TAMPEREEN TEKNILLINEN KORKEAKOULU
Teknillinen mekaniikka

VII SUOMEN MEKANIIKKAPÄIVÄT

Tampereella 25-26. 5. 2000

NIDE 2

2. kokouspäivän esitelmät

Toimittajat
Juhani Koski ja *Simo Virtanen*

Tampere 2000

TAMPEREEN TEKNILLINEN KORKEAKOULU
Teknillinen mekaniikka
PL 589
33101 TAMPERE

Puh. (03) 3652296
fax. (03) 3652107

ISBN 952-15-0414-5

Klingendahl Paino Oy
Tampere, 2000

SISÄLLYSLUETTELO

Hitsiliitokset

Elementtimenetelmä 1

Optimointi 2

Värähtelyt

Väsyminen ja lämmönsiirto

Jäännösjännitykset ja elementtimenetelmä

NIDE 2

Kutsutut esitelmät

Rakenteiden mekaniikka

Palkit

Maamekaniikka

Erityisaiheita

Elementtimenetelmä 2

ON OPTIMUM TOPOLOGY DESIGN IN MECHANICS

Niels Olhoff

Institute of Mechanical Engineering
Aalborg University, DK-9220 Aalborg, Denmark

Hans Eschenauer

Research Center for Multidisciplinary Analyses and Applied Structural Optimization
University of Siegen, Germany

Abstract. *Optimization of the basic topology or layout of a structure has decisive influence on the cost-efficiency of the structure, and this research area has been developing very rapidly over the last decade. It is the objective of this paper to present a picture of this development and the status of the field through a survey of problems of topology optimization of linearly elastic 2D and 3D continuum structures*

1 Introduction

'Topology optimization' is often referred to as 'layout optimization' or 'generalized shape optimization' in the literature (see the review by Rozvany, Bendsøe & Kirch [1] and Bendsøe's monograph [2]), and these labels will be used interchangeably in this survey paper. The importance of this type of optimization lies in the fact that the choice of the appropriate topology of a structure is generally the most decisive factor for the cost-efficiency of the structure.

Moreover, conventional sizing and shape optimization methods (see, e.g., unified surveys by Olhoff & Taylor [3], Olhoff [4] and Bendsøe [5]) cannot change the structural topology during the solution process, so a solution obtained by one of these methods will have the same topology as that of the initial design. Topology or layout optimization methods are therefore most valuable as preprocessing tools for sizing and shape optimization, see Olhoff, Bendsøe & Rasmussen [6] and Bremicker *et al.* [7].

Depending on the type of structure to be considered, two types of topology optimization exist. For inherently discrete structures, the optimum topology or layout design problem consists in determining the optimum number, positions, and mutual connectivity of the structural members. This area of research has been active for almost a century, and has been largely developed by Prager and Rozvany. For an up-to-date account of the area of layout optimization of discrete structures, the reader is referred to, e.g., the comprehensive review paper by Rozvany, Bendsøe & Kirsch [1], the monograph by Bendsøe [2], and recent proceedings edited by Bendsøe & Mota Soares [8], Pedersen [9], Olhoff & Rozvany [10], Gutkowski & Mroz [11], and Bloebaum [12].

Like the recent surveys by Bendsøe [13] and Hassani & Hinton [14], the present survey paper is restricted to topology optimization of continuum structures. This research area has been extremely active since the publication of the landmark papers by Bendsøe & Kikuchi [15] in 1988 and Bendsøe[16] in 1989.

In topology optimization of continuum structures, the shape of external as well as internal boundaries and the number of inner holes are optimized simultaneously with respect to a predefined design objective. It is assumed that the loading is prescribed and that a given amount of structural material is specified within a given 2D or 3D design domain with given boundary conditions. Chapter 2 of this paper briefly outlines the basic concepts and the mathematical-physical fundamentals of the problem; the conceptual process of topology optimization and the basic mathematical formulation of the problem are presented, and different frequently used material models are discussed.

As was originally pointed out by Lurie (see [17] for references), this type of optimization problem is not well-posed if the design space is not closed in an appropriate sense, and a regularization of the formulation of the problem is then needed , see Lurie, Cherkaev & Fedorov [18], Cheng & Olhoff [19], [20], and Kohn & Strang [21], [22]. Mathematical indications of the need for regularization are generation of anisotropy in the design and the impossibility of satisfying second order necessary conditions for optimality in certain subregions of the structural domain, see Olhoff, Lurie, Cherkaev & Fedorov [19] and Lurie,

Cherkaev & Fedorov [18]. Numerically the need manifests itself by lack of convergence or by dependence of the topology on the size of the applied finite element mesh, see Cheng & Olhoff [20], [23], Olhoff, Lurie, Cherkaev & Fedorov [19], and Cheng [24] .

There are two paths out of this dilemma: one can either extend the design space such as to include solutions with microstructure in the problem formulation, or one can restrict the space of admissible solutions in the formulation, cf. Niordson[25], Bendsøe[26], and Olhoff & Taylor[3].

The former path encounters an approach termed relaxation in which materials with periodic, perforated microstructure of continuously varying volume density and orientation are included as admissible designs (see [19], [23], [21], [22], [27] and [28]), and where their effective mechanical properties are determined via some sort of homogenization technique, e.g., a method of "smear-out" (see, e.g., [19], [23] and [29]), homogenization (see, e.g., [15], [16], [30]), or quasiconvexification ([31], [32], [33], [34]). For the different microstructure models discussed in Chapter 2 of this paper, homogenization methods for establishing their effective mechanical (stiffness) properties are presented in Chapter 3. Moreover, advantages and disadvantages of results obtained via the relaxation approach are subsequently discussed and illustrated by examples of optimum topology design of 2D and 3D structures in Chapter 3.

The latter path out of the abovementioned dilemma implies the introduction of an appropriate restriction in the problem formulation that renders the problem well-posed, see Haber, Bendsøe & Jog [35], [36], [37], [38], Fernandes, Rodrigues & Guedes[39], Petersson & Sigmund [40], Petersson [41]. The approaches of this kind admit application of simple material models like the SIMP model presented in Section 2.3.4, and they generally provide a means to control the complexity of the topology design. Approaches of this type are discussed in Chapter 4, and include the 'Perimeter method' (Haber, Bendsøe & Jog [35], [36], [37]) by which a bound constraint on the perimeter or surface area of the solid domain (of 2D or 3D designs, respectively) restricts solutions to be entirely composed of purely solid and void domains. The same can be achieved in a simpler way by use of filtering techniques known from image processing (Sigmund [42, 43]), and this approach is also discussed in Chapter 4.

The topology optimization methods cited above all use a fixed finite element mesh, and as they are all based on application of microstructured material models (cf. Chapter 2), they may be referred to as "micro-approaches". These have been used in by far the largest number of studies in the field, and are consequently devoted primary attention in this paper. However, there exists a very different, so-called "macro-approach" to topology design which will be discussed in Chapter 5 in terms of the so-called 'Bubble Method' (Eschenauer *et al.* [44], [45], [46], [47] and Schumacher [48]). In this method, a conventional shape design technique is augmented to one of simultaneous topology design by considering additional design variables that govern the number and positions of holes (called bubbles), and where in each step of an iterative hierarchical procedure the introduction and optimum positioning of a new hole is followed by a conventional fixed-topology shape optimization of the structure.

