used for the error corrections according to paragraph 5.2, the rear plane at position F_r contains 6 extra marks. The distance d_F of the front calibration plane to the front lens node is about 1200 mm, the distance d_C of the workpiece center C to the rear calibration plane equals approximately 75 mm. All distances mentioned resemble the actual values in the set-up used for the experiments of the next paragraph. In figure 5.5.2 two typical pictures of simulated scanned image coordinates are given. In picture (a) the workpiece is in the reference state and in picture (b) the workpiece is upsetted until the height approximately equals 20 mm. A magnification factor m, see paragraph 5.2, of 1150 is used. The camera is oriented such that the optical axis is approximately perpendicular to the calibration planes and through the centre of the front calibration plane marks (Δ). The values of the image distortional errors are taken smaller than 0.005 mm.
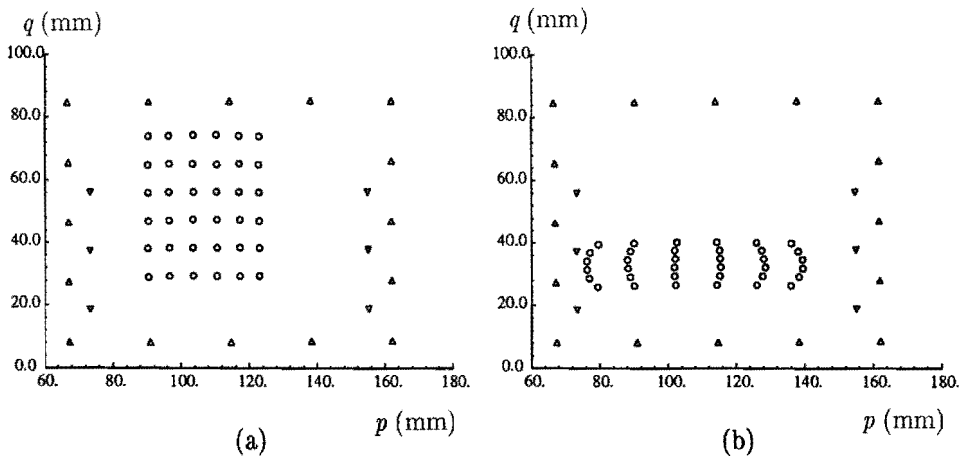
Figure 5.5.2 Image coordinates of calibration (Δ,∇) and workpiece ($\circ$) marks

In the initial state the workpiece is provided with 6×6 marks, in axial direction in rows with mutual distance 10 mm and in tangential direction in equidistant columns over a total angle of 90°. The standard deviation of the row and column distance is conservatively chosen 0.1 mm. To estimate the standard deviation of the scanned image coordinates to be expected, a formula of Peters (1987) is used which expresses the variance (in mm^2) in the scan resolution p_{sc} (in pixels/mm) and a characteristic mark image size d (mm)

$$s^2_{sc} \approx (22p^3_{sc}d)^{-1} \tag{5.5.2}$$

The scan device applied for the experiments has a value for p equal to 9.45 pixels/mm. For circular calibration marks with an image diameter of 3 mm a standard deviation of 0.004 mm results while this deviation is 0.006 mm for circular workpiece marks with an image diameter of 1.5 mm.

With these data simulations are performed of the upsetting of the workpiece to a final height of 20 mm. A typical (symmetric) example of a contour development which can be described by the interpolation for R and Z with Legendre polynomials according to appendix 5.3.1, is given in figure 5.5.3.
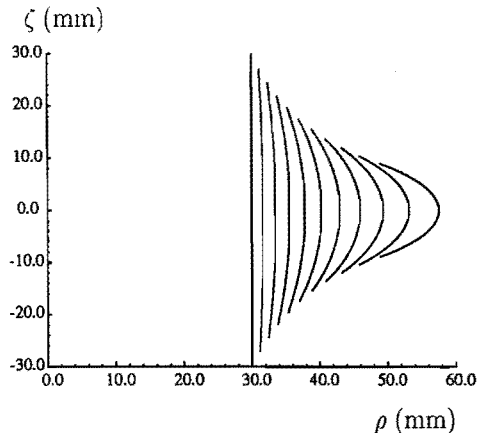
Figure 5.5.3 An example of a contour development described by the volume invariant interpolation using Legendre polynomials

The results obtained with respect to the calibration can be summarized as follows. In

general the 9 parameters figuring in equation (5.2.2), i.e. camera position and orientation, coordinates of the principal focal point (on the image) and magnification factor, are quite well estimated. The distance d_F of the front lens node to the front calibration plane is estimated within 5 mm, for the magnification factor a maximum absolute error about 5 is obtained. Rotational errors of the camera are within 0.05°, errors in the position of the principal focal point on the image within 2 mm. These errors are partly compensated by the correction factors Δp^i and Δq^i, which completely differ from their original values. Therefore the resulting minimum function value $F(\hat{\gamma})$ is only seldomly disappointing, as $P_F(F(\hat{\gamma}))$ is normally greater than 0.01. Also the normality of the resulting estimated image coordinates is only rarely rejected by a P_{KS} value smaller than 0.01.

The position and orientation of the workpiece in the reference state are within 2 mm respectively 0.02° correctly estimated. The material coordinates are estimated within 0.04 mm close to the original values. The reached minimum value $G(\hat{\delta})$ is satisfying, as well as the normality of the material coordinates. The momentary position and orientation are of the same accuracy as in the reference calculation.
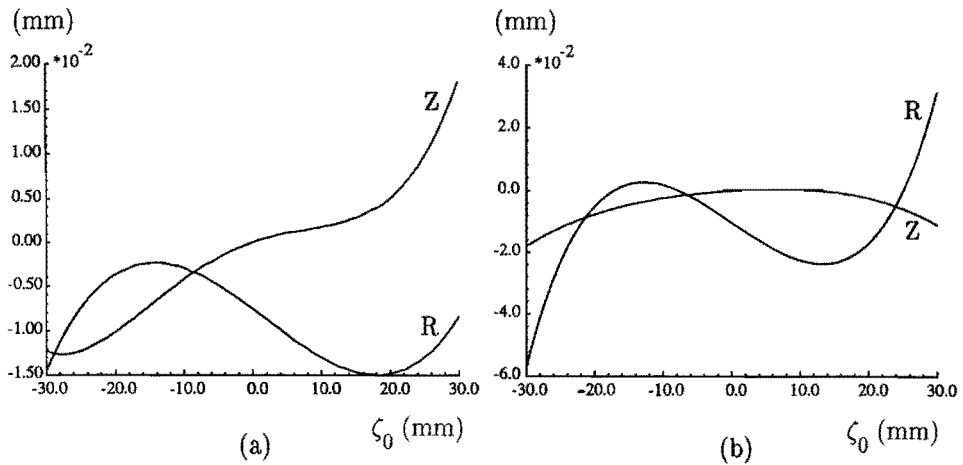
Figure 5.5.4 Axial (Z) and radial (R) displacement errors for the low (a) and high (b) upsetting range

In figure 5.5.4 an example of the axial and radial coordinate errors is given as function of ζ_0, for the low and high upsetting range.

For the low upsetting range axial and radial coordinate errors typically are about

0.01 mm and about 0.03 mm near $\zeta_0 = 30$ mm and $\zeta_0 = -30$ mm. For the high upsetting range the axial and radial coordinate errors become typically about 0.02 mm and 0.06 mm near the top and bottom of the workpiece. In all cases the reached $H(\hat{\epsilon})$ values give a P_F smaller than 0.001. Also the normality checks on the calculated image coordinates regularly deliver a P_{KS} smaller than 0.001. This is caused by the supposition of exactly known calibration parameters and material coordinates of the workpiece marks in the reconstruction of the momentary axial and radial coordinates. Nevertheless the displacement accuracy is not disappointing. The results however are too inaccurate to improve significantly the quality of the calculated contact quantities according to chapter 4. The next paragraph will show that a more accurate contour measurement cannot be performed by the present set-up, due to the physical phenomenon of surface roughening and the violation of axisymmetry resulting from inhomogeneous deformation.

5.6 Evaluation of the photogrammetric contour measurement

In this paragraph the photogrammetric method is evaluated by a number of upsetting experiments. It appears that the predicted accuracy is not reached due to several reasons. An improvement of the accuracy is possible, but is hardly relevant for the quantification of contact models according to chapter 4, as the obtainable accuracy is limited by surface roughening. The dimension of the contact area can be measured accurately enough to apply it in calculations with the stress measuring tool, presented in chapter 6.

A number of upsetting experiments is performed on soft annealed extruded aluminium Al 99.0 (AA 1050) workpieces with an initial diameter and height of 60 mm. The aluminium had an initial hardness of 21.5 N/mm^2 HB 5 and an average grain size of 0.125 mm^2 (ASTM 0). After a last fine turning process, the workpieces were sandblasted. The resulting egg-shell surface finish sufficed to prevent inconvenient light reflections on the photographs. Before upsetting, the workpieces were provided with 12×12 black marks of 2 mm diameter each. The axisymmetric upsetting tools (tool steel X210Cr12) had a hardness of 56 HRC. After fine grinding an average surface roughness value R_a resulted of approximately 0.2 μm. A 2 mm thick circular plate, of equal material and surface finish as the tools and with a hole of 3 mm diameter at the centre, was glued to one of the tools using a thin epoxy resin adhesive layer of about 5 μm thickness. A small pin in the hole centred the

workpieces during upsetting. The workpieces are upsetted on a hydraulic press in 10 steps of about 3 mm upsetting each to a final height of 30 mm. The average upsetting speed was 0.02 mm/s. After each step upsetting was interrupted in order to take a photograph of the calibration frame and workpiece, and to measure the upsetting force and the relative displacement of the tools. The displacement was registrated by an LVDT (linear variable differential transformer) and corrected afterwards for elastic deformation of the tools.

The results of three upsetting experiments are discussed. Experiment R is a Rastegaev test with a workpiece as sketched in figure 5.6.1, using lanoline as (solid) lubricant. For the experiments L and D workpieces without recesses are used, in experiment L a liquid lubricant, Ipro L34, is applied, in experiment D tools and workpiece are degreased thoroughly.
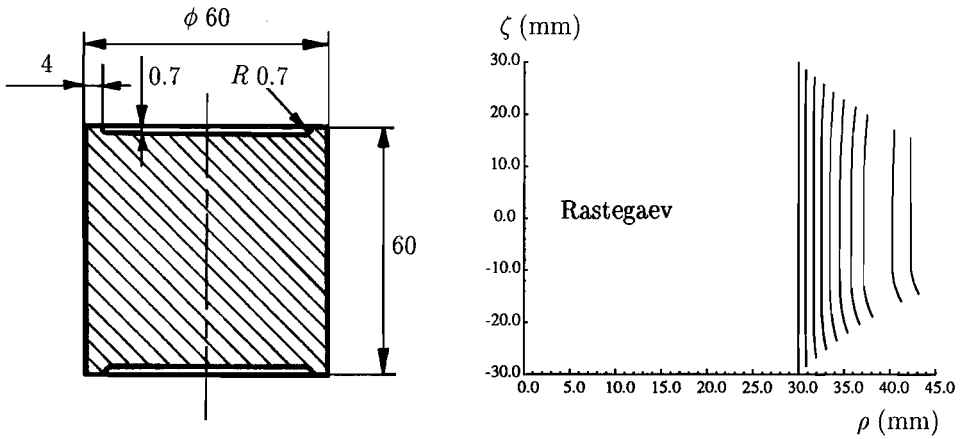
Figure 5.6.1 Rastegaev workpiece and reconstructed contour

Analysis of the photographs resulted in contour reconstructions as given in figure 5.6.1 and figure 5.6.2. For the radial function $R(\zeta_0)$ as well as the axial function $Z(\zeta_0)$, 6 Legendre polynomials (0-th until 5-th degree) are used. Of experiment L and D only the contours are presented for the deformation stages for which all workpiece mark data were available (for higher upsetting stages workpiece marks disappeared into the contact areas with the tools). The contour of the Rastegaev workpiece remained straight besides some disturbance, due to the lips of the recesses. For L and D bulging occurred.
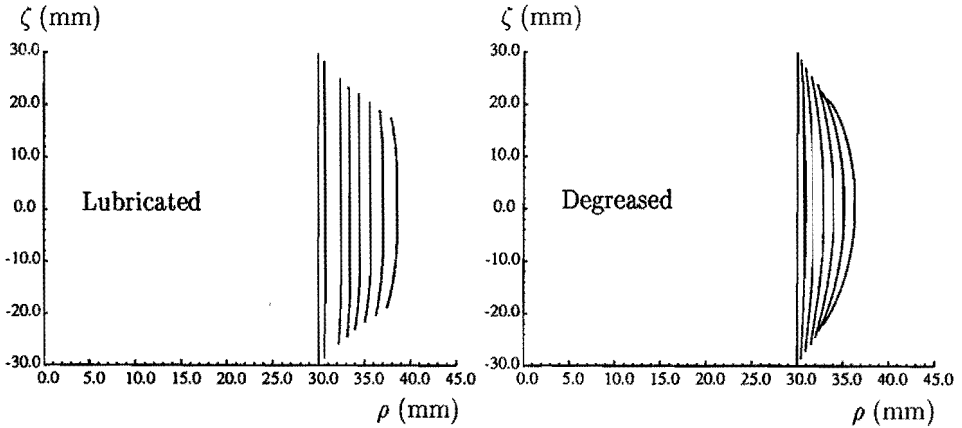
Figure 5.6.2 Reconstructed contours for the conventional workpieces

To investigate the accuracy of the contours reconstructed, the heights following from the LVDT data are subtracted from the heights as calculated from the photogrammetric data. The results are reported in figure 5.6.3. as functions of the upsetting displacement h_0-h.
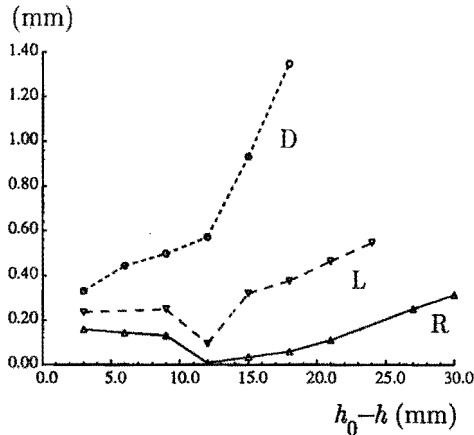
Figure 5.6.3 Height differences between photogrammetric and LVDT data

Before discussing these it has to be mentioned that the LVDT data in general underestimate the actual height of the workpieces, due to offset errors in the

reference LVDT output value used. This reference value should be the output value of the LVDT as both tools are in contact with the undeformed workpiece. Figure 5.6.3 shows that the accuracy as predicted in the previous paragraph is not reached. For the Rastegaev workpiece the height error is maximally 0.16 mm in the low upsetting range. Besides the offset error, two reasons for this can be indicated. First the distortion correction was not so successful in practice for the set-up used. The distortion errors were much larger at the image position of the calibration marks near the image edges than near the centre, resulting in an overestimation of the distortion for the centre part of the image. Secondly errors are introduced by deviation of the initial diameter and height of the workpiece from the assumed values. This results in errors in the reference coordinates of the workpiece marks, affecting all subsequent calculated contour data. A reduction of the errors in the low upsetting range is possible. A more sophisticated photogrammetric system or an alternative distortion correction will reduce errors caused by the distortion of the images. Use of the initial diameter and height, and a more refined method for supplying the workpiece with accurately positioned marks will also render smaller errors.

In the high upsetting range increasing errors occur for the Rastegaev experiment, which are caused by the disturbance of axisymmetry due to the inhomogeneity of the workpiece and to the increase of the surface roughness. With the present set-up the errors due to the violation of axisymmetry can be made smaller by taking interpolation functions also dependent on the tangential coordinate. Another possibility is to choose for a set-up with more than one camera to obtain the three-dimensional contour displacement without assumption of axisymmetry or volume invariance. Use of such data for the quantification of contact models as in chapter 4 requires a three-dimensional numerical analysis. The influence of the surface roughness may be reduced by applying a larger set of smaller marks or by applying larger marks. In the first case the influence is reduced by averaging over more marks, in the second case by the intrinsic averaging over each mark. In order to be able to apply a sufficient number of large marks, unfilled marks may be used, such as e.g. circles. Correction for surface strain gradients is needed for large marks. Use of models for the roughening and application of complex roughness characteristics (Klimczak et al. 1988) can guide the optimization of the mark pattern and mark design. As surface roughness always occurs in the high upsetting range, it sets an upper-bound to the obtainable accuracy for the contour displacements. It is therefore not to be expected that the accuracy, needed for these displacements to be useful in

the sense of chapter 4, can be established in the high upsetting range for workpieces of material with a relatively large average grain size. Despite of the errors in the height of the workpiece, the diameter of the contact surfaces is measured accurately enough for application in the calculations with the stress measuring tool as described in chapter 6. For $\zeta_0 = -30.0$ mm, the diameter resulting photogrammetrically is 86.4 mm and the average diameter measured afterwards 86.8 mm. For $\zeta_0 = 30.0$ mm, these values are 84.5 mm and 86.3 mm respectively. This quite significant difference is caused by a very local disturbance near this edge, too small to be detected with the number of marks used.

For the experiments L and D much larger and increasing errors are found for the height. This is caused by the disappearance of the visible surface into the contact zone. This error can be suppressed by assuming a fixed position of the axis of the workpiece in the object space instead of assuming volume invariance. Another possibility is of course to use a set-up with at least two cameras. The size of the visible surface, disappeared into the contact surface, can be determined by using the independently measured mutual displacement of the tools or a suited marks pattern. These possibilities are not further investigated in the present study.

6 Design and analysis of a stress measuring tool

From the previous the need raised for a measurement method to determine the surface loads on a workpiece in the upsetting test. In this chapter the design and analysis of a stress measuring tool are presented. It is based on the fact that the normal and shear stresses in the contact region cause elastic deformation of the upsetting tool. Use of strain transducers embedded in the tool in the neighbourhood of the contact area give insight in this elastic deformation, from which the surface stresses can be calculated. The advantage of such a procedure with respect to the very few other contact stress measuring methods is that it enables *in situ* measurement without influencing the contact behaviour.

The first paragraph considers the design of the stress measuring tool according to the above strategy. To come to a description of the behaviour of the tool, three modelling steps are used. The first considers the deformation of the tool due to the contact stresses, the second the transfer of this deformation to the strain transducers and the third the change of the electrical output signals of the transducers resulting from their strain. These steps are dealt with in paragraphs 6.2, 6.3 and 6.4 respectively. In paragraph 6.5 these steps are combined into a model for the behaviour of the measuring tool, offering the possibility to estimate the contact stresses from measured strain transducer signals. Calculations indicate the feasibility of the tool. In paragraph 6.6 experiments for the evaluation of the model of the strain transducers are discussed and the remaining steps in order to apply the tool in upsetting experiments in the future are considered.

6.1 Design aspects of the stress measuring tool

In this paragraph the design of the stress measuring tool is discussed. First considerations of feasibility and requirements for use of a tool, suited for local stress measurements, are given. A choice is made for the type of strain transducer to be applied. The steps in the analysis of the measuring tool, mentioned in the introduction above, become concrete and are explained.

The layout of the measuring tool, used as an upsetting tool, is restricted with respect to various aspects. In the sequel a number of considerations is presented with respect to the determination of the relationship between (transducer) output signals and surface loads, the size and position of the transducers and the transducer behaviour.

Theoretically the relationship between output signals and surface loads may be determined in various ways. One possibility would be to subject the tool to a series of calibration loads. A practical difficulty however is the generation of a variety of well-known surface loads. To avoid this problem a simple design of the tool is adapted. This makes it possible to describe the relationship between surface stresses and transducer output signals by combining a number of plain models. A separate evaluation of these models can be performed. In the geometrical layout of the tool ingenious constructions may help to fulfil general wishes like high sensitivity. However, these constructions usually introduce other difficulties, theoretically or practically.

The dimensions of the strain transducers preferably should be small. A large size limits the number of transducers to be used. Furthermore in the case of strain gradients, a relatively large size makes it difficult to identify the material point to which the measured strain applies. The minimum size of the sensing part of the transducers has to be one order of magnitude larger than some limiting physical dimension. This may be a characteristic length of the grains of the tool steel applied. The strain transducers should be positioned such that the desire for a high sensitivity is realized. A sufficient variation of the output signals in the range of the surface loads is needed for accuracy reasons. To accommodate this demand the transducers are located in the neighbourhood of the contact area. Further the output signals have to be insensitive to unknown or only partly known influences, e.g. the deformational behaviour of the support of the tool. These influences may be neglected if strains are measured at a sufficient large distance from these disturbances.