In Chapter 6, we present an overview of different kinds of problems, design objectives and constraints that can be currently handled in the area of topological optimization of structures. While the topology optimization problems dealt with in the initial papers by Bendsøe and Kikuchi [15] and Bendsøe [16] concerned stiffness maximization for a single load case, subsequent extensions include, e.g., handling of multiple load cases, Diaz & Bendsøe [49]; bi-material structures, Olhoff, Thomsen & Rasmussen [50]; plate and shell bending problems, Suzuki & Kikuchi [51], Soto & Diaz [52], [53], and Diaz, Lipton & Soto [54]; eigenfrequency optimization, Diaz & Kikuchi [49], Soto & Diaz [55], and Krog & Olhoff [56]; buckling eigenvalue optimization problems, Folgado, Rodriques & Guedes [57]; and stress minimization problems Cheng *et al.* [58], [59], and Duysinx *et al.* [60], [61].

Finally, in Chapter 6 we demonstrate by way of examples that in very recent years, new avenues have been opened for the application of topology optimization in areas of design of materials for prescribed mechanical properties (Sigmund [62] , [42]), and design of compliant mechanisms (Sigmund [63]), including design of MicroElectroMechanical Systems (MEMS), Sigmund [64], [65].

2 Basic Concepts and Mathematical-Physical Fundamentals

2.1 Conceptual Process of Topology Optimization

A conceptual presentation of a topology optimization process for a continuum structure is illustrated in Figure 1, where an admissible design domain (the grey area), boundary conditions and loading as shown in Figure 1a are assumed to be given. Let it be our design objective to find the structural topology that minimizes the compliance (maximizes the integral stiffness) subject to a prescribed amount of structural material. We assume that in solid form the amount of material is less than the amount that would be needed to cover the entire admissible domain for the structure, and hence for the initial design shown in Figure 1a, we have chosen to distribute the material evenly in some porous, microstructured form over the admissible design domain.

With exception of some adaptive techniques in topology optimization (Maute *et al.* [66], [67], [68], [69]) it is customary that approaches to topology optimization use the same finite element mesh to approximate the geometry and the mechanical response fields. Typically, the mesh is a uniform, rectangular partition of space, and the design variables are assumed to be constant within each finite element. Thus, a fixed finite element mesh is embedded in the entire admissible design domain in Fig. 1.

For the mechanical response analysis we apply finite elements with constitutive properties that reflect relationships between stiffness components and material density based on physical modelling of the porous microstructures whose orientation and density are described by continuous variables over the admissible domain. Thus, the density of material and the effective material properties related to the density is controlled via geometric variables which govern the material with microstructure that is constructed in order to relate correctly material density with effective material property.

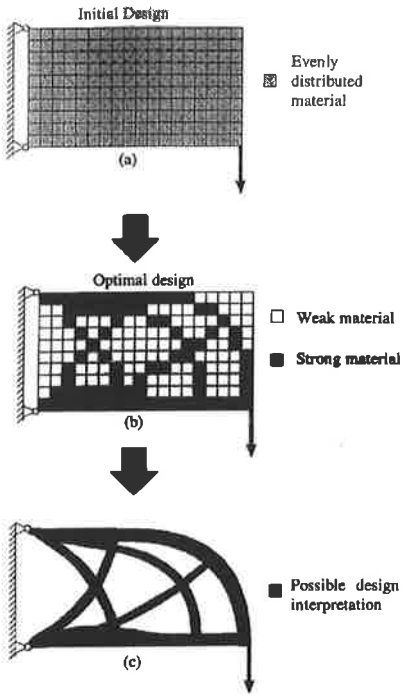


Figure 1: Conceptual representation of a possible optimization process from the initial design to the optimum design

The overall purpose of the problem is to determine the optimum structural topology within the admissible design domain that fulfills all the design specifications. To this end, the density of material within each finite element is used as a design variable defined between limits 1 (solid material, shown in black in Figure 1) and 0 (void or very weak material indicated by white). The optimization consists in determining whether each point (i.e., finite element) in the structure should contain material or not. In this process, the design variables tend to attain one of their limiting values, thereby forming a design with aggregations of points (finite elements) with solid material and void, respectively, see Figure 1b. The result is a rough description of outer as well as inner boundaries of the structure that represents the overall optimum topology of the structure.

Depending on the amount of material available, the generated topology will basically define either the rough shape of a continuum structure, normally with macroscopic interior holes (which usual shape optimization procedures cannot create), or the skeleton of a truss- or beam- like structure with slender members. The important point to be made here is that the optimum topology can be used as a basis (Figure 1c) for procedures of refined shape optimization by means of boundary variations techniques (shape optimizations)

when the topology represents a continuum structure, or as a basis for optimization of member sizes and positions of joints by means of sizing optimization techniques when the optimum topology represents a truss- or beam-like structure, see Olhoff, Bendsøe & Rasmussen [6] and Bremicker *et. al.* [7].

2.2 Basic Formulation of a Topology Optimization Problem

To discuss the basic mathematical formulation for topology optimization of a continuum structure, we now consider the classical problem of topology design for maximum stiffness of linearly elastic structures under a single loading condition. This problem is equivalent to design for minimum compliance defined as the work done by the given loading against the displacements at equilibrium, which, in turn, is equivalent to minimizing the total elastic energy at the equilibrium state of the structure. This can be verified by considering the work W done by the applied forces given by

$$W(\mathbf{u}(x)) \;=\; \int_V (\mathbf{u}(x))^T\, \mathbf{F}_V(x) dV \;+\; \int_S (\mathbf{u}(x))^T\, \mathbf{F}_S\,(x) dS \tag{1}$$

where x denotes a structural point, $\mathbf{u}(x)$ is the displacement field at equilibrium, and $\mathbf{F}_V$ and $\mathbf{F}_S$ are the body forces and surface tractions, respectively, where V and S represent the structural domain and surface. In order to use this relation the structure must be in equilibrium which requires that the equilibrium problem is solved in advance by finite element analysis. In order to express the problem in finite element terms the well known equation for the total potential energy Π is used as this is the basis for static analysis. The total potential energy defined as the strain energy U minus the work W done by the applied forces can be expressed as

$$\begin{aligned}\Pi \;=\; U - W \;=\; & \int_V (\varepsilon(x))^T\, \mathbf{E}(x)\varepsilon(x) dV \;-\; \int_V (\mathbf{u}(x))^T\, \mathbf{F}_V(x)\; \mathbf{u}(x) dV \;- \\ & \int_S (\mathbf{u}(x))^T\, \mathbf{F}_S(x)\; \mathbf{u}(x) dS\end{aligned} \tag{2}$$

where $\varepsilon(x)$ is the strain field and $\mathbf{E}(x)$ is the constitutive matrix of the material. By using standard finite element procedures of discretizing the structure into N finite elements and interpolating displacements from the nodal degrees of freedom, Eq. 2 can be transformed into the well known finite element form