The mechanical and physical behaviour of the strain transducers has to be known in order to avoid the difficult calibration of the measuring tool. However it is unavoidable to perform some measurements on a simpler configuration in order to trace this behaviour. The loads in such cases can be prescribed or measured easier. In the ideal case a strain transducer would generate a signal only depending on one strain component. For all known transducers the output signal however depends on more than one strain component. This dependency should be known. Also the influence of temperature changes on the output signal of the transducer has to be known, in order to predict and suppress disturbances by heating. Furthermore the mechanical interaction of the transducers with their surrounding is of importance. Application of transducers, based on some well-known physical principle and easily to calibrate, are therefore favoured.

According to the previous considerations a measuring tool design is proposed and further investigated. A sketch is given in figure 6.1.1.
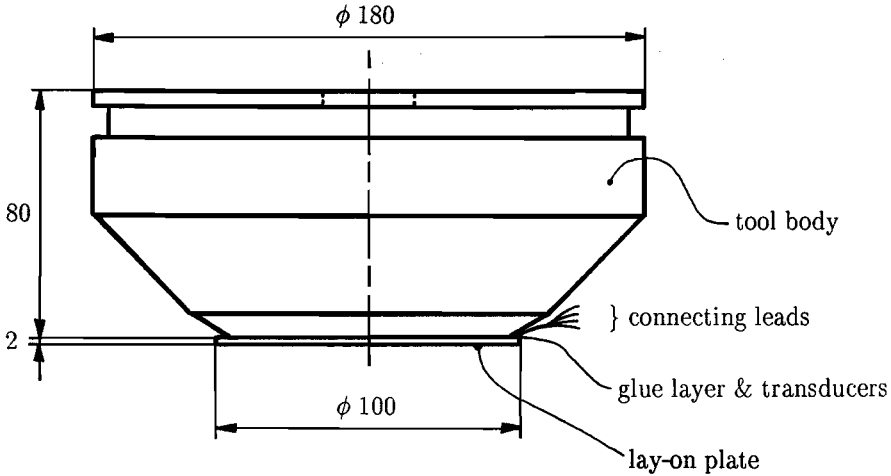
Figure 6.1.1 The stress measuring tool design

The strain transducers are positioned on one side of a thin disc (lay-on plate). At this side the lay-on plate is glued to the tool body, thus direct contact between transducers and workpieces is prevented. No special constructions are designed near the transducers, in order to prevent quantitatively not well-known mechanical behaviour as well as to maintain stiffness and strength. Applicable transducers generate an electrical output signal from piezoelectrical, piezocapacitive or piezoresistive origin. Considering strength, measuring range, application properties and calibration, piezoresistive transducers are preferred. An example of this type is the strain gauge. Application in three-dimensional stress situations are not common (Hoffmann 1987). To obtain reliable results special attention has to be paid to the glueing of the gauges. The influence of the strain in the direction perpendicular to the gauges grid may be non–linear and is only qualitativily understood. Useful application therefore can only be achieved by compensating the influence of the strain in this perpendicular direction through dummy gauges, which is quite easy in e.g. hydrostatic pressure situations but almost impossible in all other three-dimensional situations (Andreae 1974). The problem is caused by the complex geometry of the gauges and the characteristics of the glue layer, which connects the gauge to the substrate. Such a glue layer is not present using thin film transducers evaporated onto the substrate. The substrate strain is transmitted to these

6 Design and analysis of a stress measuring tool

transducers via a very thin ceramic layer. It is therefore to be expected that, in contradiction to gauges, such transducers with a simple geometry show a linear relation between the resistance and the substrate strains.

A part of a typical transducer suited for four-wire resistance measurement is sketched in figure 6.1.2. The active part of the transducer has a thickness much smaller than its inplane dimensions i.e. length and width.
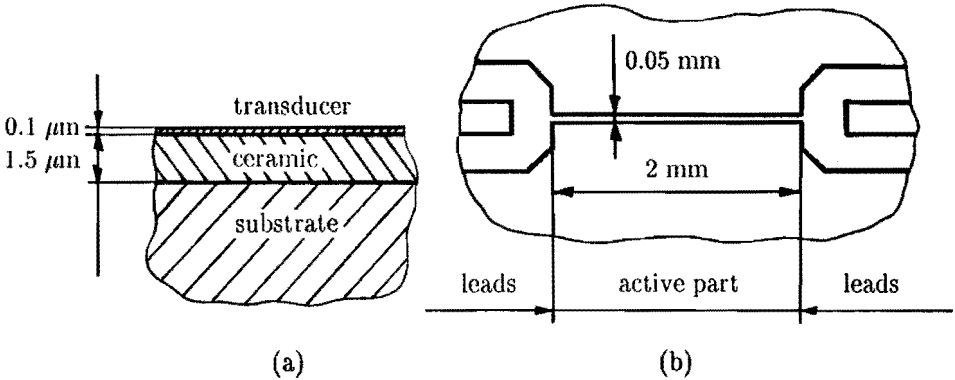
Figure 6.1.2 A typical four—wire thin film transducer in cross section (a) and top view (b)

Active is that part of the thin film of which the resistance changes are registered. Some typical dimensions are indicated in the figure. The electrical insulation from the substrate is realized by the ceramic layer. Successful applications of such transducers in elastohydrodynamic lubrication show a mechanical strength not rejecting application in the stress measuring tool (e.g. Schouten 1973, Peeken & Köhler 1979). As will be seen in the sequel a high stiffness is required for the glue layer which connects the lay-on plate to the tool body. Due to the relative low elastic modulus the layer thickness has to be small. This demand benefits the layer's strength.

As already mentioned, the determination of the relation between surface stresses and strain transducers' output signals by performing just calibrating measurements on the assembled tool is only a theoretical possibility. A feasible procedure is created with a combination of a number of separate analysis steps. These steps, elaborated in the following paragraphs, are successively indicated. A quite simple experiment with the assembled tool can evaluate this method with its underlying assumptions.
— Analysis of an elastic body with stochastic boundary conditions.

The measuring tool is first considered as a homogeneous body of which a part of the boundary conditions is disturbed by measuring errors, comparable to the situation described in chapter 4. In paragraph 6.2 a method is described to calculate the influence of these errors on stresses and strains.

— The strain behaviour of thin layers.

The stress and strain state of the thin film transducers are calculated in paragraph 6.3 from the solution for the homogeneous elastic body. This calculation is based on the assumption that the transducers can be modelled as infinitely thin inclusions. The disturbance of the global stress and strain field of the body is negligibly small. The influence of the thin glue layer is studied by numerical analyses in paragraph 6.5.

— The resistance change of the transducers due to strain and temperature changes.

Many metals show an electrical resistance behaviour which is linear over a large strain and temperature range, which is expected also to hold for evaporated thin films. This behaviour is discussed in paragraph 6.4 and combined with results from paragraph 6.3 for the strain interaction of the transducers with the surroundings.

— Combination of the previous steps offers a method to calculate the behaviour of the tool. Paragraph 6.5 presents results and shows some characteristics and accuracy aspects.

— The transducer model for the relation between resistance and strain can be evaluated through experiments with the transducers under defined strain conditions. Paragraph 6.6 discusses the experimental set-up and some premature results.

— The behaviour of the assembled measuring tool may be evaluated by Rastegaev upsetting tests. Details are dealt with in paragraph 6.6 also.

6.2 Analysis of the influence of stochastic boundary conditions

In this paragraph a method is described for taking uncertain boundary conditions into account for linear elastic mechanics' problems. These conditions generally occur if data originates from experiments. The results calculated are estimates for the in fact unknown real values. Knowledge of the statistics of the experimental data can be used to determine the reliability of these results.

In paragraph 6.2.1 the tool is modelled as a linear elastic body. Considerations with respect to the measured and exact boundary conditions are presented.

Paragraphs 6.2.2 and 6.2.3 describe the strategies with respect to the discretized linear elastic behaviour and the varied set of boundary conditions. In the last paragraph 6.2.4 some attention is paid to the numerical implementation and the accuracy of results.

6.2.1 Description of the problem

Considered is a linear elastic body B, with boundary ∂B. Both the body and its boundary consist of an invariable set of material points. Displacements of the material points are caused by mechanical loads to which the body is subjected. The displacements in the loaded situation with respect to the unloaded reference state are denoted by the vector field $\vec{u}$. For all loading cases the displacements will be small enough to neglect geometrical non–linearities. Thermal effects, inertial loads and distributed loads are neglected.

The material behaviour is supposed to be isotropic linear, according to the generalized Hooke's law

$$\sigma = 2G\left(e + \frac{\nu}{1-2\nu}\mathrm{tr}(e)I\right) \quad , \quad e = \tfrac{1}{2}\left(\vec{\nabla}\vec{u} + (\vec{\nabla}\vec{u})^c\right) \tag{6.2.1}$$

with σ the Cauchy stress tensor, e the linear strain tensor, G the shear modulus and ν the Poisson's ratio. In static problems, at each internal point of the body local equilibrium of forces is satisfied

$$\vec{\nabla}.\sigma = \vec{0} \tag{6.2.2}$$

The equations (6.2.1) and (6.2.2) theoretically can be solved if a complete set of boundary conditions is supplied. As indicated in chapter 3, in physical reality it can occur that applied boundary loads or displacements are only known implicitly. To calculate for such cases a displacement field with some wanted accuracy, sufficient information of other quantities at the boundary or at interior points has to be determined and used in the calculation. This information may be measured with a limited accuracy or may be known exactly. In the application of the stress measuring tool, the lack of information at the contact boundary is replaced by strain data measured in the neighbourhood of the contact area by strain transducers.

From literature, analyses of comparable problems are known, some to be reviewed

here. Kihara et al. (1983) describe the inverse application (in the sense of chapter 3) of the boundary element method to evaluate the unknown tractions across the contact area between tool and workpiece in upsetting. The information compensating for the lack of the boundary conditions consists of surface strains of the upsetting tool outside the contact region. The influence of measuring errors is not considered, so no reliability of the results is available. Oda & Shinada (1987) describe an inverse finite element method combined with a least squares method to calculate the unknown interaction in the contact region between two elastic bodies and the size of the contact region. The compensating data consist of boundary node forces at the supported boundary part. As all boundary conditions are considered exactly known, no reliability of the results obtained is given. A photoelastic-finite element hybrid technique is developed and examined by Chambless et al. (1986). Photoelastically measured stress data are imposed as constraints on finite element problems by a penalty method in order to account for uncertainties in boundary conditions applied. A comparable technique, combining the boundary element method and a least squares method to enforce the influence of experimental digital image correlation data on the solution, is described by Turner et al. (1988). Both methods show a large reduction of the influence of erroneous boundary conditions, but give no indication how strong the experimental data should be weighted or about the reliability of the results. Here it is chosen to combine the finite element method and a minimum covariance estimation technique to account for measured data at boundary or interior points, thus offering a method to choose the weighting of the measured data as well as a reliability estimate for the results calculated.

In the next paragraphs, the following points are discussed. The strategy with respect to the behaviour of the elastic body, as applied in the numerical examples later on, is shortly indicated in paragraph 6.2.2. A mixed finite element formulation is used. In paragraph 6.2.3 the estimation strategy for the displacement field of the body B is explained. Measures for the uncertainty of the displacement field and other derivable quantities as stresses and strains can be given. They can be used to optimize experimental set-ups. Finally some aspects of the numerical implementation and the solution strategy are considered in paragraph 6.2.4.

6.2.2 The description of elastic bodies

A well-known strategy to solve differential equations approximately is the method of weighted residuals. The equilibrium equation (6.2.2) weighted over the volume B by

6 Design and analysis of a stress measuring tool

the weighting vector field $\vec{w}$ reads in the weak formulation

$$\int_B (\vec{\nabla}\vec{w})^c : \sigma \, dV = \int_{\partial B} \vec{w} . \sigma . \vec{n} \, dA \tag{6.2.3}$$

with $\vec{n}$ the unit outward normal vector on ∂B. For discretized continuous displacement fields, elaboration of (6.2.3) and the constitutive equation (6.2.1) via the finite element method leads to a formulation resulting in general in discontinuous stress and strain fields. Because these fields are known to be continuous in reality, a smoothing technique can be applied to obtain unambiguous stress and strain values at all material points. Instead of this technique it is chosen here for a mixed finite element formulation. An advantage of the mixed formulation is that accurate results can be obtained with quite coarse meshes. Furthermore smooth stress and strain fields, desired for the analysis of the measuring tool, are directly calculated by the method. Disadvantage is the relative large computation time needed, which is reduced by substructuring in the present study as indicated in paragraph 6.2.4. In the mixed formulation, besides the nodal displacements, the nodal stresses are considered as independent variables. The constitutive equation (6.2.1) is weighted separately by a second order tensor field W

$$\int_B W : \frac{1}{2}\left(\vec{\nabla}\vec{u} + (\vec{\nabla}\vec{u})^c\right) dV - \int_B W : {}^4C : \sigma \, dV = 0 \tag{6.2.4}$$

with ${}^4C = \frac{1}{2G}\left({}^4I - \frac{\nu}{1+\nu}II\right)$ such that $e = {}^4C : \sigma$. The displacement and weighting vector field in (6.2.3) and the stress and weighting tensor field in (6.2.4) are discretized using nodal quantities. The principle of weighted residuals can be elaborated with appropriate choices for $\vec{w}$ and W. With respect to some vector base this leads to equations for the nodal displacement column $\underset{\sim}{u}$ of length n and the nodal stress column $\underset{\sim}{\sigma}$ of length m, to be denoted as

$$\begin{bmatrix} \underline{0} & \underline{A} \\ \underline{B} & \underline{C} \end{bmatrix} \begin{bmatrix} \underset{\sim}{u} \\ \underset{\sim}{\sigma} \end{bmatrix} = \begin{bmatrix} \underset{\sim}{f} \\ \underset{\sim}{0} \end{bmatrix} \tag{6.2.5}$$

with $\underline{0}$ the zero matrix, $\underset{\sim}{0}$ the zero column, $\underline{A}$ and $\underline{B}^T$ $n{\times}m$ matrices and $\underline{C}$ an $m{\times}m$ matrix. The column $\underset{\sim}{f}$ can be interpreted as the column with nodal forces.

Continuity of the stresses along element boundaries may be assured using a suitable discretization of the stress field. The stress nodal points are not necessarily coinciding with the displacement nodal points. Appropriate discretization results in a regular matrix $\underline{C}$. So unique expressions for $\underline{\sigma}$ and $\underline{f}$ in $\underline{u}$ are given by

$$\underline{\sigma} = -\underline{C}^{-1}\underline{B}\underline{u} \quad , \quad \underline{f} = -\underline{A}\,\underline{C}^{-1}\underline{B}\underline{u} \tag{6.2.6}$$

where $-\underline{A}\,\underline{C}^{-1}\underline{B}$ is a stiffness matrix.

6.2.3 Strategy for the state estimation

In the previous it is shown that the condition of body B is totally described by the displacement column $\underline{u}$, further called the state. To calculate for a given loading case the state $\underline{u}$ of length n, the values of at least n kinematical or dynamical relevant quantities of the body have to be known. These data values may contain measuring errors. Also relatively very accurate data or exactly known data may occur. The data values are called the (exact or measured) observation data, the corresponding quantities observed are called the (exact respectively measurable) state dependent quantities.

The state dependent quantities considered here are linear in $\underline{u}$, e.g. stresses, strains, displacements, forces or linear combinations. Gathering the exact observation data in the column $\underline{z}_e$ the following n_e relations are supposed to hold

$$\underline{z}_e = \underline{H}_e \underline{u} \tag{6.2.7}$$

with $\underline{H}_e$ a constant matrix which expresses the relation of the (exact) state dependent quantities to the state $\underline{u}$ of body B. The n_m measured observation data are supposed to equal the measurable state dependent quantities disturbed by an additive column with observation errors $\underline{v}$

$$\underline{z}_m = \underline{H}_m \underline{u} + \underline{v} \tag{6.2.8}$$

where $\underline{H}_m$ relates the (measurable) state dependent quantities to the state of body B. It is assumed for the random column $\underline{v}$ that the mean and the symmetric covariance matrix are known

$$E(\underset{\sim}{v}) = \underset{\sim}{0} \quad , \quad E(\underset{\sim}{v}\underset{\sim}{v}^T) = \underline{R} \tag{6.2.9}$$

Equations (6.2.9) express that systematic and model errors are negligibly small.

In the foregoing the state column $\underset{\sim}{u}$ is considered deterministic, contrary to the stochastic error column $\underset{\sim}{v}$. The observations are all linear in this state $\underset{\sim}{u}$. In literature such a statistic model is indicated as a static linear Fisher model (see e.g. Schweppe 1973). For the unknown state $\underset{\sim}{u}$ an estimate $\hat{\underset{\sim}{u}}$, linear in $\underset{\sim}{z}_e$ and $\underset{\sim}{z}_m$, will be derived

$$\hat{\underset{\sim}{u}} = \underline{W}_e \underset{\sim}{z}_e + \underline{W}_m \underset{\sim}{z}_m \tag{6.2.10}$$

The constant matrices $\underline{W}_e (n \times n_e)$ and $\underline{W}_m (n \times n_m)$ define the estimator for $\underset{\sim}{u}$. This estimator is required to be unbiased

$$E(\hat{\underset{\sim}{u}}) = \underset{\sim}{u} \tag{6.2.11}$$

Then the matrices $\underline{W}_e$ and $\underline{W}_m$ of the estimator are related by

$$\underline{W}_e \underline{H}_e + \underline{W}_m \underline{H}_m = \underline{I} \tag{6.2.12}$$

The estimator is considered to be optimal if it minimizes some norm of the covariance matrix $\underline{\Sigma}$ of the state estimate

$$\underline{\Sigma} = E\big((\hat{\underset{\sim}{u}} - E(\hat{\underset{\sim}{u}}))(\hat{\underset{\sim}{u}} - E(\hat{\underset{\sim}{u}}))^T\big) = \underline{W}_m \underline{R} \, \underline{W}_m^T \tag{6.2.13}$$

A suitable norm is the trace. For all possible measuring results it demands the estimate $\hat{\underset{\sim}{u}}$ to be in the neighbourhood of the real but unknown state $\underset{\sim}{u}$. For $\text{tr}(\underline{\Sigma})$ holds

$$\text{tr}(\underline{\Sigma}) = E\big((\underset{\sim}{u} - \hat{\underset{\sim}{u}})^T(\underset{\sim}{u} - \hat{\underset{\sim}{u}})\big) = \text{tr}(\underline{W}_m \underline{R} \, \underline{W}_m^T) \tag{6.2.14}$$

Minimizing $\text{tr}(\underline{\Sigma})$ with respect to $\underline{W}_e$ and $\underline{W}_m$ under the constraints (6.2.12) may be transformed to minimizing

$$\text{tr}(\underline{W}_m \underline{R} \, \underline{W}_m^T) + \text{tr}\Big((\underline{W}_e \underline{H}_e + \underline{W}_m \underline{H}_m - \underline{I})\underline{\Lambda}\Big) \tag{6.2.15}$$

with respect to the estimator matrices and the Lagrange multipliers $\underline{\Lambda}$. From this, after identifying $\underline{\Lambda}$ with $-\underline{\Sigma}$, can be derived

$$\begin{bmatrix} \underline{0} & \underline{H}_e \\ \underline{H}_e^T & -\underline{H}_m^T \underline{R}^{-1} \underline{H}_m \end{bmatrix} \begin{bmatrix} \underline{W}_e^T \\ -\underline{\Sigma} \end{bmatrix} = \begin{bmatrix} \underline{0} \\ \underline{I} \end{bmatrix} \quad , \quad \underline{W}_m = \underline{\Sigma}\, \underline{H}_m^T \underline{R}^{-1} \tag{6.2.16}$$

With (6.2.10) and $\underline{W}_m$ according to (6.2.16), $\hat{\underline{u}}$ may be written as

$$\hat{\underline{u}} = \begin{bmatrix} \underline{W}_e & -\underline{\Sigma} \end{bmatrix} \begin{bmatrix} \underline{z}_e \\ -\underline{H}_m^T \underline{R}^{-1} \underline{z}_m \end{bmatrix} \tag{6.2.17}$$

Again with (6.2.16) an explicit expression for $\hat{\underline{u}}$ follows

$$\hat{\underline{u}} = \begin{bmatrix} \underline{0} & \underline{I} \end{bmatrix} \begin{bmatrix} \underline{0} & \underline{H}_e \\ \underline{H}_e^T & -\underline{H}_m^T \underline{R}^{-1} \underline{H}_m \end{bmatrix}^{-1} \begin{bmatrix} \underline{z}_e \\ -\underline{H}_m^T \underline{R}^{-1} \underline{z}_m \end{bmatrix} \tag{6.2.18}$$

In appendix 6.2.1 it is shown that result (6.2.18) also follows from (6.2.7), (6.2.8) and (6.2.9) as a maximum likelihood estimate if $\underline{v}$ is assumed to be normal distributed. In that case the estimate $\hat{\underline{u}}$ is normal distributed with covariance matrix $\underline{\Sigma}$ and expected value $\underline{u}$.