$$\Pi \;=\; U \;-\; W \;=\; \frac{1}{2}\mathbf{D}^T\mathbf{K}\mathbf{D} \;-\; \mathbf{R}^T\mathbf{D} \tag{3}$$

where $\mathbf{D}, \mathbf{K}$ and $\mathbf{R}$ are the global displacement vector, stiffness matrix, and load vector. Equilibrium is obtained by using the principle of minimum total potential energy which requires stationarity of Π with respect to the displacements,

$$\frac{\partial\, \Pi}{\partial\, \mathbf{D}} \;=\; \mathbf{0} \tag{4}$$

which yields the well known equilibrium equation

$$\mathbf{KD} = \mathbf{R} \tag{5}$$

Substituting the developed finite element equations into Eq. 1 the problem of minimizing the work done by the external forces at equilibrium can be expressed as

$$\begin{array}{ll} \min & W(\mathbf{D}) \\ \chi \in B & \\ \text{subject to} & \mathbf{KD} = \mathbf{R} \end{array} \tag{6}$$

where χ denotes the design variables and B is given by

$$B = \left\{ \chi \mid \int_\Omega \chi dx = V_{\text{fixed}} \right\} \tag{7}$$

which expresses that the optimum design is to be determined for a prescribed volume of material.

The given problem can be easily expressed in terms of the total elastic energy by applying that at equilibrium $W = 2U$, which produces the following problem

$$\begin{array}{ll} \min & U(\mathbf{D}) \\ \chi \in B & \\ \text{subject to} & \mathbf{KD} = \mathbf{R} \end{array} \tag{8}$$

Here U can be given in a number of different ways, i.e.

$$U = \frac{1}{2}\mathbf{D}^T\mathbf{KD} = \sum_1^N \frac{1}{2}\int_{V_e} \varepsilon_e^T \mathbf{E}(x)\varepsilon_e dV_e = \sum_1^N \frac{1}{2}\int_{V_e} \sigma_e^T \mathbf{C}(x)\sigma_e dV_e \tag{9}$$

where $\mathbf{E}$ and $\mathbf{C}$ are finite element elasticity and compliance matrices, ε_e and σ_e element strain and stress vectors, V_e the volume, and N the total number of finite elements.

Considering the complementary energy expression at the right hand side of Eq. 9 the compliance matrix, and thus the material distribution problem, is controlled by a design variable that can be expressed by an indicator function defined as

$$\chi(x) = \begin{cases} 1 & \text{if } x \in \Omega_s \\ 0 & \text{if } x \in \Omega \setminus \Omega_s \end{cases} \tag{10}$$

where $\chi = 1$ denotes solid elastic material and $\chi = 0$ denotes void, and where Ω denotes the entire design domain and Ω_s the domain occupied by material.

This implies that the topology optimization problem is formulated as a distributed, discrete valued design problem, a so-called 0-1 problem. However, it is by now well-known that this distributed problem suffers from lack of existence of solutions, i.e., it is ill-posed, see e.g., Kohn & Strang [21] and Allaire & Kohn [70]. Numerically, the deficiency manifests itself by lack of convergence or by dependence of the topology on the size of the

applied finite element mesh, see Cheng & Olhoff [20], [23], Olhoff, Lurie, Cherkaev & Fedorov [19] and Cheng [24].

In order to obtain a well-posed problem, a regularization of the problem formulation is required. This can be done by introduction of microstructures with a continuous density of the base material as a design variable, see Section 3, and/or by including appropriate restrictions against formation of microstructure in the problem formulation as discussed in Section 4. In the latter case even the $0-1$ problem becomes well-posed, and has recently been solved by Beckers [71], [72], and Beckers & Fleury [73] by usage of dual methods.

2.3 Material Models

In this section we present some material models that allow the density of material to cover the complete range of values from 0 (void) over intermediate values (composite) to 1 (solid) and that provide regularization (well-posedness) of typical topology optimization problems when introduced in the mathematical formulation.

We firstly consider three material models with periodic, perforated microstructure, namely the "hole-in-cell" microstructure (Sub-section 2.3.1) and layered microstructures (Sub-section 2.3.2) for 2D topology optimization problems, see Figure 2, and then layered microstructures for 3D problems (Sub-section 2.3.3). These microstructural models provide regularization of the topology optimization problem via relaxation (extension) of the design space, and their periodicity implies that the effective mechanical properties of the microstructures can be determined via homogenization, "smear-out" or quasiconvexification techniques as discussed in Section 3.

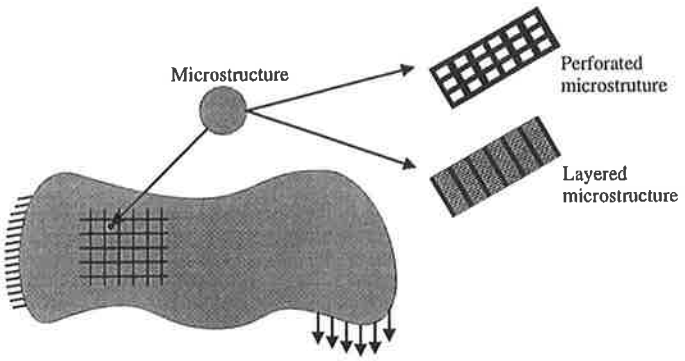


Figure 2: Perforated "hole-in-cell" and layered microstructures for solution of 2D topology optimization problems.

Figure 3 illustrates the parameterization of the perforated "hole-in-cell" and layered 2D microstructures which were first applied in Bendsøe & Kikuchi[15] and Bendsøe[16], respectively. The microstructures depicted in Figure 3 both implies orthotropic material behaviour.

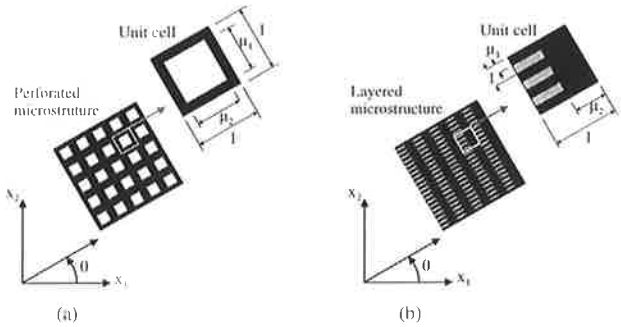


Figure 3: Microstructures for 2D continuum topology optimization problems. (a) Perforated microstructure with rectangular holes in square unit cells. (b) Layered microstructure constructed from two different isotropic materials.

In Sub-section 2.3.4 we finally present the so-called SIMP model. This model also covers the complete range of density values from 0 to 1, but does not provide regularization of the classical formulation of topology optimization problems.

However, it has been proven mathematically by Ambrosio & Butazzo [38] and Petersson [41] and will be discussed in Section 4 that the problem becomes well-posed if the SIMP material model is used in a formulation that includes a restriction in the form of, e.g., an upper constraint value on the perimeter of the topology design of a 2D structure, or on the surface area of a 3D structure.

2.3.1 'Hole-in-Cell' Microstructures

Bendsøe & Kikuchi [15], Diaz & Bendsøe[49] and several others have applied a numerically evaluated anisotropic 'hole-in-cell' microstructure as shown in Figure 3(a) that consists of a square cell of isotropic material with a rectangular hole. For the topology optimization, the orientation $\theta(x)$ of the microscopic cells and their geometry defined by $\mu_1(x)$ and $\mu_2(x)$, see Figure 3(a), are applied as design variables. Thus, the model governs void and solid for $\mu_1 = \mu_2 = 0$ and $\mu_1 = \mu_2 = 1$, respectively, and composite for intermediate values of μ_1 and μ_2. The volume density $\rho(x)$ of material $0 \leq \rho(x) \leq 1$ is given by

$$\rho(x) = 1 - \mu_1(x)\mu_2(x) \tag{11}$$

The components of the effective stiffness matrix for the microstructure can be obtained numerically for different sets of values of μ_1 and μ_2 on the basis of homogenization, see Section 3.1.

2.3.2 Layered 2D Microstructures

Referring to Figures 3(b) and 4 we now consider the construction of the class of so-called ranked layered planar microstructures which has been used by a large number of authors for solution of a variety of topology optimization problems for 2D continuum structures.

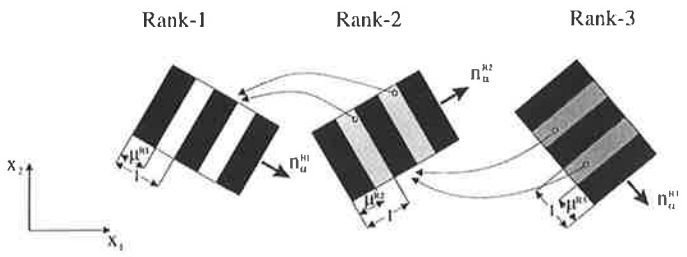


Figure 4: Construction of first, second, and third-rank microstructures by successive layering along different directions. Black and white domains are occupied by isotropic materials with the stiffness tensors $E^{+}_{\alpha\beta\kappa\gamma}$ and $E^{-}_{\alpha\beta\kappa\gamma}$, respectively, while grey domains represent areas occupied by layered materials composed of the two base materials.

The microstructures are obtained by a repetitive process in each step of which a new layering of given direction is added to the microstructure. The number of times this is performed is per definition the *rank* of the microstructure. Hence, given two isotropic materials with elasticity tensors $E^{+}_{\alpha\beta\kappa\gamma} > E^{-}_{\alpha\beta\kappa\gamma}$, a first-rank microstructure is obtained by stacking alternately thin layers of the stiffer material[1] with layers of the more compliant material in the proportion given by μ^{R1} and $1-\mu^{R1}$, respectively, along a direction characterized by a unit normal n^{R1}_{α}. Similarly, a second-rank microstructure is obtained by stacking alternately thin layers of the stiff isotropic material and the first-rank microstructure in the proportion μ^{R2} and $1-\mu^{R2}$, respectively, along a new direction characterized by the unit normal n^{R2}_{α}. The process may, as indicated in Fig. 4, be continued to build layered microstructures of any finite rank RI, and the effective stiffness tensor for the resulting material is symbolically obtained by the sequence

$$E^{Ri}_{\alpha\beta\kappa\gamma} = E(\mu^{Ri}, n^{Ri}_{\alpha}, E^{+}_{\alpha\beta\kappa\gamma}, E^{R(i-1)}_{\alpha\beta\kappa\gamma}) \quad ; \quad i = 1, \ldots, I \tag{12}$$

It is a particular advantage of these layered planar microstructures that their effective material stiffness properties, in the limit where the width of the layers tends to zero, may be calculated analytically applying either a mathematically based homogenization procedure as used by e.g. Bendsøe[16], or a more physically based smear-out procedure as used by e.g. Olhoff, Lurie, Cherkaev & Fedorov[19] and Thomsen[29].

The class of layered planar microstructures just described are of particular importance to problems of topology optimization of plane, 2D continuum structures, because second-rank planar microstructures with orthogonal layers are optimum for single load case stiffness design problems, while third-rank planar microstructures with non-orthogonal layers are optimum for multiple load case stiffness design problems, see Avellaneda[27]. With regard to other topology optimization problems for 2D continuum structures parameterized by means of layered planar microstructures, it is worth mentioning that no further generality is obtained by application of more than a third-rank planar microstructure with

[1]This material will be termed "stiff" in the following but has finite stiffness

non-orthogonal layers, as the whole range of effective stiffness properties for all finite rank planar microstructures is obtainable by application of third-rank planar microstructures with non-orthogonal layers. This follows from work by Avellaneda & Milton[74] and Lipton[28].

2.3.3 Layered 3D Microstructures

Layered microstructures have recently been used also for topology optimization of 3D continuum structures (Cherkaev & Palais [32], Allaire *et al.* [30]), Dias & Lipton [75], Jacobsen *et al.* [76], Olhoff *et al.* [77]), and multiple material problems Burns & Cherkaev [78].

We now consider a 3D layered composite microstructure of rank-three as shown in Figure 5. The three sets of layers are assumed to be mutually orthogonal, and the composite is made of two isotropic, linear elastic materials, and the volume fractions of these are given as m_1 and m_2 ($m_2 = 1 - m_1$).

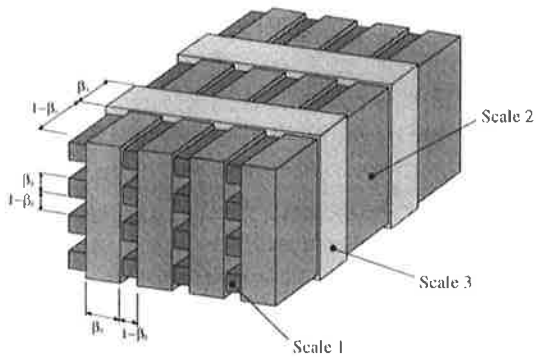


Figure 5: Model of a spatial rank three laminate where β_i denotes the ith length scale.

The length scales (Fig.5) describe the relative thickness of each layer which yields the following relation between volume density ρ and length scales for a rank n material

$$\rho_n = \beta_n + (1 - \beta_n)\rho_{n-1} \tag{13}$$

where, in the case of a rank three material,

$$\beta_1 = m_1\alpha_1 \quad , \quad \beta_2 = \frac{m_1\alpha_2}{1 - m_1\alpha_1} \quad , \quad \beta_3 = \frac{m_1\alpha_3}{1 - m_1(1 - \alpha_3)} \tag{14}$$

The introduced parameter $\alpha_i, i = 1, 2, 3$, is the spatial thickness of layer i which means that it can be considered as a volume density for the layer. By introducing this parameter the basic idea is to equalize the stresses in the different directions by varying the spatial thickness of each layer. The design variables for this 3D microstructure is the volume density ρ, the parameters $\alpha_i, i = 1, 2, 3$, and angles governing the spatial orientation.