For the measuring tool, the state $\underline{u}$ itself, containing all nodal displacements, is hardly interesting. Interesting quantities are for instance normal and tangential surface stresses, to be considered as linear expressions in $\underline{u}$. The statistical behaviour of these output quantities with respect to measured input quantities such as the signals of the strain transducers is of importance. The input and output quantities can be gathered in the column $\underline{q}$ according to $\underline{q} = \underline{M}_q \underline{u}$ with constant matrix $\underline{M}_q$. The estimate $\hat{\underline{q}}$ simply follows from replacement of $\underline{u}$ by $\hat{\underline{u}}$. The covariance matrix $\underline{\Sigma}_q$ of $\underline{q}$ is related to $\underline{\Sigma}$ by

$$\underline{\Sigma}_q = E\Big((\hat{\underline{q}} - E(\hat{\underline{q}}))(\hat{\underline{q}} - E(\hat{\underline{q}}))^T\Big) = \underline{M}_q \underline{\Sigma}\, \underline{M}_q^T \tag{6.2.19}$$

Notice that $\hat{\underset{\sim}{q}}$ is an unbiased estimate because $\hat{\underset{\sim}{u}}$ is unbiased.

6.2.4 Numerical implementation

In this paragraph some aspects of the numerical implementation, used for calculations on the measuring tool design, are discussed.

A six-node axisymmetric isoparametric element with quadratic interpolation for the displacements and stresses is applied. According to Zienkiewicz et al. (1986) this element is stable and consistent. In order to be able to choose the distribution of the elements in the mesh for the tool calculations, some simple calculations are performed for evaluating the accuracy of the element. A too coarse mesh results in relatively large modelling errors of the discretized model with respect to the continuous linear elastic theory. Mesh refinement results in smaller model errors but increases computational effort. Model errors one order of magnitude smaller than measuring errors are acceptable. The calculations showed that a relative high accuracy can be obtained with the mixed element for quite coarse element meshes.

A disadvantage of the mixed method is given by the structure of the resulting equation (6.2.5). In order to save computing time special solution procedures may be used, see Zienkiewicz et al. (1985). For the measuring tool many different calculations with equal configurations are performed. The technique of substructuring seemed advantageous. In practice it showed indeed useful to obtain reasonable program execution times.

6.3 The mechanical behaviour of thin elastic films

In the design of the measuring tool as presented in paragraph 6.1 some thin films can be appointed. The glue layer and the insulating ceramic layer can be mentioned, but also the active part of the transducers and the connecting leads. For the evaporated parts the thickness in comparison to inplane dimensions is so small that serious problems can be expected in a numerical analysis of the stress and strain state using the finite element method. Therefore an analytical approach, according to Eshelby (1957), is used, considering the mechanical behaviour of an infinite elastic matrix with an elastic inclusion. In paragraph 6.3.1 this method is resumed and extended with thermal influence. The special case of an inclusion with very small thickness with respect to length and width, i.e. a layer, is derived in paragraph 6.3.2. An alternative analysis confirms the formulas found. The translation to the situation of

the evaporated parts of the measuring tool is given. In the last paragraph the influence of some disturbances, not contained in the theory, is established.

6.3.1 Eshelby's method for calculating the elastic behaviour of a matrix with inclusion

In this paragraph Eshelby's approach is reviewed and extended to account for thermal effects. These effects are of importance for thin film transducers.

The purpose of Eshelby's analysis is to describe the stress and strain fields in an infinite body, consisting of a linear elastic matrix with a linear elastic inclusion whose material characteristics differ from the matrix. The material of the matrix as well as the inclusion is supposed to be homogeneous. The loading of the infinite body is such that the stress and strain field would be uniform, if the body was totally homogeneous. To determine the desired stress and strain fields of matrix and inclusion, Eshelby uses three artificial steps, to be described in the following. Two extra steps are needed to determine the stress and strain fields due to thermal loading.

1) Considered is an unloaded homogeneous linear elastic body with material behaviour according to Hooke's law (6.2.1) with material tensor ^4L_m. A part of this body, further called the homogeneity, is cut out and removed from the rest of the body, called the matrix in the sequel. For the homogeneity a uniform transformation strain e_T is induced, due to some fictitious physical process not to be specified. This process is such that the homogeneity remains stress free and its material behaviour does not change. By a surface load the original shape of the homogeneity is restored in order to replace it in the matrix. After connection this load is relaxed. An unloaded system results with an inhomogeneous strain field $e_C(\vec{x})$ with respect to the original state, caused by e_T. This "constrained" strain field $e_C(\vec{x})$ at a position $\vec{x}$ in the matrix or homogeneity can, according to the theory of linear elasticity, be written as

$$e_C(\vec{x}) = {}^4M(\vec{x}){:}e_T \qquad (6.3.1)$$

The fourth order tensor ^4M, as specified in appendix 6.3.1 for arbitrary shaped homogeneities, can be calculated numerically for arbitrary $\vec{x}$. The strain field

$e_C(\vec{x})$ of the matrix fades out rapidly with increasing distance from the homogeneity and is not considered here further. For an ellipsoidal homogeneity, to which the rest of this paragraph is restricted, analytical results can be given. The strain e_C in an ellipsoidal homogeneity is uniform and given by

$$e_C = {}^4S : e_T \qquad\qquad (6.3.2)$$

The constant components S_{pqrs}, $p,q,r,s \in \{1,2,3\}$, of ${}^4S$ with respect to the orthonormal vector base defined by the principal axes of the ellipsoid, are unequal zero for $p=q \wedge r=s$, $p=s \wedge q=r$, $p=r \wedge q=s$. The non-trivial components of ${}^4S$ depend on Poisson's ratio and the lengths of the principal axes of the ellipsoid. Explicit relations are given in appendix 6.3.2.

2) A uniform strain field e_A is applied to the assembly of matrix and homogeneity as remaining after step 1. The resulting strain field is the superposition of e_C and e_A. The shape of the homogeneity is then determined by the strain $e_C + e_A$. The stress σ in the homogeneity follows from the generalized Hooke's law

$\sigma = {}^4L_m : e$ with substitution of $e = e_C + e_A - e_T$.

3) The homogeneity can be replaced by an inhomogeneity or inclusion with material tensor ${}^4L_i$, without transformation strain, and unloaded of the same size as the untransformed homogeneity at the beginning of step 1. This replacement is performed such that the state of the matrix is not changed. The requirement for this is that by the strain $e_C + e_A$ of the inhomogeneity the same stress σ is generated as present after step 2. Continuity of displacement and stress at the inclusion–matrix interface is then guaranteed. The mathematical formulation reads

$$ {}^4L_i : (e_C + e_A) = {}^4L_m : (e_C + e_A - e_T) \qquad\qquad (6.3.3)$$

With (6.3.2) and (6.3.3) the desired relation $e_C = e_C(e_A)$ between applied strain e_A and constrained strain e_C for ellipsoidal inclusions can be determined. The resulting strain field $e_C + e_A$ is caused by loads, which induce a uniform strain field e_A in a totally homogeneous body. The equations (6.3.2) and (6.3.3) are presented with respect to the previously defined orthonormal vector base, coinciding with the principal axes of the ellipsoid, in appendix 6.3.3.

4) The relation $e_C = e_C(e_A)$ derived above is not yet suited when thermal strain occurs in matrix and inclusion. The steps 1 and 2 for the homogeneity are modified first. After isolating the homogeneity it is, together with the matrix, subjected to a temperature change ΔT_m with respect to the reference state, inducing the thermal strain $e_{th,m} = \frac{1}{3}\alpha_m \Delta T_m I$ with α_m the thermal coefficient of cubical expansion. The homogeneity also acquires the transformation strain e_T, such that it remains stress free and the material behaviour does not change. The load needed for connection is again relaxed afterwards. In the homogeneity the resulting strain is $e_C + e_{th,m}$ with $e_C = {}^4S{:}e_T$ and the resulting stress equals ${}^4L_m{:}(e_C - e_T)$. The tensor ${}^4S$ is the same as in (6.3.2). On the strain field of matrix and homogeneity again the applied strain e_A is superposed. This is generated e.g. by suppressing the thermal expansion of the matrix at a large distance from the homogeneity. The total strain in the homogeneity becomes $e_C + e_A + e_{th,m}$ and the stress ${}^4L_m{:}(e_C + e_A - e_T)$.

5) The homogeneity is replaced by an inhomogeneity, characterized by the material properties ${}^4L_i$ and α_i, with thermal strain $e_{th,i} = \frac{1}{3}\alpha_i \Delta T_i I$ (ΔT_i is the temperature change with respect to the reference state) but without transformation strain. The dimensions of the inhomogeneity without thermal strain are equal to the original dimensions of the homogeneity. Maintaining the shape, prescribed by the strain $e_C + e_A + e_{th,m}$, the difference between the strain $e_C + e_A + e_{th,m}$ and the strain $e_{th,i}$ generates the stress ${}^4L_i{:}(e_C + e_A + e_{th,m} - e_{th,i})$. Continuity of stress is guaranteed if

$$
{}^4L_i{:}(e_C + e_A + e_{th,m} - e_{th,i}) = {}^4L_m{:}(e_C + e_A - e_T) \quad , \quad e_C = {}^4S{:}e_T \tag{6.3.4}
$$

The expression for the contribution of the extra thermal difference strain $e_{th,m} - e_{th,i}$ to e_C is given in appendix 6.3.3.

Resuming, the uniform strain field $e_{tot,i}$ of a linear elastic ellipsoidal inclusion in an infinite linear elastic matrix equals $e_C + e_A + e_{th,m}$ with e_C according to (6.3.4). Furthermore, the uniform stress field is given by ${}^4L_i{:}(e_{tot,i} - e_{th,i})$. In the next paragraph the above results are used to study the special case of infinitely flat ellipsoidal inclusions.

6.3.2 Specialization of Eshelby's method to infinitely thin inclusions

First the case of a thin inclusion is considered to derive the desired equations. An alternative approach is presented, which confirms the formulas found. This approach offers an extension with respect to the applicability of the theory. For a thin inclusion in a layer, representing a model of a transducer on top of a ceramic layer on a substrate, results are obtained.

Considered is the case of an ellipsoidal inclusion with a thickness $2a_t$ which is small in comparison to its width $2a_w$ and length $2a_l$, i.e. $a_t/a_w \ll 1$, $a_t/a_l \ll 1$. The quantities a_p with $p \in \{l,w,t\}$ are equal to the lengths of the half principal axes of the inclusion. Taking the limit of the integrals appearing in ${}^4S$ for a_t/a_w and a_t/a_l approaching zero, with appendix 6.3.2 explicit expressions for the components S_{pqrs}, $p,q,r,s \in \{l,w,t\}$, result. Only the following components appear to be unequal to zero

$$S_{tttt} = 1 \quad , \quad S_{ttpp} = \frac{\nu_m}{1-\nu_m} \quad , \quad S_{pttp} = S_{tppt} = S_{tptp} = S_{ptpt} = \frac{1}{2} \qquad (6.3.5)$$

with $p \in \{l,w\}$ and ν_m Poisson's ratio of the matrix. For finite but small thickness, (6.3.5) constitutes an approximation with relative errors of order $(a_t/a_p)^2$.

In the sequel for a concise notation the components of a symmetric second order tensor T are gathered in columns as

$$\underset{\sim}{T} = [\ T_{ll} \quad T_{ww} \quad T_{tt} \quad T_{lw} \quad T_{wt} \quad T_{tl}\ ]^T \qquad (6.3.6)$$

According to this convention, the strain column $\underset{\sim}{e}_{C,i}$ is defined containing the components of e_C with respect to the thin inclusion and $\underset{\sim}{e}_{A,m}$ with the components of the strain e_A applied at the matrix. Substitution of (6.3.5) in (6.3.4) yields an expression for $\underset{\sim}{e}_{C,i}$ as function of the strain $\underset{\sim}{e}_{A,m}$ and the temperature changes ΔT_m and ΔT_i of the matrix and inclusion respectively, with respect to the reference state. Assuming isotropic behaviour for matrix and inclusion it follows

$$\underset{\sim}{e}_{C,i} = \underset{\sim}{C}_{im}\underset{\sim}{e}_{A,m} + \underset{\sim}{\Delta}_{im} \quad , \quad \underset{\sim}{C}_{im} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \varphi & \varphi & \chi-1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \psi-1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \psi-1 \end{bmatrix} \quad , \quad \underset{\sim}{\Delta}_{im} = \begin{bmatrix} 0 \\ 0 \\ \eta \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad ,$$

$$\varphi = \frac{G_m}{G_i}\frac{1-2\nu_i}{1-2\nu_m}\frac{\nu_m}{1-\nu_i} - \frac{\nu_i}{1-\nu_i} \quad , \quad \chi = \frac{G_m}{G_i}\frac{1-2\nu_i}{1-2\nu_m}\frac{1-\nu_m}{1-\nu_i} \quad , \quad \psi = \frac{G_m}{G_i} \quad ,$$

$$\eta = \frac{1+\nu_i}{1-\nu_i}(\tfrac{1}{3}\alpha_i\Delta T_i - \tfrac{1}{3}\alpha_m\Delta T_m) \tag{6.3.7}$$

where G_m, ν_m and G_i, ν_i are the shear modulus and Poisson's ratio of the matrix and inclusion material respectively. The total strain $\underset{\sim}{e}_{tot,i}$ of the inclusion equals

$$\underset{\sim}{e}_{tot,i} = \underset{\sim}{e}_{C,i} + \underset{\sim}{e}_{A,m} + \underset{\sim}{e}_{th,m} = (\underline{C}_{im} + \underline{I})\underset{\sim}{e}_{A,m} + \underset{\sim}{\Delta}_{im} + \underset{\sim}{e}_{th,m} \tag{6.3.8}$$

with $\underline{I}$ the unit matrix of rank 6 and $\underset{\sim}{e}_{th,m}$ a representation of $e_{th,m}$ according to (6.3.6). The stress column $\underset{\sim}{\sigma}_i$ in the inclusion follows from $\underset{\sim}{\sigma}_i = \underline{L}_i(\underset{\sim}{e}_{tot,i} - \underset{\sim}{e}_{th,i})$ with the isotropic linear elastic material matrix

$$\underline{L}_i = \begin{bmatrix} \lambda+2\mu & \lambda & \lambda & & & \\ \lambda & \lambda+2\mu & \lambda & & \underline{0} & \\ \lambda & \lambda & \lambda+2\mu & & & \\ & & & 2\mu & 0 & 0 \\ & \underline{0} & & 0 & 2\mu & 0 \\ & & & 0 & 0 & 2\mu \end{bmatrix} \tag{6.3.9}$$

with Lamé constants $\lambda = 2G_i\nu_i/(1-2\nu_i)$ and $\mu = G_i$.

An alternative method can be followed to derive the equations (6.3.8). Considered is a thin layer in a matrix, both isotropic linear elastic. The generalized Hooke's law for the layer, extended for thermal strains reads

$$\underset{\sim}{\sigma}_i = \underline{L}_i(\underset{\sim}{e}_{tot,i} - \underset{\sim}{e}_{th,i}) \tag{6.3.10}$$

with $\underline{L}_i$ according to (6.3.9). Assumptions can be made with respect to the interaction between the matrix and a thin layer of large inplane dimensions, if the Lamé constants λ and μ of the matrix material are of the same order as those of the layer. At first it is supposed that the layer does not influence the mechanical behaviour of the matrix, while the matrix, also called the substrate, prescribes the strain in the plane of the layer

$$(e_{tot,i})_{pp} = (e_{A,m})_{pp} + \tfrac{1}{3}\alpha_m\Delta T_m \quad , \quad (e_{tot,i})_{pq} = (e_{A,m})_{pq} \tag{6.3.11}$$

with $p, q \in \{l, w\}$, $p \neq q$. Further, the stress and strain in the layer are supposed to be uniform over the thickness. At the layer–matrix interfaces the stress vector $\vec{p}_i$ on the layer is the opposite of the stress vector $\vec{p}_m$ on the matrix because of the law of action and reaction. The components $\underset{\sim}{p}_m$ of $\vec{p}_m$ are directly related to $\underset{\sim}{\sigma}_m$ by

$$\underset{\sim}{p}_m = [\ (p_m)_l\ (p_m)_w\ (p_m)_t\ ]^T = [\ (\sigma_m)_{tl}\ (\sigma_m)_{wt}\ (\sigma_m)_{tt}\ ]^T\ ,$$

$$\underset{\sim}{\sigma}_m = \underline{L}_m(\underset{\sim}{e}_{A,m} - \underset{\sim}{e}_{th,m}) \tag{6.3.12}$$

The components $\underset{\sim}{p}_i$ of $\vec{p}_i$ are related to $\underset{\sim}{\sigma}_i$ by

$$\underset{\sim}{p}_i = -[\ (\sigma_i)_{tl}\ (\sigma_i)_{wt}\ (\sigma_i)_{tt}\ ]^T \tag{6.3.13}$$

The equations (6.3.10) through (6.3.13) combined with $\underset{\sim}{p}_m = -\underset{\sim}{p}_i$ deliver expression (6.3.8) for $\underset{\sim}{e}_{tot,i}$ as found by Eshelby's method. No assumptions about the particular shape of the layer and the position with respect to the matrix are made in this second approach. Obviously, provided the elastic constants of matrix and layer are of the same order, the results of Eshelby's method are also valid for arbitrary shaped included thin layers and thin layers at the surface of a matrix (substrate). For surface layers the strain $\underset{\sim}{e}_{A,m}$ obeys (6.3.12) with $\underset{\sim}{p}_m$ the components of the external stress vector acting on the layer.

Not only the case of one surface layer is of interest, but also the case of twofold surface layers. The top layer is for instance a strain transducer, while the other layer can represent the intermediate ceramic insulator. It is assumed in this paragraph that Eshelby's formulas remain valid irrespective to the real geometrical situation. The next paragraph considers the validity of this assumption. The total strain $\underset{\sim}{e}_{tot,i} = \underset{\sim}{e}_{C,i} + \underset{\sim}{e}_{A,m} + \underset{\sim}{e}_{th,m}$ of the intermediate layer follows from the formulas presented before. The same applies to the total strain $\underset{\sim}{e}_{tot,o} = \underset{\sim}{e}_{C,o} + \underset{\sim}{e}_{A,i} + \underset{\sim}{e}_{th,i}$ of the outside layer for which the intermediate layer acts as substrate. The strain $\underset{\sim}{e}_{A,i}$ is the strain induced in the intermediate layer by the substrate, and is therefore equal to $\underset{\sim}{e}_{tot,i} - \underset{\sim}{e}_{th,i}$. It results for $\underset{\sim}{e}_{tot,o}$

$$\underset{\sim}{e}_{tot,o} = (\underline{C}_{om} + \underline{I})\underset{\sim}{e}_{A,m} + \underset{\sim}{\Delta}_{om} + \underset{\sim}{e}_{th,m} \tag{6.3.14}$$

This result is analogous to the equations (6.3.8) for $e_{tot,i}$ with the quantities of the inclusion replaced by those of the outside layer. The properties of the intermediate layer are not represented in (6.3.14). Reminding the alternative approach presented before this is obvious. The requirements the intermediate layer has to obey for a transfer of stress and strain from the substrate, count also for the outside layer.

Until now isotropy for the layers is assumed. Because of the production method of the layers, anisotropy may occur. Most likely the inplane behaviour of the layer will be isotropic while the behaviour in the thickness direction may be different. The modified version of equation (6.3.7) for such a transversal isotropic case is given in appendix 6.3.4.

As stated before, in the next paragraph some assumptions, made in the previous, will be considered with respect to the real geometry of the measuring tool.

6.3.3 The influence of disturbances and the glue layer

In this last paragraph on the mechanical behaviour of thin layers, it is discussed whether the modelling of the active parts of the strain transducers, as described in paragraph 6.3.2, is allowed. Also the assumption that the glue layer is thin is regarded shortly.

A typical layout of the active part of a strain transducer is sketched in figure 6.1.2, which shows that the ceramic layer is very thin with respect to its inplane dimensions. The global strain variations of the substrate have a characteristic length in the order of these inplane dimensions. The strain of the substrate may therefore be considered uniform over a subarea of the ceramic layer, with dimensions much larger than its thickness. The same remarks can be made with respect to the transducer. The strain field for the active part can be assumed uniform. Disturbances may occur from the constrained strain of the leads of the transducers. This disturbing field however fades out over a short distance from the leads. This influence is negligible for transducers with large inplane dimensions with respect to the thickness of the leads.

The glue layer, necessary to connect the lay-on plate to the tool body, is very thin. The elastic properties of the glue however are so different from those of the tool body, that only for extreme thin layers the alternative approach will be valid. As such thin layers cannot be realized in practice, the influence of the glue layer is studied numerically in paragraph 6.5.

6.4 The strain dependency of resistive strain transducers

A phenomenological description is given for the resistance change of the active part of thin film strain transducers with strain and temperature variations. The dependency on strain and temperature change is supposed to be linear. Most of the metals, commonly used as transducer material, exhibit such a linear dependency in a large area of application. The behaviour of a thin film strain transducer is comparable to the behaviour of the strain gauge, often used in experimental mechanics. The advantages of evaporated thin film transducers over strain gauges for application in the measuring tool are indicated.

Changes of the strain and temperature of materials influence the physical properties. Piezoelectrical and piezoresistive phenomena may be observed as examples of this behaviour. Resistance change can be used to determine the strain at a surface with strain gauges or thin film transducers of conducting or semiconducting material. In this paragraph a relation for the resistance as function of strain and temperature is derived for thin conducting films, comparable to equations used in literature (see e.g. Barsis et al. 1970, Dössel 1984), and combined with the results of paragraph 6.3, expressing the strain of a thin film transducer in the substrate strain.

A theory for the resistance strain temperature dependency is not needed to set up a phenomenological description of this phenomenon. It is only stated here that such theories predict non–linear relations for very thin transducers (e.g. Bedda et al. 1986). This is in accordance with experimental data (e.g. Schouten 1973). Because a linear dependency is much more convenient, films of sufficient thickness, i.e. of the order 10^{-1} μm, will be used. Semiconductors and some particular conductors nevertheless show a non-linear characteristic, which are disapproved for application here.

For many conductors the resistance behaviour is more pronounced than other phenomena such as piezoelectricity. At a material point of such a conductor the generalized Ohm's law holds

$$\vec{E} = \rho.\vec{J} \qquad (6.4.1)$$

The electrical field vector (gradient of the electrical potential) is indicated by $\vec{E}$, the vector of the electrical current density by $\vec{J}$. In (6.4.1) these vector quantities are related by the symmetric second order specific resistance tensor ρ. It is assumed here

that significant change of ρ occurs only due to straining and temperature changes. Other influences are considered absent or negligibly small. As strain quantity the linear strain e with respect to some reference state is used. The temperature change ΔT is defined as the difference between the actual temperature T of the material and the temperature T_0 at the reference state. Over some range of strain and temperature the change of ρ will be linear in e and ΔT

$$\rho = \rho_0 + {}^4\pi{:}e + \epsilon\Delta T \quad , \quad {}^4\pi = (\frac{\partial\rho}{\partial e^c})^c\bigg|_{T=T_0, e=0} \quad , \quad \epsilon = \frac{\partial\rho}{\partial T}\bigg|_{T=T_0, e=0} \qquad (6.4.2)$$

In (6.4.2) $^4\pi$ is the fourth order piezoresistive tensor and ϵ the second order thermoresistive tensor.

In order to derive from (6.4.1) and (6.4.2) a relationship between resistance change and strain and temperature change of the active part of the transducer, information about the strain, temperature, geometry and electrical boundary conditions is needed. The active part is the part of the thin film of which the resistance change is registered. Paragraph 6.3 showed that the strain field may be considered uniform over the active part. Because of its small dimensions also the temperature may be considered uniform. In a lengthy rectangular part of a thin film of quite uniform composition the vector of current density coincides with the longitudinal axis when an electrical potential difference is induced in that direction. Using an orthonormal base $\{\vec{e}_l, \vec{e}_w, \vec{e}_t\}$ with indices for length, width and thickness direction, only the component J_l of $\vec{J}$ will be unequal to zero. With respect to the piezoresistive tensor it is noticed that it may be isotropic or transversal isotropic with $\vec{e}_t$ as direction of the symmetry axis. Transversal isotropy can be caused by the production method of the thin film. If it is evaporated, $\vec{e}_t$ is the growth direction and in that direction therefore the material behaviour may be different from the behaviour in the inplane directions. For the interesting component E_l of the electrical field vector $\vec{E}$ in longitudinal direction, it can be shown that

$$E_l = (\rho_0 + \pi_l e_{ll} + \pi_w e_{ww} + \pi_t e_{tt} + \epsilon\Delta T)J_l \qquad (6.4.3)$$

for the transversal isotropic case, with ρ_0 the reference specific resistance in longitudinal direction, π_l, π_w and π_t piezoresistive characteristics for the

longitudinal, width and thickness direction, and ϵ the relevant thermoresistive material parameter. For the isotropic case holds $\pi_w = \pi_t$. Some materials show an even simpler strain dependency. The resistance change of such ideal resistors only occurs through volume strain and temperature change, thus $\pi_l = \pi_w = \pi_t = \pi$. Assuming E_l uniform, integration of (6.4.3) over the volume of the active part can be performed easily. For the resistance of this part of length L and cross section area A it follows

$$R = \frac{\int\limits_L E_l dL}{\int\limits_A J_l dA} = \frac{L}{A}(\rho_0 + \pi_l e_{ll} + \pi_w e_{ww} + \pi_t e_{tt} + \epsilon \Delta T) \qquad (6.4.4)$$

The assumption of a uniform E_l will hold globally for the active part. In the vicinity of the leads disturbances will occur, but the contribution to R is supposed to be so small that it may be neglected.

The actual dimensions L and A of the transducer are determined by the strains

$$L = L_0(1 + e_{ll}) \quad , \quad A = A_0(1 + e_{ww})(1 + e_{tt}) \qquad (6.4.5)$$

with L_0 and A_0 the length and cross section area in the reference state. The relative resistance change $\Delta R / R_0$ with $\Delta R = R - R_0$ and R_0 the resistance in the reference state follows with (6.4.4) and (6.4.5)

$$\frac{\Delta R}{R_0} = \frac{(\rho_0 + \pi_l e_{ll} + \pi_w e_{ww} + \pi_t e_{tt} + \epsilon \Delta T)(1 + e_{ll})}{\rho_0(1 + e_{ww})(1 + e_{tt})} - 1 \qquad (6.4.6)$$

For small strains linearization is allowed

$$\frac{\Delta R}{R_0} = (\frac{\pi_l}{\rho_0} + 1)e_{ll} + (\frac{\pi_w}{\rho_0} - 1)e_{ww} + (\frac{\pi_t}{\rho_0} - 1)e_{tt} + \frac{\epsilon}{\rho_0}\Delta T \qquad (6.4.7)$$

Equation (6.4.7) expresses a linear dependency of the relative resistance change on the relevant strain components and temperature change.

The strain components as used in (6.4.7) are those of the active part of the transducer. In the application as strain transducer, one is interested in the strains of its surroundings. It is therefore needed to transfer the transducer strain e into the at

the surroundings applied strain $e_{A,m}$. The relationship is derived in paragraph 6.3, where $e_{tot,o}$ is used instead of e. Substitution of (6.3.14) in (6.4.7) yields

$$\frac{\Delta R}{R_0} = \left(\frac{\pi_l}{\rho_0}+1+\left(\frac{\pi_t}{\rho_0}-1\right)\varphi\right)(e_{A,m})_{ll}+\left(\frac{\pi_w}{\rho_0}-1+\left(\frac{\pi_t}{\rho_0}-1\right)\varphi\right)(e_{A,m})_{ww}+\left(\frac{\pi_t}{\rho_0}-1\right)\chi(e_{A,m})_{tt}+$$

$$+\left(\frac{\epsilon}{\rho_0}+\left(\frac{\pi_t}{\rho_0}-1\right)\eta^o\right)\Delta T_o+\left(\left(\frac{\pi_l}{\rho_0}+\frac{\pi_w}{\rho_0}\right)\frac{1}{3}\alpha_m+\left(\frac{\pi_t}{\rho_0}-1\right)\left(\frac{1}{3}\alpha_m-\eta^m\right)\right)\Delta T_m \qquad (6.4.8)$$

where φ, χ, η^m and η^o are given by (A6.3.4.3). In general the value of the various material parameters of the transducer such as π_l or ϵ, differ from the usual bulk values as a result of the production method. Calibrating measurements are necessary to obtain the specific values of the multiplication constants for strain and temperature change in (6.4.8).

For a strain transducer, all temperature influences have to be absent or compensated. In the case of small heat production by the electrical current, thus $\Delta T_o = \Delta T_m$, the transducer is self temperature compensated if

$$\frac{\epsilon}{\rho_0}+\left(\frac{\pi_t}{\rho_0}-1\right)\left(\frac{1}{3}\alpha_m+\eta^o-\eta^m\right)+\left(\frac{\pi_l}{\rho_0}+\frac{\pi_w}{\rho_0}\right)\frac{1}{3}\alpha_m = 0 \qquad (6.4.9)$$

For metals and ceramic materials α_m is of the order 10^{-5} K^{-1}. The piezoresistive parameters $\frac{\pi_l}{\rho_0}$, $\frac{\pi_w}{\rho_0}$ and $\frac{\pi_t}{\rho_0}$ are of the order 1 for many metals. For temperature fluctuations of substrate and transducer of about 10^{-1} K the resistance change due to dilatational effects will be of the order of the change due to microstrains. The thermoresistive parameter $\frac{\epsilon}{\rho_0}$ is only small for special alloys, which are often used for commercial strain gauges. The resulting deviation from (6.4.9) can be compensated using bridge type measuring methods.

For strain gauges the resistance strain temperature dependency is also given by equation (6.4.7). The strain of the substrate $(e_{A,m})_{ll}$ in the strain sensitive direction of the gauge, here called the length direction, is transmitted generally by a glue layer and a plastic carrier. Stress free contraction occurs in width and thickness direction, resulting theoretically in a simpler linear resistance strain dependency than for thin

film transducers. Due to the low stiffness of the glue, the substrate strain is in reality not completely transferred to the gauge, which would result in non-linear behaviour for simple strain gauge configurations. Special layouts are needed to obtain linearity. The glue layer also introduces non-linear behaviour for gauges under three-dimensional loading cases. This behaviour is difficult to quantify or compensate (Andreae 1974, Hoffmann 1987). For this reason such gauges are not applicable in the stress measuring tool. Evaporated transducers however are very suited because of the better resemblance of the elastic behaviour of matrix, strain transferring ceramic and transducer.

6.5 Analysis of the stress measuring tool design

In this paragraph the results previously derived are used to evaluate the design of the stress measuring tool. After a description of the assumptions with respect to the different tool parts, some calculations are performed to quantify the influence of systematic errors as well as stochastic errors. Systematic errors could be introduced by the unknown contact behaviour between the tool and tool holder, and by the glue layer with only approximately known elastic behaviour and thickness. It is shown that these errors remain within 10 % in the present design. The stochastic errors are of importance for the reconstruction of the contact stresses from the measured strain data. A reliable prediction of the discretized contact stresses (see paragraph 6.2.2) turns out to be possible if the number of stress unknowns is taken considerably smaller than the number of transducers.

In figure 6.5.1 the different parts of the construction are indicated. Also the element mesh for the tool body is given. The body is constructed of a normal tool steel (X210Cr12) with Young's modulus $2.12*10^5$ N/mm^2 and Poisson's ratio 0.28. The lay-on plate material may be constructed of some different material, in the present design also X210Cr12 is chosen. The thickness is restricted by the need for small systematic errors and a high sensitivity of the strain transducers. From calculations, of which some results are presented in the sequel, it followed that a thickness of about 2 mm is suitable. Before glueing the lay-on plate to the tool body, the plate should be provided with evaporated strain transducers, with the longitudinal axis of the active part in radial and tangential direction. A possible layout is visualized in figure 6.5.2. In the calculations it is supposed that 16 radial and 16 tangential transducers are used, with positions starting at a minimum radius of 2.5 mm and a mutual radial distance of 2.5 mm.
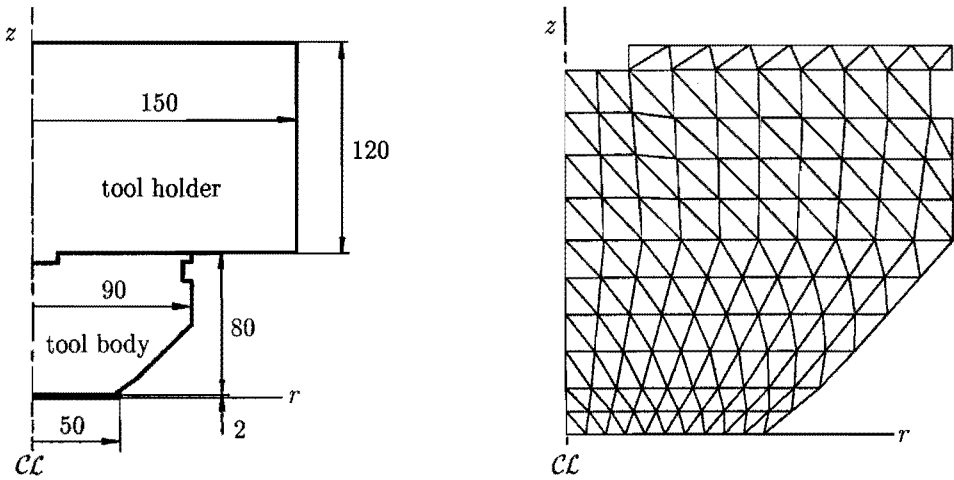
Figure 6.5.1 The construction analysed and element mesh of the tool body
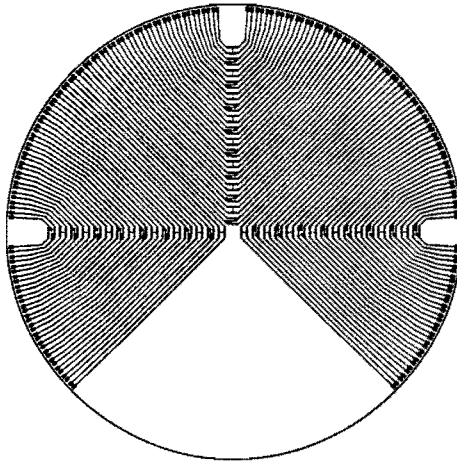
Figure 6.5.2 Possible layout of the transducer pattern

Because of the small temperature effects, as indicated in the previous paragraph, a suitable choice for the transducer material is Ni/Cr 80/20. For the time being the transducers are assumed ideal piezoresistive with isotropic elastic behaviour equal to the bulk behaviour of Ni/Cr 80/20. With the property values $E_o = 2.19*10^5$ N/mm^2, $\nu_o = 0.29$ and $\frac{\pi}{\rho_0} = 1.66$, from (6.4.8) can be derived

$$\frac{\Delta R}{R_0} = 2.63(e_{A,m})_{ll} + 0.63(e_{A,m})_{ww} + 0.62(e_{A,m})_{tt} \qquad (6.5.1)$$

with $(e_{A,m})_{ll}$, $(e_{A,m})_{ww}$ and $(e_{A,m})_{tt}$ the strain components of the lay-on plate in the length, width and thickness direction of the transducer. The lay-on plate is glued to the tool body by an epoxy resin adhesive, with an approximate Young's modulus of $4*10^3$ N/mm^2 and Poisson's ratio 0.35. Calculations in which the glue layer is varied in thickness show that for a small influence of the layer on the response of the transducers a thickness smaller than about 0.01 mm is necessary. In the calculations the loading of the tool is prescribed as visualized in figure 6.5.3.

Figure 6.5.3 The normal stress σ_n and shear stress τ at the contact surface

Figure 6.5.4 The theoretical resistance changes

Assuming the tool holder to be rigid, no slip between tool body and tool holder and the absence of the glue layer, the theoretical relative resistance change of the transduce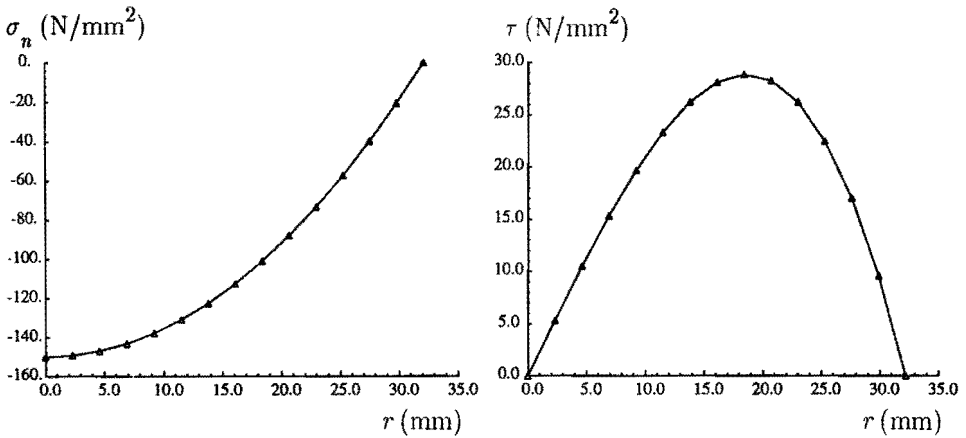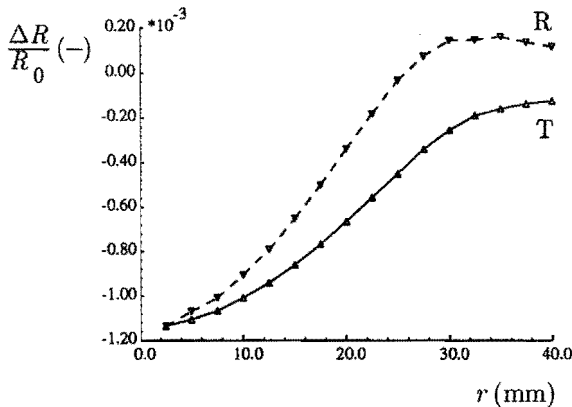rs with the longitudinal axis in radial (R) and tangential (T) direction is presented in figure 6.5.4. It appears that the resistance changes are very well measurable by bridge type methods. The responses in reality will be affected by systematic errors. Assuming frictionless slip between tool and tool holder results in the relative differences of the transducers responses with respect to the theoretical resistance changes, as given in figure 6.5.5 (a). In figure 6.5.5 (b) the relative differences due to the assumption that the tool holder has the same elastic behaviour as the tool body, are shown. Relative response differences of about 2.5% result.
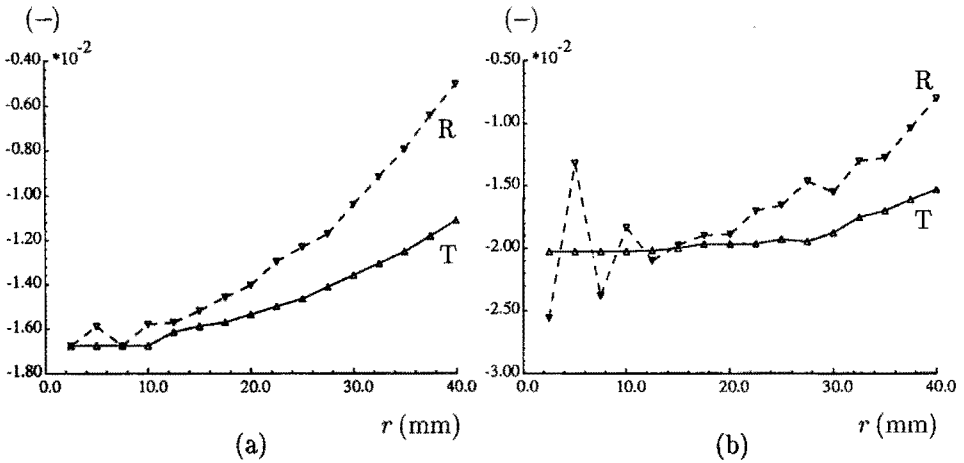
Figure 6.5.5 The relative response differences caused by frictionless slip between tool and tool holder (a) and by an elastic tool holder (b)

The relative response differences due to a glue layer of 0.01 mm thickness are given in figure 6.5.6. Maximum relative response differences of about 8% result. Although a thickness of 0.01 mm and even smaller can be established quite well, problems may occur with the insulation of the transducers to the tool body. Therefore an insulating coating or application of an insulating polycarbonate foil of for instance 0.002 mm thickness is necessary. The above results indicate that an upper-bound for the systematic errors of about 10% is possible with the present design. For the sequel calculations in this paragraph these errors will be neglected. For future research it will be necessary to study their influence in more detail and to account for this influence in the prediction of the contact stresses.
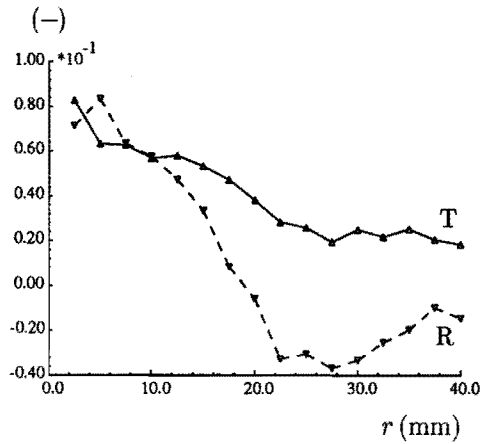
Figure 6.5.6 Relative response differences due to a 0.01 mm thick glue layer