2.3.4 SIMP model

For the SIMP (Solid Isotropic Material with Penalty) model, see e.g., Bendsøe [3], Zhou & Rozvany [79], Mlejnek & Schirrmacher [80], and Rozvany *et al.* [81], the elasticity tensor E_{ijkl} for and the volume of a structure made of the material are given by

$$E_{ijkl}(x) = \rho(x)^p E^{\circ}_{ijkl}, \quad p > 1; \quad \text{Volume} = \int_{\Omega} \rho(x) dx \tag{15}$$

where $\rho(x), x \in \Omega, 0 \leq \rho(x) \leq 1$, is a density function of the material and E°_{ijkl} the elasticity tensor of a given solid isotropic reference material. The density function $\rho(x)$ enters the stiffness relation in a power $p > 1$ which has the effect of penalizing intermediate densities $0 < \rho < 1$, since at such densities the SIMP material has lower stiffness than the reference material at the same cost. Figure 6 displays the relative stiffness ratio $\frac{E}{E^{\circ}}$ vs. volume density ρ for different values of the penalization power p, and clearly illustrates that the use of the SIMP material model will force the topology design towards limiting values $\rho = 0$ (void) and $\rho = 1$ (solid) and thereby prompt the creation of more distinctive 0-1 designs. When applying the model, the value of the penalization power p usually is gradually increased from 1 to 4 during the topology design process (the so-called continuation technique, see Sigmund [81]).

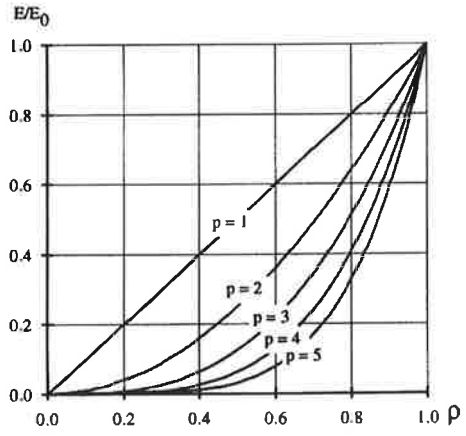


Figure 6: Relative stiffness vs. volume density for the SIMP material model for different values of the penalization power p.

It is a drawback of the SIMP model that the topology designs obtained not only exhibit dependence on the value of p, but also on the finite element mesh applied. However, as will be discussed in Section 4, this drawback disappears, and topology optimization problems based on usage of the SIMP model become well-posed if, e.g., a perimeter or surface area constraint is included in the formulation of the problem.

Even if such a constraint is not included in the formulation, the SIMP model can yield very nice results, and the simplicity of the model greatly facilitates the implementation of topology design in commercial finite element codes. In research, the method has been very useful in investigations into extending the scope of topology optimization, where this simple model allows for emphasizing other aspects of the problems. In recent years, this approach has been employed by, a.o., Bendsøe [2], Bendsøe & Diaz [82], Duysinx & Bendsøe [60], Hinton *et al.* [83], Magister & Post [84], Maute & Ramm [67], [85], Maute *et al.* [68], Petersson & Sigmund [40], Sigmund [63], Sigmund & Torquato [86], Sigmund *et al.* [87] and Yang & Chen [88].

The SIMP model is very often called the 'Artificial' or the 'Fictituous' material model or -interpolation scheme in the literature. Thus, up to now, it has not been apparent whether the 'SIMP material' of intermediate density, $0 < \rho < 1$, could be interpreted physically. However, in very recent work, Bendsøe & Sigmund [89] analyzed the SIMP and similar material models and compared them to variational bounds on effective properties of composite materials, such as the Hashin-Shtrikman bounds [90].

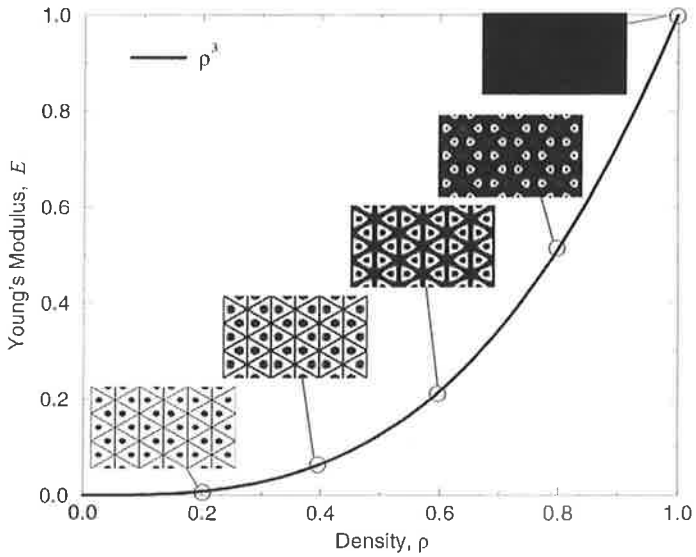


Figure 7: Microstructures of material and void which realize the material properties of the SIMP model with $p = 3$ for a base material with Poisson's ratio $\nu = 1/3$. As stiffer material microstructures (that are closer to the Hashin-Shtrikman upper bound) can be constructed from the given densities ρ, non-structural areas are found at the cell centers. (By courtesy of Bendsøe and Sigmund.)

The investigation by Bendsøe & Sigmund [89] led to the remarkable result that under fairly simple conditions on the penalization power p, any of the isotropic elasticity tensors of the SIMP model (Eq. 15) can be physically realized as the elasticity tensor of a composite

material made of void and an amount of the reference (base) material corresponding to the relevant density, ρ. For example, for a Poisson's ratio of the base material $\nu = 1/3$, both for 2D and 3D problems a penalization power p satisfying $p > 3$ ensures that the SIMP model can be physically realized within the framework of microstructured composite materials [89]. Fig. 7 shows some microstructured realizations in 2D for $\nu = 1/3$ and $p = 3$.

Thus, due to these very interesting recent results of Bendsøe and Sigmund [89], when assuming proper values of p, we may refer to the SIMP model as a microstructured material model in the sequel.

3 Homogenization Techniques for the Relaxation Approach

The relaxation approach to achieve well-posedness of problems of topology optimization is based on the ability to model porous materials with microstructure such as those considered in Sub-sections 2.3.1 - 2.3.3. Each of these porous composite materials are constructed from a particular unit cell which at a macroscopic level consists of material and void, and the composite materials consist of infinitely many of such cells that are now assumed to be infinitely small and to be repeated periodically through the composite material. At this limit, we can also have continuously varying density of material through the porous composite material as required for topology optimization.

The resulting material can be described by effective, macroscopic material properties which depend on the geometry of the particular unit cell. These effective macroscopic properties can be computed on the basis of a mathematically based method of homogenization which will be briefly discussed in Section 3.1, but homogenization may be also carried out via of 'smear-out' or quasi convexification techniques as will be illustrated in Section 3.2 and 3.3, respectively.