Response data of the transducers are used as input to reconstruct the normal and frictional shear stress at the contact side. For this purpose the 32 responses are supposed to have a standard deviation of $1*10^{-6}$. The number of stress unknowns at the contact side equals 30. As the response data are only disturbed by small truncation errors, the estimates calculated for the normal and shear stress appear to be equal to the initial normal and shear stress values.
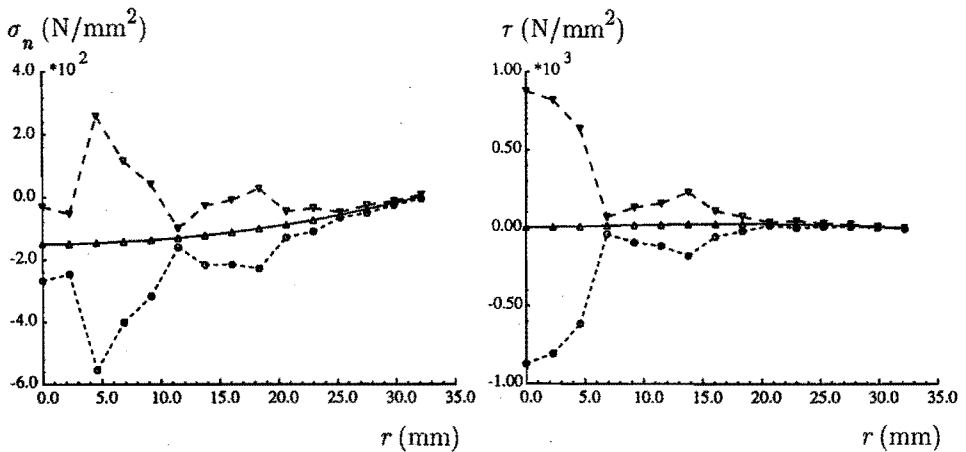
Figure 6.5.7 95% Confidence intervals for normal and shear stress

The covariance estimates for the normal and shear stress, calculated according to paragraph 6.2.3, are used to produce 95% confidence intervals. The confidence intervals calculated for the current case are given in figure 6.5.7. Reliable results are obtained only at the outer radius of the contact zone, which is quite unsatisfactory. In an attempt to improve the accuracy of the results, one may propose to take data of strain gauges at the tool body or the total upsetting force into account. The first proposal is performed using 24 strain gauges with a standard deviation of $1*10^{-6}$ for their relative resistance changes, the second with an assumed standard deviation of 250 N. The total upsetting force in the current case is 245 kN. The first method delivers an insufficient improvement, the second almost none, as can be seen from the confidence intervals for the normal stress in figure 6.5.8.
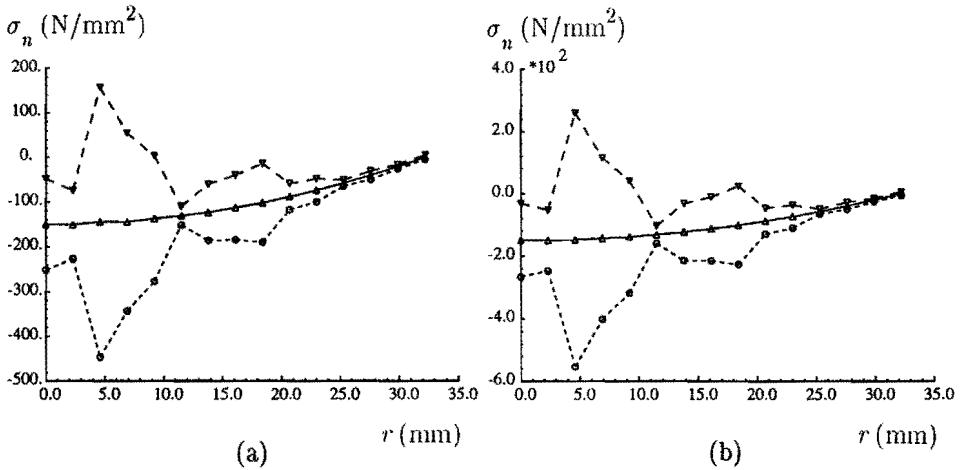
Figure 6.5.8 95% Confidence intervals for the normal stress with application of strain gauges (a) or the upsetting force (b)

A large improvement is obtained when the number of contact stress unknowns is reduced with respect to the number of transducers. The number of unknowns is reduced to 10, by demanding the normal as well as the shear friction stress to be polynomial expansions of fourth degree in the radial coordinate r along the contact zone. This is established by prescription of (linear) relationships between contact nodal stresses. The calculated estimates for the stresses then very closely resemble the original stress pattern and show a satisfying small 95% confidence interval, see figure 6.5.9.

σ_n (N/mm^2) τ (N/mm^2)

r (mm) r (mm)
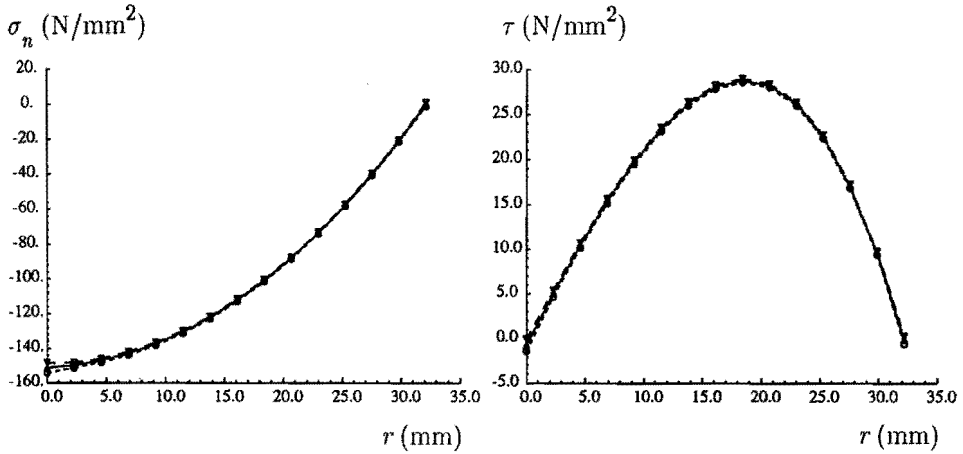
Figure 6.5.9 95% Confidence intervals for a reduced set of stress unknowns

That the original stress pattern is well reconstructed is not surprising as the normal and shear stress, given in figure 6.5.3, are polynomial expansions of second and third degree respectively. The credibility of the estimated contact stress pattern in reality, when the real pattern is unknown, can be judged by the χ^2-test with respect to the expression (A6.2.1.2). A nearer description of the application of this test is given in paragraph 5.4.

Some calculations are performed with a reduced set of transducers. With 8 radial and 8 tangential transducers, equidistantly distributed in radial direction, the confidence intervals of figure 6.5.9 become approximately twice as large. However, with 16 equidistant tangential transducers only, the confidence intervals become much larger (about 40 N/mm^2 for σ_n and 20 N/mm^2 for τ). With 16 equidistant radial transducers only, the confidence intervals remain smaller (about 20 N/mm^2 and 5 N/mm^2 respectively). Although obviously the radial transducers deliver more information about the contact stresses than the tangential transducers, transducers in both orientations are necessary for a reliable reconstruction of the contact stress pattern.

From the previous the following is concluded. In the present design the systematic errors are smaller than 10%, which can be improved in the future by taking their influence into account. A reliable reconstruction of the contact stress pattern seems possible if the number of stress unknowns is significantly smaller than the number of

transducer signals. A nearer investigation is needed to determine the optimal value of the ratio of stress unknowns to transducer signals.

6.6 Calibration of the transducers and the measuring tool

In order to achieve a prediction of the behaviour of the tool according to the current design, the resistance strain dependency of the (Ni/Cr 80/20) transducers on the lay-on plate has to be established. Further it is needed, after assembly, to investigate the resulting tool response in order to evaluate the true influence of the neglected systematic errors.

According to paragraph 6.4, the resistance changes of the transducers depend on three substrate strains. To establish the three relevant constants, see (6.4.8)

$$\frac{\pi_l}{\rho_0}+1+(\frac{\pi_t}{\rho_0}-1)\varphi \quad , \quad \frac{\pi_w}{\rho_0}-1+(\frac{\pi_t}{\rho_0}-1)\varphi \quad , \quad (\frac{\pi_t}{\rho_0}-1)\chi \qquad (6.6.1)$$

for each lay-on plate transducer, in principle it has to be subjected to three independent strain states. As this can not be performed, an alternative approach is proposed, based on two assumptions with respect to the constants in (6.6.1). The validity of these assumptions can be investigated by experiments in which the strain state of the transducers can be prescribed easily, and thereafter used to calibrate the lay-on plate transducers in one experiment.

At first it is assumed that for Ni/Cr 80/20 transducers on a steel substrate the values of φ and χ may be taken equal to the isotropic bulk values as given by equation (6.3.7). Assuming Young's modulus $E_m = 2.12*10^5$ N/mm^2 and Poisson's ratio $\nu_m = 0.28$ for the steel substrate and $E_o = 2.19*10^5$ N/mm^2, $\nu_o = 0.29$ for the transducers it follows by approximation

$$\varphi = -0.04 \quad , \quad \chi = 0.94 \qquad (6.6.2)$$

Secondly it is supposed that the transducers will be ideal piezoresistive, so $\pi_l = \pi_w = \pi_t = \pi$. This will be true if no orientation in any direction of the thin film material occurs. A support for the above assumptions can be found in literature. For instance Holland (1961) reports that the temperature coefficient of the resistivity of baked evaporated nichrome films indicates that the composition is similar to that of

the evaporated alloy. A substrate heated during evaporation to about 600 K will benefit the diffusion of the nickel and chromium in the thin film, thus producing a more uniform distribution of the components. Heating of the substrate further results in transducers showing no resistivity changes during aging. Otherwise artificial aging after evaporation is necessary to produce stable resistivity levels (Belser 1957, Stein 1962). With the assumptions the resistivity strain dependency is simplified to

$$\frac{\Delta R}{R_0} = (0.96\frac{\pi}{\rho_0} + 1.04)(e_{A,m})_{ll} + (\frac{\pi}{\rho_0} - 1)\left(0.96(e_{A,m})_{ww} + 0.94(e_{A,m})_{tt}\right) \qquad (6.6.3)$$

To investigate whether equation (6.6.3) is useful in practice, experiments with transducers on a substrate under uniaxial and uniform pressure are most easily to perform. Some premature results of such experiments indeed suggest that equation (6.6.3) is suitable to describe the true resistance strain dependency. In these experiments however no temperature compensation was applied which disturbed the results in an unknown way. Future experiments should be executed using a bridge measuring device e.g. according to Bolk (1982), thus enabling temperature compensation for a range of transducer resistances using reference transducers. With (6.6.3) the calibration of the transducers on the lay-on plate easily can be performed in a uniform pressure experiment, thus delivering a value for $\frac{\pi}{\rho_0}$ per transducer. Such a calibration is needed as the reproducibility of equal thin film transducers still is somewhat disappointing nowadays (Bethe 1988, Meijer 1988).

After assembly of the tool an experiment is needed to compare the actual response to the predicted one. Very suited for this are Rastegaev experiments. With such experiments the tool can be loaded with an almost uniform normal stress. It is recommendable to perform such evaluating experiments in the future using different diameters for the Rastegaev workpieces.

7 Conclusions and recommendations

In this chapter conclusions which can be drawn from the research performed are presented and recommendations for the continuation are indicated.

In this thesis theoretical and experimental tools are described with the aim to assess the applicability of extended friction models and to quantify them in the case of quasi-static mild cold forming contacts. These tools, successively considered in the sequel, include the thermodynamical framework in chapter 2, the experimental-numerical method in chapter 4 and the measuring methods for contour displacements and contact stresses for the upsetting configuration in chapter 5 and chapter 6. With respect to the quantification of the contact behaviour in the upsetting configuration, it is expected that this will be possible by the use of the stress measuring tool. Improvement of the tools, and extension of their applicability in the future is regarded.

Maxwell friction models show to fit within the thermodynamical framework under restrictions that are easily accommodated. To obtain more severe restrictions for constitutive models it is needed to specify the different thermodynamical quantities in more detail. Results from structural modelling of the real contacts may be used to arrive at such a specification. The main problem in a structural approach is that contact models on a microscopic scale are needed. Hence also results from research on the microscopic physical behaviour of contacting surfaces should be available to obtain more restrictive statements.

The experimental-numerical method is an elaboration for the geometrical and physical non-linearities, based on the maximum likelihood method in combination with the finite element method. Simulations with this method of the upsetting experiment on cylindrical test pieces show that use of the upsetting displacement and measured displacements of the visible unloaded surface only, is insufficient to determine reliable contact quantity values. Taking the upsetting force into account also hardly improves the results. With the upsetting displacement and the normal contact stress distribution, assumed to be measured with a conservatively low accuracy, much better results are obtained, while accurate contour data further increase the reliability. To obtain useful results from the upsetting test on cylindrical test pieces for the quantification of contact models, an increase of the predicted reliability of the friction stress is needed. A possibility to reach this increase is the use of assumptions with respect to the contact stress pattern. The credibility of these assumptions can be determined by statistical tests. The

reliability of the friction shear stress further can be increased by the application of a measuring method, more accurate than assumed for the calculations, or by using a method for measuring the friction stress itself. The last two suggestions are further investigated. An interesting feature offered by the simulation method is the possibility to obtain specifications for upsetting tests, improved with respect to the measuring accuracy in a different way than indicated above. It may be possible that for special shaped workpieces, other than the cylindrical workpieces analysed, contour displacements result in a higher accuracy for the contact quantities. It can be thought of hyperboloidal or ring-shaped workpieces. Also workpieces with indented contact surfaces in combination with non-flat tool contact surfaces may be considered. Also other forming processes can be examined in order to come to a specification of useful experimental set-ups. An example may be the hemispherical stretching. of sheet material. Extension of the experimental-numerical method to other than axisymmetric configurations further widens the class of set-ups, e.g. the draw-bending experiment of sheet material.

The photogrammetric set-up for the measurement of the contour displacements is theoretically not able to deliver the accuracy as required from the upsetting simulations indicated above. Although improvement of this accuracy is quite well possible by using a more sophisticated set-up and more optimal workpiece mark patterns and mark design, this is hardly significant for the upsetting test. Experimental evaluation shows that disturbances due to non-axisymmetric behaviour and surface roughening (caused by global and local inhomogeneities) can occur, much larger than the required displacement accuracy. This does not reject the axisymmetric upsetting test for quantification of contact models, as the deviations from axisymmetry for the contact stress pattern and radial displacement of contact points may be quite large before the influence is really disturbing. The contour displacement measurement remains of importance for assessing the dimension of the contact area with the tools. Further, an (improved) set-up can be used for measuring displacements in a contactless way in other forming processes such as sheet metal forming.

The stress measuring tool using embedded strain transducers is analysed by a numerical model of the mechanical behaviour of the tool and the mechanical and physical behaviour of the transducers. Systematic errors of about 10 % are not imaginary, mainly caused by the glue layer connecting the lay-on plate to the tool body. These are so large that it is needed to study these errors nearer in order to reduce or compensate for their influence. For the reduction of systematic errors

thinner glue layers are necessary. If the glue layer thickness and behaviour is well-known or can be measured, the influence of the layer can be accounted for. Otherwise Rastegaev tests with the tool may quantify the layer's influence and the results can be used for compensation of the systematic errors. Neglecting the systematic errors, the contact stress pattern can be reconstructed with a satisfying reliability if the number of stress unknowns is significantly smaller than the number of transducer signals. Radial as well as tangential transducers are necessary for this reconstruction. The numerical method, by which the stress measuring tool is analysed, may be used for the analysis of tools for measuring other quantities and with other application purposes than in the present study. It can be thought of force transducers as an example. For the stress measuring tool an insulating ceramic layer or other non-conducting layer has to be applied to insulate the transducers from the tool body. The most simple solution for this is to provide the transducers with a protective coating. The resistance strain dependency for Ni/Cr 80/20 transducers has to be examined further. A complicating factor may be that the composition of the transducer material cannot be maintained constant for different batches. It is therefore recommendable to study the influence of composition variations on the behaviour of the transducers and to incorporate this in the resistance strain dependency, in order to increase the calibration accuracy.

The present research delivers a number of methods to investigate the mechanical interaction between die and workpiece in forming processes further in the future. However, completion of the stress measuring method and application in experiments, extension of the methods to other set-ups than the upsetting test, as well as combination with results from associated tribological research, such as e.g. structural investigations, are indispensable for a full understanding. Despite the extensive research in the past, considerable effort is still required before the classification and quantification of the contact behaviour between two contact partners under well-defined lubrication conditions, will have become a standard procedure.

Appendices

Appendix 2.2.1 The balance law for contact surfaces

In this appendix the balance law (2.2.5) is derived for some mass associated quantity φ. For that purpose the averaging volume V_E belonging to a point $\underset{\sim}{\eta}$ is followed. Considered is a quantity Q defined on this volume

$$Q = \int_{V_E} \rho\varphi dV = A_E <\rho\varphi>_A \qquad (A2.2.1.1)$$

with ρ the mass density. The change $Q^{\bullet} = A_E <\rho\varphi>_A^{\bullet}$ of Q in time is caused by material flow (convection), immaterial flow (flux) and sources.

1) Material flow through the border ∂V_E of V_E. The local material velocity is indicated by $\vec{v}$. The decrease of Q per unit of time due to convection then equals

$$\int_{\partial V_E} \rho\varphi(\vec{v}-\vec{u}_s).\vec{n}dA = \int_{A_E} \rho^+\varphi^+(\vec{v}^+-\vec{u}_s).\vec{n}^+ dA + \int_{A_E} \rho^-\varphi^-(\vec{v}^--\vec{u}_s).\vec{n}^- dA +$$

$$+ \int_{\pi D_E^t l_E} \rho\varphi(\vec{v}-\vec{u}_s).\vec{\nu}dA \qquad (A2.2.1.2)$$

with $\vec{n}^+$ and $\vec{n}^-$ the unit outward normal vectors on V_E at the S^+ and S^- surfaces respectively, parallel to $\vec{n}_s$, and $\vec{\nu}$ the unit outward normal vector on the cylindrical surface of V_E, perpendicular to $\vec{n}_s$. The convection term is decomposed into surface integrals for the surfaces S^+ and S^- and for the cylindrical surface. The superscripts $+$ and $-$ indicate bulk continuum quantities at the $+$ and $-$ side of the averaging volume V_E. The continuum quantities at S^+ and S^- vary with characteristic length L along these surfaces. Therefore the integrals over A_E may be written as

$$\vec{n}^+.(\vec{v}^+-\vec{u}_s)\rho^+\varphi^+ A_E \quad , \quad \vec{n}^-.(\vec{v}^--\vec{u}_s)\rho^-\varphi^- A_E \qquad (A2.2.1.3)$$

The integral over the cylindrical surface of V_E can (see appendix 2.2.2) be

replaced by

$$(\vec{\nabla}_s \cdot [(I - \vec{n}_s \vec{n}_s) \cdot <\vec{v}\rho\varphi>_A]) A_E - \vec{u}_s \cdot (\vec{\nabla}_s \cdot [(I - \vec{n}_s \vec{n}_s) <\rho\varphi>_A]) A_E \qquad (A2.2.1.4)$$

with $\vec{n}_s$ chosen equal to $\vec{n}^+$ or $\vec{n}^-$ arbitrarily. The expression in (A2.2.1.4) may be reformulated as

$$(\vec{\nabla}_s \cdot [(I - \vec{n}_s \vec{n}_s) \cdot <\vec{v}\rho\varphi>_A]) A_E - (\vec{\nabla}_s \cdot [(I - \vec{n}_s \vec{n}_s) \cdot \vec{u}_s <\rho\varphi>_A]) A_E +$$

$$+ (\vec{\nabla}_s \cdot \vec{u}_s) A_E <\rho\varphi>_A \qquad (A2.2.1.5)$$

2) Flux through the boundary ∂V_E. The local flux is indicated by $\vec{\psi}$. The decrease of Q resulting from fluxes is

$$\int_{\partial V_E} \vec{n} \cdot \vec{\psi} dA = \int_{A_E} \vec{n}^+ \cdot \vec{\psi}^+ dA + \int_{A_E} \vec{n}^- \cdot \vec{\psi}^- dA + \int_{\pi D_E t_E} \vec{\nu} \cdot \vec{\psi} dA \qquad (A2.2.1.6)$$

Similarly as before the right hand side integrals may be transferred to respectively

$$\vec{n}^+ \cdot \vec{\psi}^+ A_E \quad , \quad \vec{n}^- \cdot \vec{\psi}^- A_E \quad , \quad (\vec{\nabla}_s \cdot [(I - \vec{n}_s \vec{n}_s) \cdot <\vec{\psi}>_A]) A_E \qquad (A2.2.1.7)$$

3) Sources within V_E. Only the case of a distributed mass associated φ–production Φ is considered. For the increase of Q holds

$$\int_{V_E} \rho\Phi dV = A_E <\rho\Phi>_A \qquad (A2.2.1.8)$$

Combining the above relations results in the general balance law (2.2.5), an equation only valid if the conditions (2.1.2) for the dimensions of the averaging volume are satisfied.