3.1 Mathematically based Homogenization

This approach is based on analysis of the unit cell together with an asymptotic expansion of the displacement field of the composite material, see Bensousson *et.al.* [91], Sanchez-Palencia [92] and Bourgat [93].

Assuming that two coordinate systems, X and Y, are defined to describe the material on the macroscopic and on the microscopic scale, respectively, it is assumed hat a periodic microstructure exists in the vicinity of an arbitrary point $\boldsymbol{x}$ of a linearly elastic structure, see Figure 3. The periodicity is represented by a parameter δ which is very small, and the elasticity tensor E^{δ}_{ijkl} is given in the form

$$E^{\delta}_{ijkl}(\boldsymbol{x}) \;=\; E_{ijkl}(\boldsymbol{x}, \boldsymbol{x}/\delta) \tag{16}$$

where $\boldsymbol{y} \rightarrow E_{ijkl}(\boldsymbol{x}, \boldsymbol{y})$ is Y-periodic with unit cell $[0,1] \times \boldsymbol{R}$ of periodicity. Here, $\boldsymbol{x}$ is the macroscopic variation of the material parameters, while $\boldsymbol{x}/\delta$ gives the microscopic, periodic variations.

When the composite structure is subjected to macroscopic loading, the displacement field $\boldsymbol{v}^{\delta}(\boldsymbol{x})$ can be written in the form of an asymptotic expansion with respect to the cell size δ,

$$\boldsymbol{v}^{\delta}(\boldsymbol{x}) \;=\; \boldsymbol{v}_0(\boldsymbol{x}) + \delta \boldsymbol{v}_1(\boldsymbol{x}, \boldsymbol{x}/\delta) \tag{17}$$

Here, the leading term $\boldsymbol{v}_0(\boldsymbol{x})$ is the macroscopic displacement field, which is an average over the cell and hence independent of the microscopic variable $\boldsymbol{y}$. It can be shown (see [91],[92]) that the effective displacement field $\boldsymbol{v}_0(\boldsymbol{x})$ will arise from the applied macroscopic loading if the effective (homogenized) elasticity tensor $E^H_{ijkl}(\boldsymbol{x})$ has the form

$$E^H_{ijkl}(\boldsymbol{x}) \;=\; \frac{1}{|Y|}\int_Y \left[E_{ijkl}(\boldsymbol{x},\boldsymbol{y}) - E_{ijpq}(\boldsymbol{x},\boldsymbol{y})\frac{\partial \chi^{kl}_p}{\partial y_q} \right] d\boldsymbol{y} \tag{18}$$

Here, χ^{kl}_p is a microscopic displacement field which is given as the Y-periodic solution to the cell equilibrium equations

$$\int_Y E_{ijpq}(\boldsymbol{x},\boldsymbol{y})\frac{\partial \chi^{kl}_p}{\partial y_q}\frac{\partial v_i}{\partial y_j}\, d\boldsymbol{y} \;=\; \int_Y E_{ijkl}(\boldsymbol{x},\boldsymbol{y})\frac{\partial v_i}{\partial y_j}\, d\boldsymbol{y} \quad \text{for all } \boldsymbol{v} \tag{19}$$

where $\boldsymbol{v}$ are Y-periodic virtual displacement fields.

The coefficients in the homogenized elasticity tensor E^H_{ijkl} can be determined by solving three and six prestrain analysis problems for the unit cell Y in cases of 2D and 3D, respectively. In most cases this has to be done numerically using finite element analysis, see Bourgat [93] and Guedes & Kikuchi [94]. For the subsequent application in topology optimization, it is then expedient to represent the dependence of the components of the constitutive tensor on the microstructural parameters by means of approximation formulas.

Referring to Chapter 2, the constitutive tensor for the 'hole-in-cell' microstructure (Section 2.3.1) must be determined numerically [9], whereas the tensor can be established analytically for the layered 2D microstructure (Section 2.3.2) by the above approach. With a view to include presentation of other homogenization techniques in this paper, a smear-out and a quasiconvexification approach, respectively, will be presented subsequently by way of the layered 2D microstructure and the layered 3D microstructure described in Chapter 2.

3.2 Layered 2D Microstructures – Homogenization via Smear-Out Technique

Referring to Subsection 2.3.2, we shall now demonstrate how the effective stiffness tensor for layered 2D microstructures can be derived analytically by a simple averaging technique based on the constitutive relationships between average stress and strain tensors and use of appropriate continuity and discontinuity conditions for components of these tensors at the interfaces between the two materials. Such so-called *smear-out* techniques have previously been used by e.g. Olhoff, Lurie, Cherkaev & Fedorov [19] for the derivation of effective bending stiffness tensors for microstructurally layered Kirchhoff plates, later by Thomsen [29] for the derivation of effective material stiffness tensors for first-rank

planar disk microstructures, and recently also by Soto & Diaz [95] and Soto [96] for the derivation of effective bending and transverse shear stiffness tensors for microstructurally layered Mindlin plates.

3.2.1 Microstructure of First Rank

Following the method used in [19] we start by considering a small rectangular plane element Ω as depicted in Figure 8. The size of this element, which consists of a finite

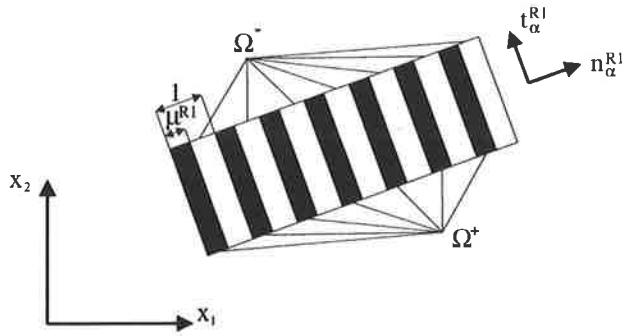


Figure 8: Plane element with an underlying first-rank microstructure.

number of parallel, alternating layers of the stiff and the compliant isotropic materials, respectively, is assumed to be small in comparison with the dimensions of the entire plane structure, but to be large relative to its underlying first-rank microstructure. The orientation of the element in the plane coordinate system x_1x_2, see Figure 8, is given by the unit vectors n_α^{R1} and t_α^{R1} which are perpendicular to and parallel with the layers, respectively, i.e., $n_\alpha^{R1}n_\alpha^{R1} = 1$, $t_\alpha^{R1}t_\alpha^{R1} = 1$, and $n_\alpha^{R1}t_\alpha^{R1} = 0$. The stress-strain field within the element is assumed to be homogeneous from a macroscopic point of view, while at the microscroscale, i.e., at the level of subdomains Ω^+ and Ω^-, the stress-strain field is assumed to be piece-wise homogeneous. We shall denote the stiffness tensors for the stiff and compliant layers located in subdomains Ω^+ and Ω^-, by $A^+_{\alpha\beta\kappa\gamma}$ and $A^-_{\alpha\beta\kappa\gamma}$, respectively. Furthermore, assuming that plane stress conditions prevail, and taking the stiff and the compliant isotropic materials used in the layers to have the Young's moduli E^+ and E^- and a common Poisson's ratio ν, we have the following general relationships between the non-zero components of the stiffness tensors for the two types of layering