Appendix 2.2.2 A special surface integral

In this appendix surface integrals over the cylindrical surface of an elementary averaging volume V_E of the form

$$\int_{\pi D_E t_E} \vec{\nu}.q \, dA \qquad\qquad (A2.2.2.1)$$

are elaborated, with q a tensor of first or higher order and $\vec{\nu}$ the unit outward normal vector. By separating the integration over the surface into an integration over the thickness t_E of V_E and the circumference πD_E, (A2.2.2.1) can be reformulated as

$$\int_{\pi D_E} \vec{\nu}.(\int_{t_E} q \, dh) \, dl \qquad\qquad (A2.2.2.2)$$

with h a coordinate in thickness direction and l a coordinate in circumferential direction. Assuming the integral of q over t_E smooth, such that the divergence theorem on the surface may be applied, it follows

$$\int_{\pi D_E t_E} \vec{\nu}.q \, dA = \int_{A_E} \vec{\nabla}_s.[(I-\vec{n}_s \vec{n}_s).\int_{t_E} q \, dh] \, dA \qquad\qquad (A2.2.2.3)$$

Because of (2.1.2) and the definition of $<q>_A$ according to (2.1.1) it results

$$\int_{\pi D_E t_E} \vec{\nu}.q \, dA = (\vec{\nabla}_s.[(I-\vec{n}_s \vec{n}_s).<q>_A]) A_E \qquad\qquad (A2.2.2.4)$$

To derive (A2.2.2.4) a certain smoothness for the integral of q in thickness direction was assumed. The result suggests that this is a too strong demand, as the right hand side of (A2.2.2.4) is already properly defined if $<q>_A$ is continuously differentiable. According to the considerations of paragraph 2.1, this demand for $<q>_A$ is guaranteed for all relevant cases.

In the same way as above it shows for a scalar quantity q

$$\int_{\pi D_E t_E} \vec{\nu} q \, dA = (\vec{\nabla}_s.[(I-\vec{n}_s \vec{n}_s)<q>_A]) A_E \qquad\qquad (A2.2.2.5)$$

Appendix 2.4.1 Some aspects of objectivity

In this appendix the principle of frame indifference is shortly discussed. Considered are two observers O and $\overline{O}$, each with an own orthonormal vector base. The observers are supposed equivalent, i.e. with equal notion of time and distance. The time difference between O and $\overline{O}$ is arbitrarily chosen equal zero. Quantities observed by observer $\overline{O}$ are overstriked, quantities observed by O are not. As example the position vector of some point P in space as observed by O is indicated with $\vec{x}$, and as observed by $\overline{O}$ with $\vec{\overline{x}}$.
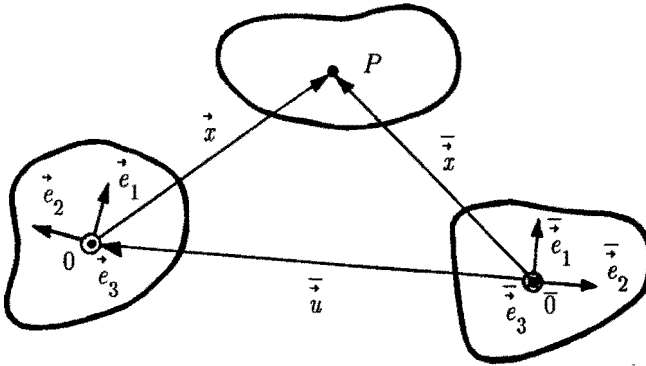
Figure A2.4.1.1 Two equivalent observers O and $\overline{O}$

The components of some vectorial quantity $\vec{v}$ with respect to the vector base of O combine into the column $\underset{\sim}{v}$, called the column representation of $\vec{v}$ with respect to $\{\vec{e}_1,\vec{e}_2,\vec{e}_3\}$. Equivalently the components of $\vec{\overline{v}}$ with respect to the vector base of $\overline{O}$ combine into the column representation $\underset{\sim}{\overline{v}}$. In the same way some second order tensorial quantity T has a matrix representation $\underset{\sim}{T}$ with respect to $\{\vec{e}_1,\vec{e}_2,\vec{e}_3\}$, and $\overline{T}$ a matrix representation $\underset{\sim}{\overline{T}}$ with respect to $\{\vec{\overline{e}}_1,\vec{\overline{e}}_2,\vec{\overline{e}}_3\}$.

The coordinates $\underset{\sim}{x}$ and $\underset{\sim}{\overline{x}}$ of some point P in space are related by Euclidian transformations. The transformation expressing $\underset{\sim}{\overline{x}}$ in $\underset{\sim}{x}$ reads

$$\overline{\underset{\sim}{x}}(t) = \overline{\underset{\sim}{Q}}(t)\underset{\sim}{x}(t) + \overline{\underset{\sim}{u}}(t) \qquad\qquad (A2.4.1.1)$$

with $\overline{\underset{\sim}{Q}}$ a time dependent orthonormal tensor and $\overset{\rightarrow}{\overline{u}}$ a time dependent vector. For an arbitrary time t, $\overline{\underset{\sim}{Q}}$ and $\overset{\rightarrow}{\overline{u}}$ indicate the rotation and the translation respectively, of O with respect to $\overline{O}$. Physical quantities are called objective if they are related in a defined way. For scalar, vector and second order tensor quantities the defining relations are

$$\overline{B} = B \quad , \quad \overline{\underset{\sim}{B}} = \overline{\underset{\sim}{Q}}\,\underset{\sim}{B} \quad , \quad \overline{\underset{\sim}{B}} = \overline{\underset{\sim}{Q}}\,\underset{\sim}{B}\,\overline{\underset{\sim}{Q}}^{T} \qquad\qquad (A2.4.1.2)$$

A physical quantity is called invariant if it is observed identical by two equivalent observers i.e. the representations for both observers are equal. A scalar objective quantity is also invariant. The principle of frame indifference states that constitutive equations should not change due to observer transformations. This is the case if they only contain objective or only invariant terms. Part of the principle is also the assumption that mass density, stress and force vector, heat flux vector, internal energy, entropy and heat sources are objective. It is noted that the balance laws of paragraph 2.3 do not change due to observer transformations, except for the part $\vec{n}{}^{+}.\overset{\rightarrow}{\sigma}{}^{+}.\vec{v}{}^{+} - \vec{n}{}^{-}.\overset{\rightarrow}{\sigma}{}^{-}.\vec{v}{}^{-}$ which can be formulated as $\overset{\rightarrow}{\sigma}{}^{+}.(\vec{v}{}^{+} - \vec{v}{}^{-})$ using (2.3.3) and $\overset{\rightarrow}{\sigma}{}^{+} = \vec{n}{}^{+}.\sigma^{+}$. The stress vector $\overset{\rightarrow}{\sigma}{}^{+}$ is objective, the velocity difference is this only in a restricted sense. For this velocity difference the observations of $\overline{O}$ and O are related by

$$\overline{\underset{\sim}{v}{}^{+}} - \overline{\underset{\sim}{v}{}^{-}} = \dot{\overline{\underset{\sim}{Q}}}(t)(\underset{\sim}{x}{}^{+} - \underset{\sim}{x}{}^{-}) + \overline{\underset{\sim}{Q}}(t)(\underset{\sim}{v}{}^{+} - \underset{\sim}{v}{}^{-}) \qquad\qquad (A2.4.1.3)$$

Because $t_E << L$ the first term in the right hand side is negligible, resulting in an equation satisfying objectivity. Only for transformations for which $\dot{\overline{\underset{\sim}{Q}}}(t) = 0$ objectivity is exactly guaranteed.

In paragraph 2.4 an invariant stress vector and contact layer deformation vector are introduced using some second order tensor $A = A(\eta,t)$. With the foregoing it shows that second order tensors A can be used, obeying the relationship

$$\overline{A}(\eta,t) = \overline{Q}(t)A(\eta,t) \quad , \quad \det(\underline{A}) \neq 0 \qquad \text{(A2.4.1.4)}$$

In order not to conflict with the accepted independent variables of paragraph 2.4 only tensors A containing kinematics of the surfaces S, S^+ and S^- are considered. An example of A, obeying (A2.4.1.4), is a linear combination of the inplane deformation tensors F_s, F^+, F^- and a contribution $\vec{n}_s\vec{n}_{s0}$ accounting for the rotation of the direction normal to the contact layer, with the subscript 0 referring to the reference state chosen

$$A = \alpha_s F_s + \alpha^+ F^+ + \alpha^- F^- + \beta \vec{n}_s \vec{n}_{s0} \qquad \text{(A2.4.1.5)}$$

The coefficients α_s, α^+, α^- and β may be functions of the surface change factors $\vec{V}_{s0}\cdot\vec{x}_s$, $\vec{V}_{s0}\cdot\vec{x}^+$ and $\vec{V}_{s0}\cdot\vec{x}^-$. In the case only quantities belonging to the surface S^+ are considered of importance, legitimate choices are

$$\alpha^+ F^+ + \beta \vec{n}^+ \vec{n}_0^+ \quad , \quad \alpha^+ (F^+)^{-c} + \beta \vec{n}^+ \vec{n}_0^+ \quad , \quad \alpha^+ R^+ + \beta \vec{n}^+ \vec{n}_0^+ \qquad \text{(A2.4.1.6)}$$

with R^+ such that $F^+ = R^+ \cdot U^+$ with

$$(U^+)^c = U^+ \quad , \quad R^+ \cdot (R^+)^c = I - \vec{n}^+ \vec{n}^+ \quad , \quad \det(R^+ + \vec{n}^+ \vec{n}_0^+) = 1 \qquad \text{(A2.4.1.7)}$$

Appendix 2.6.1 Objective rates for vectorial surface quantities

In paragraph 2.6 an objective rate for vectorial quantities $\vec{q}$ was resulting

$$\overset{\triangledown}{\vec{q}} = \dot{\vec{q}} + A.(\dot{A^{-1}}).\vec{q} \tag{A2.6.1.1}$$

Baaijens (1987) introduced objective rates of the type

$$\overset{\triangledown}{\vec{q}} = \dot{\vec{q}} - (\vec{\nabla}_s \vec{v}^+ + H^+)^c.\vec{q} \tag{A2.6.1.2}$$

with H^+ an objective second order tensor. Considered is the identification of the rates (A2.6.1.2) with (A2.6.1.1) for three choices of A as suggested in appendix 2.4.1. At first A is taken equal to $F^+ + \vec{n}^+\vec{n}_0^+$. Quite easily it follows

$$A.(\dot{A^{-1}}) = -(\vec{\nabla}_s \vec{v}^+ + \vec{n}^+\dot{\vec{n}}^+)^c \tag{A2.6.1.3}$$

Equality of the rates occurs if $H^+ = \vec{n}^+\dot{\vec{n}}^+$ is chosen. As A is not a rotation tensor, the constitutive equations formulated with the objective rate as in equations (2.6.8) and (2.6.9) will be more complex. This rate is a Truesdell rate type. A second choice for A is the rotation tensor $R^+ + \vec{n}^+\vec{n}_0^+$. It then follows that

$$A.(\dot{A^{-1}}) = -(R^+.(\dot{R}^+)^c + \vec{n}^+\dot{\vec{n}}^+)^c \tag{A2.6.1.4}$$

and thus equality of the rates occurs for $H^+ = \vec{n}^+\dot{\vec{n}}^+ - \vec{\nabla}_s\vec{v}^+ + R^+.(\dot{R}^+)^c$. The corresponding rate is a Dienes rate type. For his calculations Baaijens used $H^+ = 0$, completely legitimate because only quantities $\vec{q}$ perpendicular to $\vec{n}^+$ were considered. For A then has to be chosen the non–regular tensor $(I - \vec{n}^+\vec{n}^+).(F^+ + \vec{n}^+\vec{n}_0^+)$ with its pseudo-inverse $(F^+ + \vec{n}^+\vec{n}_0^+)^{-1}.(I - \vec{n}^+\vec{n}^+)$.

Appendix 4.3.1 The maximum likelihood estimate

In this appendix some notes with respect to the maximum likelihood estimate are given. At first two important concepts with respect to estimators in general are reviewed.

1) Bias :

The bias $\underset{\sim}{b}$ of an estimator for some state $\underset{\sim}{s}$ is defined as

$$\underset{\sim}{b} = E(\underset{\sim}{\hat{s}}) - \underset{\sim}{s}_t \qquad \qquad (A4.3.1.1)$$

with the expected value operator E according to (4.3.8) and $\underset{\sim}{s}_t$ the unknown true state. If $\underset{\sim}{b} = \underset{\sim}{0}$ the estimator is unbiased, otherwise biased.

2) Efficiency :

An unbiased estimator is called efficient if there exists no other unbiased estimate for which the error covariance matrix is smaller (i.e. the covariance matrix of some unbiased estimate subtracted from the covariance matrix of the efficient estimate is always positive semi-definite).

As stated in paragraph 4.3, the maximum likelihood estimator is under fairly general conditions asymptotically unbiased and efficient. This means that the estimate becomes unbiased and efficient if it is based on an increasing amount of observation data.

An approximation for the state error covariance matrix $\underset{\sim}{\Sigma}$ can be given if the estimate is assumed unbiased and not too far from the true state

$$\underset{\sim}{\Sigma} = E\left((\underset{\sim}{\hat{s}} - \underset{\sim}{s}_t)(\underset{\sim}{\hat{s}} - \underset{\sim}{s}_t)^T\right) \approx \underset{\sim}{Z}_e (\underset{\sim}{Z}_e^T \underset{\sim}{J}_m^T \underset{\sim}{R}_m^{-1} \underset{\sim}{J}_m \underset{\sim}{Z}_e)^{-1} \underset{\sim}{Z}_e^T \quad ,$$

$$\underset{\sim}{J}_m^T = \frac{\partial}{\partial \underset{\sim}{s}}(\underset{\sim}{h}_m^T) \quad , \quad \underset{\sim}{Z}_e \underset{\sim}{J}_e^T = \underset{\sim}{0} \quad , \quad \underset{\sim}{J}_e^T = \frac{\partial}{\partial \underset{\sim}{s}}(\underset{\sim}{h}_e^T) \qquad (A4.3.1.2)$$

where $\underset{\sim}{J}_m$ and $\underset{\sim}{J}_e$ are the Jacobian matrices of the measurable state dependent quantities $\underset{\sim}{h}_m$ and exact state dependent quantities $\underset{\sim}{h}_e$ respectively, belonging to the state $\underset{\sim}{s}$. The columns of $\underset{\sim}{Z}_e$ span the orthogonal complement of the space spanned by the rows of $\underset{\sim}{J}_e$. The matrix $\underset{\sim}{Z}_e$ equals the unit matrix if no exact state dependent quantities apply. An approximate covariance matrix $\underset{\sim}{\Sigma}_q$ for quantities $\underset{\sim}{q}$, non-linear in the state $\underset{\sim}{s}$ such as stresses and forces, can be gained from

$$\underline{\Sigma}_q = E\left((\hat{\underline{q}} - \underline{q}_t)(\hat{\underline{q}} - \underline{q}_t)^T\right) \approx \underline{M}_q \underline{\Sigma} \, \underline{M}_q^T \quad , \quad \underline{M}_q^T = \frac{\partial}{\partial \underline{s}}(\underline{q}^T) \tag{A4.3.1.3}$$

where $\hat{\underline{q}}$ follows by replacement of $\underline{s}$ by $\hat{\underline{s}}$ in the expression for $\underline{q}$. The matrices $\underline{Z}_e$, $\underline{J}_m$ and $\underline{M}_q$ may be evaluated in $\hat{\underline{s}}$ if $\hat{\underline{s}}$ is not too far from $\underline{s}_t$. If $\hat{\underline{s}}$ is biased, at least a term $\underline{b}\underline{b}^T$ could be added to the estimate for $\underline{\Sigma}$. Generally $\underline{b}$ is however unknown. In that case, as well as for large errors in the measurements, a Monte Carlo simulation can be used to evaluate the estimator performance (Schweppe 1973).

Appendix 4.3.2 A strategy to deal with sudden changes in material behaviour in estimation problems

In this appendix the method used by Jansen (1987[2]) to deal with material showing sudden changes in mechanical behaviour, is described. Jansen developed an estimator for reconstructing the deformation process during a uni-axial tension test on a test bar with known material properties. The input of the estimator consisted of paired measurements of tension force and elongation for a number of successive time steps.

In the case of tension tests during which no unloading occurs, the real deformation process is very well reconstructed by a maximum likelihood estimator, for the elastic as well as for the elasto-plastic part of the process. This is the case for both small and large values of the observation errors with respect to the true changes of the state dependent quantities. No significant bias errors are established. Overestimation of the history parameter (i.c. the plastic or irreversible part of the strain), caused by measurement errors at some time step, is apparently corrected at succeeding time steps by the estimator. The transition from elastic to elasto-plastic is supposed to occur when the probability density of the measurements at a given time step is larger assuming elasto-plastic behaviour than assuming elastic behaviour.

A tension test performed with alternating loading and unloading necessitates an extra decision algorithm to determine the time steps at which the material behaviour changes from elasto-plastic (loading) into elastic (unloading). If such an algorithm is not used, too frequently unloading would be assumed erroneously. A procedure delivering reliable results uses the likelihood ratio test to choose between the two material models. The procedure of Jansen is summarized in the sequel.

1) At first it is assumed that a given time step is executed in the elasto-plastic region. The result based on this assumption further will be called an elasto-plastic estimate. If the probability density p_{1p} of the elasto-plastic estimate for the next time step is larger than some critical value β_0, that estimate is supposed to be an estimate for the true state of the tension bar.

2) If however p_{1p} is smaller than β_0, also an estimate supposing elastic behaviour is calculated, further called an elastic estimate. The probability density p_{1e} of that estimate can be very high, as observed data can equal exactly the state dependent quantities. It is supposed that the tension test is performed such that there will be at least two time steps between changes from loading into

unloading and vice versa. A decision between elastic and elasto-plastic behaviour is postponed to the second time step.

3) If the probability density p_{2_p} of the elasto-plastic estimate for the second time step is also larger than β_0, the elasto-plastic estimates are supposed to be estimates for the true states. If p_{2_p} is smaller than β_0, the probability density p_{2_e} of the elastic estimate on the unloading path for the elastic estimate of the first time step is calculated. With p_{2_e} and p_{2_p} the likelihood ratio λ

$$\lambda = \frac{p_{2_e}}{p_{2_p}} \tag{A4.3.2.1}$$

is determined. This ratio originates from the theory of hypothesis testing and is in the case considered used to decide between the hypothesis of elastic and elasto-plastic material behaviour. If λ is larger than some critical value β_e, the elastic estimates are supposed to be valid and an improved unloading path with the measurements of the two time steps can be calculated. If λ is smaller than some critical value β_p, the elasto-plastic estimates are supposed to be valid. For values of λ between β_p and β_e, the decision is based on data of a third time step. Choices for the critical values β_0, β_p and β_e depend on considerations with respect to computation time, variances of the observation errors, magnitude of the changes of the true values of the state dependent quantities per time step, the loading and unloading frequency and measuring frequency. For details it is referred to Jansen (1987[2]). To the above scheme indicated roughly, refinements can be made to improve the estimates. These refinements will not be discussed here.