$$\begin{aligned} A^{(+/-)}_{1111} = A^{(+/-)}_{2222} = A^{(+/-)}_0 \quad ; \quad A^{(+/-)}_{1122} = A^{(+/-)}_{2211} = \nu A^{(+/-)}_0 \\ A^{(+/-)}_{1212} = A^{(+/-)}_{1221} = A^{(+/-)}_{2121} = A^{(+/-)}_{2121} = \frac{1-\upsilon}{2} A^{(+/-)}_0 \end{aligned} \tag{20}$$

where the stiffness constants A^+_0 and A^-_0 for the stiff and the compliant layers, are given

by

$$A_0^{(+/-)} = \frac{E^{(+/-)}}{1-\nu^2} \tag{21}$$

We now determine the effective stiffness tensor $A^{R1}_{\alpha\beta\kappa\gamma}$ for our first-rank microstructure shown in Figure 8, by firstly considering the constitutive relationship for the small rectangular domain Ω,

$$\sigma^{avr}_{\alpha\beta} = A^{R1}_{\alpha\beta\kappa\gamma}\epsilon^{avr}_{\kappa\gamma} \tag{22}$$

Here, the Greek indices refer to the axes of a global coordinate system x_1x_2 embedded in the plane domain, see Figure 8, and the tensors $\epsilon^{avr}_{\alpha\beta}$ and $\sigma^{avr}_{\alpha\beta}$ are direct averages of strain and stress tensors for the small plane element Ω. Defining homogeneous strain tensors $\epsilon^{+}_{\alpha\beta}$ and $\epsilon^{-}_{\alpha\beta}$ and stress tensors $\sigma^{+}_{\alpha\beta}$ and $\sigma^{-}_{\alpha\beta}$ within subdomains Ω^+ and Ω^-, respectively, we have the following constitutive relations for the subdomains Ω^+ and Ω^-,

$$\sigma^{+}_{\alpha\beta} = A^{+}_{\alpha\beta\kappa\gamma}\epsilon^{+}_{\kappa\gamma} \quad ; \quad \sigma^{-}_{\alpha\beta} = A^{-}_{\alpha\beta\kappa\gamma}\epsilon^{-}_{\kappa\gamma} \tag{23}$$

and the following expressions for the direct averages of the strain and stress tensors within the domain Ω,

$$\epsilon^{avr}_{\alpha\beta} = \mu^{R1}\epsilon^{+}_{\alpha\beta} + (1-\mu^{R1})\epsilon^{-}_{\alpha\beta} \quad ; \quad \sigma^{avr}_{\alpha\beta} = \mu^{R1}\sigma^{+}_{\alpha\beta} + (1-\mu^{R1})\sigma^{-}_{\alpha\beta} \tag{24}$$

The problem is now to establish an analytical expression for the effective stiffness tensor $A^{R1}_{\alpha\beta\kappa\gamma}$ as a function of the density variable μ^{R1} and the layering orientation n^{R1}_{α}, see Figure 8. Such an expression can be derived by usage of Eqs.(22)-(24) and two sets of interface conditions that express the mechanics of the microstructure.

Firstly, from considerations of static equilibrium across the interfaces between adjacent subdomains Ω^+ and Ω^- we can write the following set of continuity/discontinuity conditions for the components of the stress tensors acting parallel and orthogonal to the interface,

$$(\sigma^{+}_{\alpha\beta} - \sigma^{-}_{\alpha\beta})n^{R1}_{\alpha}n^{R1}_{\beta} = 0 \quad ; \quad (\sigma^{+}_{\alpha\beta} - \sigma^{-}_{\alpha\beta})t^{R1}_{\alpha}t^{R1}_{\beta} \neq 0 \quad ; \quad (\sigma^{+}_{\alpha\beta} - \sigma^{-}_{\alpha\beta})n^{R1}_{\alpha}t^{R1}_{\beta} = 0 \tag{25}$$

Similarly, by geometric compatibility considerations we may set up the following continuity/discontinuity conditions for the components of the strain tensors acting parallel and orthogonal to the interface:

$$(\epsilon^{+}_{\alpha\beta} - \epsilon^{-}_{\alpha\beta})n^{R1}_{\alpha}n^{R1}_{\beta} \neq 0 \quad ; \quad (\epsilon^{+}_{\alpha\beta} - \epsilon^{-}_{\alpha\beta})t^{R1}_{\alpha}t^{R1}_{\beta} = 0 \quad ; \quad (\epsilon^{+}_{\alpha\beta} - \epsilon^{-}_{\alpha\beta})n^{R1}_{\alpha}t^{R1}_{\beta} \neq 0 \tag{26}$$

However, these interface conditions may be written in the more compact form

$$\epsilon^{+}_{\alpha\beta} - \epsilon^{-}_{\alpha\beta} = \beta_1 n^{R1}_{\alpha}n^{R1}_{\beta} + \beta_2(t^{R1}_{\alpha}n^{R1}_{\beta} + n^{R1}_{\alpha}t^{R1}_{\beta}) \tag{27}$$

where the scalars β_1 and β_2 designate the jumps of those components of the strain tensor that exhibit discontinuity across the interfaces between the Ω^+ and Ω^- subdomains.

Eqs. (22)-(25) and Eq. (27) together with Eqs. (20)-(21) constitute the set of equations needed for the derivation of the effective stiffness tensor. The remaining part of the derivation consists of simple but cumbersome algebraic manipulations of these equations (the reader is referred to Krog [97] and Krog & Olhoff [98]) which yield the following analytical expressions for the effective stiffness tensor for the first-rank 2D microstructure,

$$A^{R1}_{\alpha\beta\kappa\gamma} = A^{+}_{\alpha\beta\kappa\gamma} - (1-\mu^{R1})\left((A^{+}_{\alpha\beta\kappa\gamma} - A^{-}_{\alpha\beta\kappa\gamma})^{-1} - \frac{\mu^{R1}}{A^{+}_{0}} \Lambda^{R1}_{\alpha\beta\kappa\gamma} \right)^{-1} \tag{28}$$

with specific expressions for stiffness entries defined in Eq. (20). Here, the tensor $\Lambda^{R1}_{\alpha\beta\kappa\gamma}$ contains information about layer orientation and is defined as

$$\begin{aligned} \Lambda^{R1}_{\alpha\beta\kappa\gamma} = \; & n^{R1}_{\alpha} n^{R1}_{\beta} n^{R1}_{\kappa} n^{R1}_{\gamma} + \frac{1}{2(1-\nu)} \big(\, t^{R1}_{\alpha} n^{R1}_{\beta} n^{R1}_{\kappa} t^{R1}_{\gamma} + n^{R1}_{\alpha} t^{R1}_{\beta} n^{R1}_{\kappa} t^{R1}_{\gamma} \\ & + \; t^{R1}_{\alpha} n^{R1}_{\beta} t^{R1}_{\kappa} n^{R1}_{\gamma} + n^{R1}_{\alpha} t^{R1}_{\beta} t^{R1}_{\kappa} n^{R1}_{\gamma} \, \big) \end{aligned} \tag{29}$$

It is worth mentioning that the isotropy of the stiffness tensor $A^{+}_{\alpha\beta\kappa\gamma}$ for the material in the subdomains Ω^{+} has been used in the derivation of Eq. (28), while no such restriction has been imposed on the stiffness tensor $A^{-}_{\alpha\beta\kappa\gamma}$ for the material in the subdomains Ω^{-}. This important feature makes these equations applicable for calculation of effective stiffness properties of multi-rank microstructures.