Appendix 5.3.1 Interpolation for the contour functions

The special interpolation for the functions R and Z as introduced in paragraph 5.3, preserving volume invariance, is described. First the dimensionless quantities ζ_0, $\bar{\rho}$ and ζ are introduced by

$$\zeta_0 = 2\frac{\zeta_0}{h_0} \quad , \quad \bar{\rho}(\zeta_0) = \frac{R(\zeta_0)}{R_0} \quad , \quad \zeta(\zeta_0) = 2\frac{Z(\zeta_0)}{h_0} \tag{A5.3.1.1}$$

with h_0 and R_0 the initial height and radius of the workpiece. The demand of volume invariance then can be stated as

$$\frac{1}{2} \int_{\zeta_0=-1}^{1} \bar{\rho}^2 \frac{d\zeta}{d\zeta_0} d\zeta_0 = 1 \tag{A5.3.1.2}$$

This demand is obeyed by applying (spherical) Legendre polynomials to approximate $\bar{\rho}^2$ and $\dfrac{d\zeta}{d\zeta_0}$

$$\bar{\rho}^2 = \sum_{i=0}^{n} \alpha_i P_i(\zeta_0) \quad , \quad \frac{d\zeta}{d\zeta_0} = \sum_{i=0}^{m} \beta_i P_i(\zeta_0) \tag{A5.3.1.3}$$

where for P_i holds

$$\int_{\xi=-1}^{1} P_q(\xi) P_r(\xi) d\xi = 0 \quad , \quad \int_{\xi=-1}^{1} P_q^2(\xi) d\xi = \eta_q \quad , \quad P_q(1) = 1 \tag{A5.3.1.4}$$

with $\eta_q = \dfrac{2}{2q+1}$ and $q,r \in \{0,1,2,...\}$, $q \neq r$ (see e.g. Abramowitz & Stegun 1965). Elaborating (A5.3.1.2) using (A5.3.1.3) and (A5.3.1.4) a condition for the parameters α_i and β_i appears to be

$$\sum_{k=0}^{\min(n,m)} \eta_k \alpha_k \beta_k = 1 \tag{A5.3.1.5}$$

It is always possible to choose the interpolation such that α_0 is unequal zero. In that case β_0 can be expressed as

$$\beta_0 = (1 - \sum_{k=1}^{\min(n,m)} \eta_k \alpha_k \beta_k) / \eta_0 \alpha_0 \qquad (A5.3.1.6)$$

Appendix 6.2.1 A maximum likelihood estimate for the linear elastic body

The expression (6.2.18) for the estimate $\hat{u}$ of the state $\underset{\sim}{u}$ is derived as a maximum likelihood estimate.

For the observation errors $\underset{\sim}{v}$ a (multivariate) normal distribution density function with mean $\underset{\sim}{0}$ and covariance matrix $\underline{R}$ are supposed to hold. The likelihood function $p(\underset{\sim}{z}_m : \underset{\sim}{u})$ for the measured observation data given some state $\underset{\sim}{u}$, follows as

$$p(\underset{\sim}{z}_m : \underset{\sim}{u}) = \left((2\pi)^{n_m} \det(\underline{R}) \right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2} (\underset{\sim}{z}_m - \underline{H}_m \underset{\sim}{u})^T \underline{R}^{-1} (\underset{\sim}{z}_m - \underline{H}_m \underset{\sim}{u}) \right) \qquad (A6.2.1.1)$$

with n_m the length of $\underset{\sim}{z}_m$. The maximum likelihood estimate $\hat{u}$ for $\underset{\sim}{u}$ is that particular $\underset{\sim}{u}$ which maximizes $p(\underset{\sim}{z}_m : \underset{\sim}{u})$ for the actual measured observation data. Because the state $\underset{\sim}{u}$ is constrained by $\underset{\sim}{z}_e = \underline{H}_e \underset{\sim}{u}$, the estimate $\hat{u}$ follows from minimizing J according to

$$J = \frac{1}{2} (\underset{\sim}{z}_m - \underline{H}_m \underset{\sim}{u})^T \underline{R}^{-1} (\underset{\sim}{z}_m - \underline{H}_m \underset{\sim}{u}) + \underset{\sim}{\lambda}^T (\underset{\sim}{z}_e - \underline{H}_e \underset{\sim}{u}) \qquad (A6.2.1.2)$$

with respect to $\underset{\sim}{u}$ and the Lagrangian multipliers $\underset{\sim}{\lambda}$. It results

$$\begin{bmatrix} \underline{0} & \underline{H}_e \\ \underline{H}_e^T & -\underline{H}_m^T R^{-1} \underline{H}_m \end{bmatrix} \begin{bmatrix} \underset{\sim}{\lambda} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} \underset{\sim}{z}_e \\ \underline{H}_m^T R^{-1} \underset{\sim}{z}_m \end{bmatrix} \qquad (A6.2.1.3)$$

The solution for $\hat{u}$ from (A6.2.1.3) is in conformity with equation (6.2.18).

Appendix 6.3.1 The constrained strain field for arbitrary homogeneities

The relation (6.3.1) will be derived. This relation expresses the constrained strain $e_C(\vec{x})$ at some position $\vec{x}$ inside or outside a homogeneity in the transformation strain e_T.

At the boundary of the isolated homogeneity a load $\sigma_T.\vec{n}$ is needed in order to restore the original shape, where the stress tensor σ_T is related to e_T by the generalized Hooke's law, $\vec{n}$ is the unit outward normal vector. Hooke's law is written as $\sigma_T = {}^4L\!:\!e_T$ while in the isotropic case ${}^4L$ is depending on the shear modulus G and Poisson's ratio ν only. After connection this load is relaxed.

The displacement $\vec{u}$ at some point $\vec{x}$ of an infinite, homogeneous and isotropic linear elastic body, due to a point load $\vec{F}$ with the point of application $\vec{y}$, is given by (Love 1944)

$$\vec{u}(\vec{x},\vec{y}) = A(\vec{x},\vec{y}).\vec{F} \quad , \quad A = \frac{1}{16\pi G(1-\nu)}\frac{1}{r}\Big((3-4\nu)I+\vec{\epsilon}\vec{\epsilon}\Big) \quad ,$$

$$r = \|\vec{y}-\vec{x}\| \quad , \quad \vec{\epsilon} = \frac{\vec{y}-\vec{x}}{r} \tag{A6.3.1.1}$$

With this relation the so-called constrained displacement $\vec{u}_C$ after relaxation of $\sigma_T.\vec{n}$ can be determined. It follows, using the divergence theorem

$$\vec{u}_C(\vec{x}) = -\int_{\partial V} A.(\sigma_T.\vec{n})\,dA = -\int_V {}^3G(\vec{x},\vec{y})\,dV\!:\!e_T \quad ,$$

$$^3G = \vec{\nabla}_y.A.{}^4L = \frac{1}{8\pi(1-\nu)r^2}\Big((1-2\nu)(\vec{I}\vec{\epsilon}+(\vec{I}\vec{\epsilon})^{rc}-\vec{\epsilon}I)+3\vec{\epsilon}\vec{\epsilon}\vec{\epsilon}\Big) \tag{A6.3.1.2}$$

with ∂V the boundary and V the volume of the homogeneity. Notice that the integrations over ∂V and V are performed with $\vec{x}$ fixed, while $\vec{y}$ is a local variable. The gradient operator $\vec{\nabla}_y$ has $\vec{y}$ as differentiation variable. For the constrained strain $e_C(\vec{x}) = \frac{1}{2}\Big(\vec{\nabla}_x\vec{u}_C+(\vec{\nabla}_x\vec{u}_C)^c\Big)$ results

$$e_C(\vec{x}) = {}^4M(\vec{x}){:}e_T \quad , \quad {}^4M = -\frac{1}{2}\Big(\vec{\nabla}_x \int_V {}^3GdV + (\vec{\nabla}_x \int_V {}^3GdV)^{lc}\Big) \qquad \text{(A6.3.1.3)}$$

For large d, with d the distance of $\vec{x}$ to the centre of the homogeneity, e_C is proportional to d^{-3}, so approaches 0.

Appendix 6.3.2 The constrained strain field for ellipsoidal homogeneities

The expression for ^4S of equation (6.3.2) is derived for the case of an ellipsoidal homogeneity. For some fixed point $\vec{x}$ inside an ellipsoidal homogeneity, the volume integrals in (A6.3.1.3) can be elaborated

$$\int_V {}^3G(\vec{x},\vec{y})\,dV = \int_{4\pi} {}^3G^*(\vec{\epsilon})\,d\omega \int_0^{r^*(\vec{x},\vec{\epsilon})} dr = \int_{4\pi} {}^3G^*(\vec{\epsilon})\,r^*(\vec{x},\vec{\epsilon})\,d\omega \qquad (A6.3.2.1)$$

An infinitesimal volume dV is taken as $r^2\,d\omega\,dr$ with $d\omega$ an infinitesimal surface part of the unit sphere with its centre at position $\vec{x}$. The third order tensor $^3G^*$, defined by $^3G^*(\vec{\epsilon}) = {}^3G(\vec{x},\vec{y})r^2$, is independent of r, $r^* = r^*(\vec{x},\vec{\epsilon})$ is the distance between the point $\vec{x}$ and the point at the boundary of the homogeneity in the direction $\vec{\epsilon}$. Considered is the ellipsoidal homogeneity with its centre at the origin 0 and the principal axes coinciding with the orthonormal base vectors $\{\vec{e}_1,\vec{e}_2,\vec{e}_3\}$. Positions $\vec{r}^* = \vec{x}+r^*\vec{\epsilon}$ of points at the surface of the homogeneity obey

$$\vec{r}^*.D.\vec{r}^* = 1 \quad , \quad D = a_1^{-2}\vec{e}_1\vec{e}_1 + a_2^{-2}\vec{e}_2\vec{e}_2 + a_3^{-2}\vec{e}_3\vec{e}_3 \qquad (A6.3.2.2)$$

with a_1, a_2, a_3 the half principal axes lengths. Combination of the explicit expression for $r^*(\vec{x},\vec{\epsilon})$, following from (A6.3.2.2), with equations (A6.3.2.1) and appendix 6.3.1 yields an expression for e_C

$$e_C = {}^4S\!:\!e_T \quad , \quad {}^4S = \frac{1}{16\pi(1-\nu)}\int_{4\pi} \frac{(D.\vec{\epsilon})(^3G^*) + \left((D.\vec{\epsilon})(^3G^*)\right)^{lc}}{\vec{\epsilon}.D.\vec{\epsilon}}\,d\omega \qquad (A6.3.2.3)$$

Because ^4S does not depend on $\vec{x}$, the strain e_C is uniform in the homogeneity. The components S_{pqrs} with respect to $\{\vec{e}_1,\vec{e}_2,\vec{e}_3\}$ are unequal zero for $p=q \wedge r=s$, $p=s \wedge q=r$, $p=r \wedge q=s$. It can be written

$$S_{pqqp} = S_{pqpq} = \tfrac{1}{2}Q(a_p^2 + a_q^2)J_{pq} + \tfrac{1}{2}R(J_p + J_q) \quad ,$$

$$S_{pppp} = Qa_p^2 J_{pp} + RJ_p \quad , \quad S_{ppqq} = Qa_q^2 J_{pq} - RJ_p \tag{A6.3.2.4}$$

with the coefficients Q and R only depending on Poisson's ratio ν according to

$$Q = \frac{3}{8\pi(1-\nu)} \quad , \quad R = \frac{1-2\nu}{8\pi(1-\nu)} \tag{A6.3.2.5}$$

and the integrals J_p, J_{pp} and J_{pq}

$$J_p = \int_{4\pi} \frac{\epsilon_p^2}{a_p^2 g} d\omega \quad , \quad J_{pp} = \int_{4\pi} \frac{\epsilon_p^4}{a_p^4 g} d\omega \quad , \quad J_{pq} = \int_{4\pi} \frac{\epsilon_p^2 \epsilon_q^2}{a_p^2 a_q^2 g} d\omega \tag{A6.3.2.6}$$

with $\epsilon_p = \vec{\epsilon}.\vec{e}_p$, $g = \vec{\epsilon}.D.\vec{\epsilon}$ and $p,q \in \{1,2,3\}$, $p \neq q$. If the quantities a_p are mutual different, the above integrals are elliptic of the first and second kind. More explicit results can be obtained if mutual equality occurs. For details it is referred to Eshelby (1957).

Appendix 6.3.3 The constrained strain components for ellipsoidal inhomogeneities

The relationships between the strain components $(e_C)_{11}$, $(e_C)_{22}$, $(e_C)_{33}$ and $(e_A)_{11}$, $(e_A)_{22}$, $(e_A)_{33}$ are

$$
\begin{bmatrix} (e_C)_{11} \\ (e_C)_{22} \\ (e_C)_{33} \end{bmatrix} = \underline{S}^* \underline{T}^{-1} \underline{U} \begin{bmatrix} (e_A)_{11} \\ (e_A)_{22} \\ (e_A)_{33} \end{bmatrix} + \underline{S}^* \underline{T}^{-1} \kappa_i \begin{bmatrix} \eta \\ \eta \\ \eta \end{bmatrix}
\tag{A6.3.3.1}
$$

where the matrix $\underline{S}^*$ contains the components of the fourth order tensor ^4S of appendix 6.3.2, such that $S^*_{pq} = S_{ppqq}$, $p,q \in \{1,2,3\}$. For the components of the matrices $\underline{T}$ and $\underline{U}$ holds

$$
T_{pq} = 2(G_i - G_m)S_{ppqq} + 2(\frac{G_i \nu_i}{1-2\nu_i} - \frac{G_m \nu_m}{1-2\nu_m}) \sum_{r=1}^{3} S_{rrqq} + 2G_m \frac{\nu_m}{1-2\nu_m} + 2G_m \delta_{pq} \quad ,
$$

$$
U_{pq} = 2(\frac{G_m \nu_m}{1-2\nu_m} - \frac{G_i \nu_i}{1-2\nu_i}) + 2(G_m - G_i)\delta_{pq}
\tag{A6.3.3.2}
$$

with $p,q \in \{1,2,3\}$ and δ_{pq} the Kronecker delta so $\delta_{pq} = 1$ for $p = q$ and $\delta_{pq} = 0$ for $p \neq q$. The scalars κ_i and η are defined by

$$
\kappa_i = \frac{2G_i(1+\nu_i)}{3(1-2\nu_i)} \quad , \quad \eta = \frac{1}{3}\alpha_m \Delta T_m - \frac{1}{3}\alpha_i \Delta T_i
\tag{A6.3.3.3}
$$

If $\eta = 0$ no thermal effect results in the strains $(e_C)_{11}$, $(e_C)_{22}$ and $(e_C)_{33}$. This is always the case for the shear strains

$$
(e_C)_{pq} = \frac{2S_{pqqp}(G_m - G_i)}{2S_{pqqp}(G_i - G_m) + G_m}(e_A)_{pq}
\tag{A6.3.3.4}
$$

with $p,q \in \{1,2,3\}$, $p \neq q$.

Appendix 6.3.4 The transversal isotropic thin layer

The results (6.3.14) are extended to a transversal isotropic layer. With the layer isotropic in its plane, Hooke's law reads

$$
\underset{\sim}{\sigma}_o =
\begin{bmatrix}
C_1 & C_2 & C_3 & & & \\
C_2 & C_1 & C_3 & & \underline{0} & \\
C_3 & C_3 & C_4 & & & \\
& & & C_1 - C_2 & 0 & 0 \\
& \underline{0} & & 0 & C_5 & 0 \\
& & & 0 & 0 & C_5
\end{bmatrix}
\left(\underset{\sim}{e}_{tot,o} -
\begin{bmatrix}
\alpha_T \\
\alpha_T \\
\alpha_L \\
0 \\
0 \\
0
\end{bmatrix}
\Delta T_o \right)
\tag{A6.3.4.1}
$$

The quantities α_T and α_L are the thermal coefficients of linear expansion in the so-called transversal and longitudinal direction respectively. The elastic constants C_p, $p \in \{1,...,5\}$, can be expressed in the longitudinal Poisson's ratio ν_L, elongation modulus M_L, shear modulus G_L and the transversal plane strain compression modulus K_T and shear modulus G_T

$$
C_1 = K_T + G_T \quad , \quad C_2 = K_T - G_T \quad , \quad C_3 = 2K_T\nu_L \quad , \quad C_4 = M_L + 4K_T\nu_L^2 \quad ,
$$

$$
C_5 = 2G_L
\tag{A6.3.4.2}
$$

The expression for $\underset{\sim}{e}_{tot,o}$ becomes equivalent to (6.3.12), however with modified constants φ, χ, ψ, η^o and η^m

$$
\varphi = \frac{1}{C_4}\left(2G_m \frac{\nu_m}{1-2\nu_m} - C_3\right) \quad , \quad \chi = \frac{2G_m}{C_4}\frac{1-\nu_m}{1-2\nu_m} \quad , \quad \psi = \frac{2G_m}{C_5} \quad ,
$$

$$
\eta^o = \frac{1}{C_4}\left(2C_3\alpha_T + C_4\alpha_L\right) \quad , \quad \eta^m = \frac{1}{3C_4}\left(2C_3 + C_4\right)\alpha_m
\tag{A6.3.4.3}
$$

References

Abramowitz, M. & I.A.Stegun (eds.) 1965. *Handbook of mathematical functions.* New York: Dover Publications.

Andreae, G. 1974. On the influence of hydrostatic pressure on strain gauges. (in German). *Materialpruefung 16*:98–102.

Arnold, C.R., P.J.Rolls & J.C.J. Stewart 1971. *Applied photography.* D.A.Spencer (ed.) London: The Focal Press.

Baaijens, F.P.T., W.A.M.Brekelmans, F.E.Veldpaus & F.J.M.Starmans 1986. A constitutive equation for frictional phenomena including history dependency. In K.Mattiasson, A.Samuelsson, R.D.Wood & O.C.Zienkiewicz (eds.), *Proceedings of the 2nd International Conference on Numerical Methods in Industrial Forming Processes*, p. 91–96. Rotterdam: Balkema.

Baaijens, F.P.T. 1987. *On a numerical method to solve contact problems.* Ph.D.Thesis. Eindhoven University of Technology, The Netherlands.

Barsis, E., E.Williams & C.Skoog 1970. Piezoresistivity coefficients in manganin. *Journal of Applied Physics 41*:5155–5162.

Bathe, K.–J. 1982. *Finite element procedures in engineering analysis.* Englewood Cliffs: Prentice-Hall.

Bedda, M., S.Messaadi, C.R.Pichard & A.J.Tosser 1986. Analytical expression for the total conductivity of unannealed and annealed metal films. *Journal of Materials Science 21*:2643–2647.

Belser, R.B. 1957. Electrical resistances of thin metal films before and after artificial aging by heating. *Journal of Applied Physics 28*:109–116.

Bethe, K. 1988. Thin film strain gauges for the construction of sensors. (in German). *Technica 7*:44–47.

Bolk, W.T. 1982. The opposite current circuit – A proposal for strain gauge measurements. (in German). *Messen + Pruefen / Automatik 18*:296–298,308.

Brandes, E.A. (ed.) 1983. *Smithells metals reference book.* London: Butterworths.

Brouha, M., J.E.de Jong & K.J.A.van der Weide 1979. Experimental verification of finite element analysis on axisymmetric deformation processes. *Proceedings of the North American Metalworking Research Conference VII*, p. 57–64. Dearborn: Society of Manufacturing Engineers.

Chambless, D.A., W.F.Swinson, J.C.Suhling & J.L.Turner 1986. New hybrid photoelastic finite element technique for stress analysis. *Proceedings of the 1986 SEM Spring Conference on Experimental Mechanics,* p. 991–998. Bethel: Society for Experimental Mechanics.

Chan, K.S., S.R.Bodner & U.S.Lindholm 1988. Phenomenological modeling of hardening and thermal recovery in metals. *Journal of Engineering Materials and Technology 110*:1–8.

Cheng, J.–H. & N.Kikuchi 1985. An incremental constitutive relation of unilateral contact friction for large deformation analysis. *Journal of Applied Mechanics 52*:639–648.

Contreras, H. 1980. The stochastic finite–element–method. *Computers & Structures 12*:341–348.

DIANA, 1988. *DIANA finite element analysis user's manual.* Rijswijk: Institute TNO for Building Materials and Building Structures.

Dössel, O. 1984. Longitudinal and transverse gauge factors of polycristalline strain gauges. *Sensors and Actuators 6*:169–179.

Drewniak, J. & Z.Pawicki 1985. Boundary element method for fluctuations of boundary conditions. *Mechanics Research Communications 12*:113–118.

Eshelby, J.D. 1957. The determination of the elastic field of an ellipsoidal inclusion, and related problems. *Proceedings of the Royal Society Series A 241*:376–396.

Francis, H.A. 1977. Application of spherical indentation mechanics to reversible and irreversible contact between rough surfaces. *Wear 45*:221–269.

Fredriksson, B. 1976. Finite element solution of surface nonlinearities in structural mechanics with special emphasis to contact and fracture mechanics problems. *Computers & Structures 6*:281–290.

Frost, H.J. & M.F.Ashby 1982. *Deformation mechanism maps: The plasticity and creep of metals and ceramics.* Oxford: Pergamon Press.

Galenkamp, H. & H.van Wijngaarden 1985. Determining the Von Mises stress from the neck of the tensile–test specimen. *Philips Technical Review 42*:11–19.

Gill, P.E., W.Murray & M.H.Wright 1981. *Practical optimization.* London: Academic Press.

Gräbener, Th. 1981. Tribological problems in cold bulk forming. (in German). *Draht 32*:559–564.

Griffiths, D.F. 1979. Finite elements for incompressible flow. *Mathematical Methods in the Applied Sciences 1*:16–31.

Handa, K. & K.Andersson 1981. Application of finite element methods in the statistical analysis of structures. In T.Moan & M.Shinozuka (eds.), *Proceedings of the 3rd International Conference on Structural Safety and Reliability*, p. 409–417. Amsterdam: Elsevier.

Herbertz, R. & H.Wiegels 1981. A method for frictionless upsetting of cylindrical workpieces for registering flow curves. (in German). *Stahl und Eisen 101*:89–92.

Hoffmann, K. 1987. *An introduction to the measuring technique with strain gauges.* (in German). Darmstadt: Hottinger Baldwin Messtechnik GmbH.

Holland, L. 1961. *Vacuum deposition of thin films.* London: Chapman & Hall.

Hunter, S.C. 1976. *Mechanics of continuous media.* New York: John Wiley & Sons.

Huyghe, J.M.R.J. 1986. *Non–linear finite element models of the beating left ventricle and the intramyocardial coronary circulation.* Ph.D.Thesis. Eindhoven University of Technology, The Netherlands.

IMSL, 1987. *MATH/LIBRARY user's manual, version 1.0.* Houston: IMSL Inc.

Jansen, H.J. 1987. *Estimators: properties, construction and overview.* (in Dutch). M.Sc.Thesis, part 1. Report WFW87076, Department of Mechanical Engineering, Eindhoven University of Technology, The Netherlands.

Jansen, H.J. 1987. *A statistical method to reconstruct a deformation process in case of history dependent material behaviour.* (in Dutch). M.Sc.Thesis, part 2. Report WFW87077, Department of Mechanical Engineering, Eindhoven University of Technology, The Netherlands.

Kihara, J., G.Shen, T.Yamauchi, H.Mimura, H.Makino & B.Liu 1983. Some applications of BEM for evaluation of the deformation of tool and calculation of the residual stress distribution in product in the plastic working processes. In C.A.Brebbia, T.Futagami & M.Tanaka (eds.), *Boundary Elements. Proceedings of the Fifth International Conference*, p. 393–405. Berlin: Springer–Verlag.

Klarbring, A. 1985. *Contact problems in linear elasticity. Friction laws and mathematical programming applications.* Ph.D.Thesis. Linköping University, Sweden.

Klimczak, T., J.H.Dautzenberg & J.A.G.Kals 1988. On the roughening of a free surface during sheet metal forming. *Annals of the CIRP 37*:267–270.

Kobayashi, A.S. (ed.) 1987. *Handbook on experimental mechanics.* Englewood Cliffs: Prentice-Hall.

Kubo, S. 1988. Inverse problems related to the mechanics and fracture of solids and structures. *JSME International Journal, series I 31*:157–166.

Kunisch, K. & L.White 1986. Parameter estimation for elliptic equations in multidimensional domains with point and flux observations. *Nonlinear Analysis, Theory, Methods & Applications 10*:121–146.

Landheer, D., A.J.G.Dackus & J.A.Klostermann 1980. Fundamental aspects and technological implications of the solubility concept for the prediction of running properties. *Wear 62*:255-286.

Lange, K. 1985. *Handbook of metal forming.* New York: McGraw–Hill.

Lee, E.H. 1969. Elastic–plastic deformations at finite strain. *Journal of Applied Mechanics 36*:1–6.

Liu, W.K., A.Mani & T.Belytschko 1987. Finite element methods in probabilistic mechanics. *Probabilistic Engineering Mechanics 2*:201–213.

Love, A.E.H. 1944. *A treatise on the mathematical theory of elasticity.* New York: Dover Publications.

MARC, 1988. *MARC general purpose finite element program, revision K3, user's manuals A–E.* Palo Alto: MARC Analysis Research Corporation.

Meijer, H.A. 1988. Personal communication.

Meyer–Nolkemper, H. 1982. Flow curves of metallous working materials. (in German). *HFF–Report nr.4.* Hannover: Hannoverisches Forschungsinstitut für Fertigungsfragen.

Michalowski, R. & Z.Mróz 1978. Associated and non–associated sliding rules in contact friction problems. *Archives of Mechanics 30*:259–276.

Müller, I. 1985. *Thermodynamics.* London: Pitman Publishing.

Mura, T., B.Cox & Z.Gao 1986. Computer–aided nondestructive measurements of plastic strains from surface displacements. In G.Yagawa & S.N.Atluri (eds.), *Computational Mechanics 1986. Proceedings of the International Conference on Computational Mechanics,* p. II.43–II.48. Tokyo: Springer–Verlag.

NAG, 1987. *NAG Fortran library manual, MARK 12.* Oxford: Numerical Algorithms Group Ltd.

Norton, J.P. 1986. *An introduction to identification.* London: Academic Press.

Oda, J. & T.Shinada 1987. An inverse analysis technique to obtain contact stress distributions. (in Japanese). *Transactions of the Japanese Society of Mechanical Engineers 53*:1614–1621.

Oden, J.T. & J.A.C.Martins 1985. Models and computational methods for dynamic friction phenomena. *Computer Methods in Applied Mechanics and Engineering 52*:527–634.

Peeken, H. & A.Köhler 1979. Measuring technique with evaporated transducers on machine parts. (in German). *Antriebstechnik 18*:99–101.

Peters, G.W.M. 1987. *Tools for the measurement of stress and strain fields in soft tissue.* Ph.D.Thesis. University of Limburg, The Netherlands.

Petersson, H. 1977. *Constitutive relations for contacting solids.* Division of Structural Design, publication 77–10. Chalmers University of Technology, Göteborg, Sweden.

Rabinowicz, E. 1965. *Friction and wear of materials.* New York: John Wiley & Sons.

Rasmussen, S.N., W.Nester & K.Pöhlandt 1984. A further development of the Rastegaev upsetting test for measuring flow curves. (in German). *Zeitschrift fuer Industrielle Fertigung 74*:667–670.

Reiss, W. & K.Pöhlandt 1985. Behaviour and manipulation of wrought materials under homogeneous deformation. (in German). *Materialpruefung 27*:100–104.

Ruina, A.L. 1985. Constitutive relations for frictional slip. Chapter 9 in Z.Bažant (ed.), *Mechanics of Geomaterials.* New York: John Wiley & Sons.

Schey, J.A. 1983. *Tribology in metalworking: friction, lubrication and wear.* Metals Park: American Society for Metals.

Schoofs, A.J.G. 1987. *Experimental design and structural optimization.* Ph.D.Thesis. Eindhoven University of Technology, The Netherlands.

Schouten, M.J.W. 1973. Elastohydrodynamic lubrication. (in German). *Forschungskuratorium Maschinenbau e.V. research report 24.* Frankfurt/M: Maschinenbau Verlag.

Schweppe, F.C. 1973. *Uncertain dynamic systems.* Englewood Cliffs: Prentice-Hall.

Sidoroff, F. 1973. The geometrical concept of intermediate configuration and elastic-plastic finite strain. *Archives of Mechanics 25*:299–309.

Sokal, R.R. & F.J.Rohlf 1969. *Biometry.* San Francisco: W.H.Freeman and Company.

Stacey, A., H.J.MacGillivary, G.A.Webster, P.J.Webster & K.R.A.Ziebeck 1985. Measurement of residual stresses by neutron diffraction. *Journal of Strain Analysis 20*:93–100.

Stein, P.K. 1962. *Measurement engineering.* s.l.: Strain Gage Readings.

Tangena, A.G. 1987. *Tribology of thin film systems.* Ph.D.Thesis. Eindhoven University of Technology, The Netherlands.

Theocaris, P.S. & C.I.Razem 1979. Intensity, slope and curvature discontinuities in loading distributions at the contact of two plane bodies. *International Journal of Mechanical Sciences 21*:339–353.

Timothy, S.P., J.M.Pearson & I.M.Hutchings 1987. The contact pressure distribution during plastic compression of lead spheres. *International Journal of Mechanical Sciences 29*:713–719.

Tuncer, C. & T.A.Dean 1987. A new pin design for pressure measurement in metal forming processes. *International Journal of Machine Tools and Manufacture 27*:325–331.

Turner, J.L., S.R.McNeill, Y.J.Chao, M.A.Sutton & W.H.Peters 1988. Hybrid methods in experimental mechanics. In S.N.Atluri & G.Yagawa (eds.), *Computational Mechanics 1988. Proceedings of the International Conference on Computational Engineering Science*, p. 48.ix.1-48.ix.4. New York: Springer-Verlag.

Van de Vosse, F.N. 1987. *Numerical analysis of carotid artery flow*. Ph.D.Thesis. Eindhoven University of Technology, The Netherlands.

Van Rooyen, G.T. & W.A.Backofen 1960. A study of interface friction in plastic compression. *International Journal of Mechanical Sciences 1*:1:27.

Van Wijngaarden, H. 1988. *Constitutive equations for metals with an application to the extrusion of lead*. Ph.D.Thesis. Eindhoven University of Technology, The Netherlands.

Vater, M. & G.Nebe 1965. On stress and strain distribution in upsetting. (in German). *VDI−Z Progress report series 2*(5). Düsseldorf: Verein Deutscher Ingenieure.

Whitaker, S. 1969. Advances in theory of fluid motion in porous media. *Industrial and Engineering Chemistry 61*(12):14–28.

Zienkiewicz, O.C. 1977. *The finite element method*. London: McGraw-Hill.

Zienkiewicz, O.C., J.P.Vilotte, S.Toyoshima & S.Nakazawa 1985. Iterative method for constrained and mixed approximation. An inexpensive improvement of FEM performance. *Computational Methods in Applied Mechanical Engineering 51*:3–29.

Zienkiewicz, O.C., S.Qu, R.L.Taylor & S.Nakazawa 1986. The patch test for mixed formulations. *International Journal for Numerical Methods in Engineering 23*:1873–1883.

Zmitrowicz, A. 1987. A thermodynamical model of contact friction and wear: I Governing equations; II Constitutive equations for materials and linearized theories; III Constitutive equations for friction, wear and frictional heat. *Wear 114*:135–168; *114*:169–197; *114*:199–221.

Samenvatting

In veel omvormprocessen spelen contactfenomenen een belangrijke rol met betrekking tot de benodigde proceskrachten, optredende mechanische spanningen in gereedschappen en product, resulterende productgeometrie, restspanningen en oppervlaktegesteldheid. Om te komen tot een juiste dimensionering van gereedschappen en smeermiddelkeuze in de ontwerpfase, is het van belang om te beschikken over modellen welke de mechanische interactie voldoende betrouwbaar beschrijven. Bekende fenomenologische modellen voor het mechanisch contactgedrag zijn het Coulomb wrijvingsmodel, het Von Mises wrijvingsmodel en het in het laatste decennium ontwikkelde elasto-plastische analogie model. Dit proefschrift beschrijft een aantal gereedschappen om de bruikbaarheid van zulke contactmodellen te onderzoeken, alsmede de contactparameters uit deze modellen te kwantificeren.

Na een inleiding en afbakening van het beoogde onderzoeksgebied in hoofdstuk 1, wordt in hoofdstuk 2 een thermodynamisch continuümsmodel beschreven voor de contactsituatie bij omvormprocessen. De bovengenoemde mechanische contactmodellen blijken onder tamelijk algemene condities binnen dit thermodynamische kader te passen. Experimenteel onderzoek is derhalve noodzakelijk om tot een nadere specificatie van bruikbare contactmodellen te komen.

In hoofdstuk 3 worden afwegingen gegeven met betrekking tot mogelijke experimenten en meetmethoden, resulterend in de keuze om *in situ* het stuikproces op axiaalsymmetrische werkstukken nader te onderzoeken. Hoofdstuk 4 beschrijft een experimenteel-numerieke werkwijze waarmee gemeten randvoorwaarden in rekening gebracht kunnen worden bij analyses van omvormprocessen, alsmede hun invloed bestudeerd kan worden op de uiteindelijke resultaten. Met dit gereedschap is het stuikproces met verschillende sets randvoorwaarden geanalyseerd. Het blijkt dat meting van de normaalspanning, en eventueel de afschuifspanning, in het contactgebied tussen stempel en stuikproduct noodzakelijk is om enigszins betrouwbaar de contactgrootheden te bepalen ter evaluatie van het contactgedrag. Deze betrouwbaarheid kan vergroot worden als ook de verplaatsingen van het vrije buitenoppervlak, mits voldoende nauwkeurig gemeten, verdisconteerd worden.

Hoofdstuk 5 presenteert een contactloze fotogrammetrische meetmethode voor de verplaatsingen van het vrije buitenoppervlak. Experimentele evaluatie leert dat de theoretisch voorspelde nauwkeurigheid niet haalbaar is als gevolg van tekortkomingen in de fysische modelvorming. De invloed van een aantal van deze tekortkomingen kan met een verbeterde proefopzet gereduceerd worden. Bij het

optreden van ernstige oppervlakteverruwing is het echter twijfelachtig of de benodigde nauwkeurigheid bereikt kan worden. De verplaatsingsmeting blijft van belang voor de bepaling van de afmeting van het contactgebied, welke een invoervariabele is bij de kwantificering van de contactspanningen volgens hoofdstuk 6. Dat hoofdstuk geeft een ontwerp en analyse van een meetgereedschap voor de contactspanningen bij het stuikproces. Meting van de elastische rekken in het inwendige van het gereedschap, gecombineerd met een rekenmodel leidt tot een uitspraak over de contactspanningen tussen gereedschap en werkstuk. De rekken worden gemeten met dunne film rekopnemers. Voor deze opnemers is een fysisch model opgesteld. Een eerste aanzet is gegeven ter evaluatie van dit model alsmede van het voorspelde gedrag van het totale gereedschap.

Tot slot geeft hoofdstuk 7 conclusies met betrekking tot het huidige onderzoek, alsmede adviezen voor een voortzetting.

Nawoord

Bij deze wil ik mijn dank doen toekomen aan iedereen die heeft bijgedragen aan het tot stand komen van dit proefschrift. De vakgroep WFW zorgde met haar fijne werksfeer voor een aantal prettige onderzoeksjaren. De grote lijn in het onderzoek werd op stimulerende wijze door Jan Janssen bewaard. Marcel Brekelmans toonde zich, na overname van de begeleidende taak van Frans Veldpaus, een deskundig en inspirerend klankbord. Zijn inbreng op dit werk is van onschatbare waarde. Jaap Jansen droeg als afstudeerder zijn steentje bij aan het onderzoek. Het typewerk is gedeeltelijk verzorgd door Marleen van Boxtel. Hoewel dit uit de inhoud niet zo blijkt, is een behoorlijke tijd geïnvesteerd in experimenten. Met betrekking tot de realisatie van proefopstellingen zijn met name Theo van Duppen, Toon van Gils, Karel Koekkoek en Julius IJzermans te vermelden. Laatstgenoemde, alsmede Roel van de Brink en Leo Wouters waren behulpzaam bij het verzamelen en verwerken van experimentele gegevens. Bijdragen van buiten de vakgroep werden geleverd door Sjaak Cauwenberg, Rob König, Henk Meijer en Thieu Smeets. Door vele anderen, hier niet met naam genoemd, heeft dit werk gestalte gekregen. Mijn erkentelijkheid komt eveneens hen toe. Tot slot gaat mijn dank uit naar de achterban voor haar bewuste en onbewuste steun. Hierbij wil ik mijn ouders en naaste familie afzonderlijk vermelden. Hun steun en inzet heeft onmiskenbaar bijgedragen aan dit resultaat.

Levensbericht

17–3–1961   Geboren te Schaesberg

1973–1979   Gymnasium B aan de scholengemeenschap Sint Michiel te Geleen

1979–1984   Studie werktuigbouwkunde aan de Technische Universiteit
Eindhoven

1984–1988   Wetenschappelijk assistent aan de Technische Universiteit
Eindhoven, afdeling werktuigbouwkunde

Stellingen
behorende bij het proefschrift
ON FRICTION IN FORMING

1) Continuümsthermodynamica levert in het algemeen slechts nevencondities ten aanzien van constitutieve vergelijkingen. Theoretisch en experimenteel structureel onderzoek is onontbeerlijk voor een volledige modelvorming.
   - Dit proefschrift, hoofdstuk 2.
   - **Rooyackers, H.F.L. 1988.** *A numerical implementation of the Schapery model for nonlinear visco-elasticity.* Proefschrift Technische Universiteit Eindhoven.

2) Bij de cilinderstuikproef kunnen op basis van de verplaatsingen van het onbelaste buitenoppervlak van het proefstuk geen significante uitspraken worden gedaan met betrekking tot het contactgedrag ter plaatse van het belaste buitenoppervlak.
   - Dit proefschrift, hoofdstuk 4.
   - **Herbertz, R. & H.Wiegels 1981.** Ein Verfahren zur Verwirklichung des reibungsfreien Zylinderstauchversuches für die Ermittlung von Fließkurven. *Stahl und Eisen 101:*89 –92.

3) Bij het ontwerp en de analyse van meetmethoden en de verwerking van meetdata binnen de experimentele mechanica vormen hybride experimenteel-numerieke werkwijzen waardevolle gereedschappen.
   - Dit proefschrift, hoofdstukken 4 en 6.
   - Hybrid experimental computational methodology in solid mechanics. Hoofdstuk 48 in S.N.Atluri & G.Yagawa (eds.), *Computational Mechanics '88. Proceedings of the International Conference on Computational Engineering Science.* New York: Springer –Verlag.

4) Resultaten van drukmeting in veel elastohydrodynamisch onderzoek moeten in twijfel worden getrokken, gezien het niet in rekening brengen van de verschillen in de relaties tussen druk en opnemerrekken voor de ijksituatie enerzijds en de meetsituatie anderzijds.
   - Dit proefschrift, hoofdstuk 6.
   - **Köhler, A. 1981.** *Die Entwicklung von aufgedampften Messwertaufnehmer und deren Anwendung zur Druck– und Temperaturmessung in geschmierten Walz– und Gleitkontakten.* Proefschrift RWTH Aachen.

5) Bij de bepaling van axiaalsymmetrische contactspanningsverdelingen met het in dit proefschrift beschreven meetgereedschap is gebruik van opnemers voor of alleen de radiale rek of alleen de tangentiale rek onvoldoende.
   - Dit proefschrift, hoofdstuk 6.

6) Voor betrouwbare temperatuurmetingen met dunne film signaalopnemers is het nodig dat het thermoresistief gedrag domineert over het piëzoresistief gedrag of dat de piëzoresistieve effecten afdoende gecompenseerd worden.
   - Dit proefschrift, hoofdstuk 6.
   - **Kannel, J.W., F.F.Zugaro & T.A.Dow 1978.** A method for measuring surface temperature between rolling/sliding steel cylinders. *Journal of Lubrication Technology 100*:110−114.

7) De waarde van een softwarepakket wordt vergroot door flexibele mogelijkheden om gegevens uit te kunnen wisselen met andere softwarepakketten.
   - **Vroom, R. 1988.** Toepasbaarheid databases in techniek nog niet probleemloos. *Ingenieur & Computer* nr.5, september 1988.

8) Gezien de toenemende noodzaak om van het huidige economische groeimodel af te stappen is het aan te bevelen zogenaamd toekomstonderzoek als verplicht onderdeel in iedere academische studie op te nemen.
   - **Stikker, A. 1988.** *De prijs van een wonder : naar nieuwe modellen voor onze toekomst.* Amsterdam: Bres.
   - **Groot Wassink, J. 1988.** Vorming in het technisch universitair onderwijs. *De Ingenieur 100*(5):26−31.

9) Vervolgonderwijs voor ingenieurs is wezenlijk en dient vooral gestalte te krijgen via part−time onderwijs door de (technische) universiteiten.
   - **De Vree, J.H.P., J.D.Janssen & D.H.van Campen 1988.** A new continuous engineering education programme for computational and experimental mechanics in the Netherlands. *First European Forum for Engineering Education : L'Investissement dans l'Avenir.* Najaar '88, Stuttgart.

10) Een defensie zonder provocatie is een middel om een vreedzame wereldsamenleving te verwezenlijken.
    - **Boeker, E. 1986.** *Europese veiligheid: alternatieven voor de huidige veiligheidspolitiek.* VU Studies Vrede en Veiligheid. Amsterdam: VU Uitgeverij.

11) Voor elke dienstplichtige zou de wijze van vervulling van de militaire dienstplicht een bewuste keuze moeten zijn.
    - *Alles over dienstweigeren.* Den Haag: Pax Christi, 1980.

Frans Starmans, december 1988