3.2.2 Effective Stiffness in Matrix Form

A convenient matrix description of the effective stiffness properties can be obtained if we introduce the following orthonormal basis of symmetric second order tensors

$$\xi^1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \xi^2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \xi^3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{30}$$

In this basis, any symmetric second order tensor is represented as a vector, and any symmetric fourth order tensor as a symmetric 3×3 matrix. Thus, the symmetric fourth order stiffness tensor $A_{\alpha\beta\kappa\gamma}$ for the plane disk may be represented as a symmetric 3×3 matrix $\mathbf{A}$ with components A_{ij} defined by

$$\mathbf{A} = [A_{ij}] = [\xi^{i}_{\alpha\beta} \xi^{j}_{\kappa\gamma} A_{\alpha\beta\kappa\gamma}] = \begin{bmatrix} \frac{1}{2}(A_{1111} + A_{2222}) - A_{1122} & A_{1112} - A_{2221} & \frac{1}{2}(A_{1111} - A_{2222}) \\ & 2A_{1212} & A_{1112} + A_{2221} \\ sym & & \frac{1}{2}(A_{1111} + A_{2222}) + A_{1122} \end{bmatrix} \tag{31}$$

With this transformation at hand, we are able to transform the fourth order tensor equation for the effective stiffness tensor in the preceding Section into a a simple matrix equation. The "inverse transformation" for computation of the components of the original stiffness tensor from the components of the matrix $\mathbf{A}$ can be seen in [97] and [98].

Now, applying the transformation in Eq.(31) to the tensor equation for the effective stiffness tensor, $A^{R1}_{\alpha\beta\kappa\gamma}$ in Eqs.(28), and by using the following relation

$$\xi^i_{\alpha\beta}\xi^j_{\kappa\gamma}\left(A_{\alpha\beta\kappa\gamma}\right)^{-1} = \left(\xi^i_{\alpha\beta}\xi^j_{\kappa\gamma}A_{\alpha\beta\kappa\gamma}\right)^{-1} \tag{32}$$

which holds for fourth order tensors with all major and minor symmetries, we get the following matrix equation:

$$\mathbf{A}^{R1} = [A^{R1}_{ij}] = [\xi^i_{\alpha\beta}\xi^j_{\kappa\gamma}A^{R1}_{\alpha\beta\kappa\gamma}] = \mathbf{A}^{+} - (1-\mu^{R1})\left((\mathbf{A}^{+}-\mathbf{A}^{-})^{-1} - \frac{\mu^{R1}}{A^{+}_{0}}\mathbf{\Lambda}^{R1}\right)^{-1} \tag{33}$$

Here $\mathbf{A}^{+}$, $\mathbf{A}^{-}$ and $\mathbf{\Lambda}^{R1}$ are (3x3) matrices obtained by using the transformations in Eq. (31) to the fourth order tensors $A^{+}_{\alpha\beta\kappa\gamma}$, $A^{-}_{\alpha\beta\kappa\gamma}$ and $\Lambda^{R1}_{\alpha\beta\kappa\gamma}$. Adopting the notation introduced in Eqs. (20), and using the isotropy of the tensors $A^{+}_{\alpha\beta\kappa\gamma}$ and $A^{-}_{\alpha\beta\kappa\gamma}$, we get the following expressions for the matrices $\mathbf{A}^{+}$ and $\mathbf{A}^{-}$,

$$\mathbf{A}^{(+/-)} = [A^{(+/-)}_{ij}] = [\xi^i_{\alpha\beta}\xi^j_{\kappa\gamma}A^{(+/-)}_{\alpha\beta\kappa\gamma}] = A^{(+/-)}_{0}\begin{bmatrix} 1-\nu & 0 & 0 \\ 0 & 1-\nu & 0 \\ 0 & 0 & 1+\nu \end{bmatrix} \tag{34}$$

Finally, taking the normal and tangential vectors, n^{R1}_{α} and t^{R1}_{α}, in the expression for the tensor $\Lambda^{R1}_{\alpha\beta\kappa\gamma}$ given in Eq. (29) as $n^{R1}_{\alpha} = \{\cos(\theta^{R1}), sin(\theta^{R1})\}^T$ and $t^{R1}_{\alpha} = \{-sin(\theta^{R1}), \cos(\theta^{R1})\}^T$, we obtain the following expression for the matrix $\mathbf{\Lambda}^{R1}$,

$$\mathbf{\Lambda}^{R1} = [\xi^i_{\alpha\beta}\xi^j_{\kappa\gamma}\Lambda^{R1}_{\alpha\beta\kappa\gamma}] = \begin{bmatrix} \dfrac{3-\nu-(1+\nu)\cos(4\theta^{R1})}{4(1-\nu)} & -\dfrac{(1+\nu)sin(4\theta^{R1})}{4(1-\nu)} & \dfrac{\cos(2\theta^{R1})}{2} \\ & \dfrac{3-\nu+(1+\nu)\cos(4\theta^{R1})}{4(1-\nu)} & \dfrac{sin(2\theta^{R1})}{2} \\ sym & & \dfrac{1}{2} \end{bmatrix} \tag{35}$$

The matrix expressions in Eqs.(33)-(35) constitute a simple set of matrix equations for the components of the effective stiffness tensor for the first-rank 2D microstructures describedin Sub-section 2.3.2.

3.2.3 Multi-Rank Layered 2D Microstructures

Analytical expressions for the effective stiffness matrix for layered 2D microstructures of any finite rank may, as indicated in Eq. (12), be established iteratively. Thus the effective stiffness properties of a rank-(i+1) microstructure may be determined by using the expressions for the effective stiffnesses of a first-rank microstructure given in Eqs. (33)-(35) with the effective stiffness for a rank-i microstructure as the stiffness of the compliant material.

Following derivations in Soto [96], see also Lipton [99], we obtain the following expression for the effective stiffness matrix, see [54], [96], [97] and [56],

$$\mathbf{A}^{RI} = \mathbf{A}^{+} - (1-\rho^{RI})\left((\mathbf{A}^{+}-\mathbf{A}^{-})^{-1} - \frac{\rho^{RI}}{A_0^{+}}\sum_{i=1}^{I} p_i \mathbf{\Lambda}^{Ri}\right)^{-1} \tag{36}$$

where the factors p_i for the rank-I microstructure obey the relationship

$$\sum_{i=1}^{I} p_i = 1 \;\; ; \quad p_i \geq 0 \;\; , \quad i = 1, \ldots, I \tag{37}$$

and the layer orientation matrix $\mathbf{\Lambda}^{Ri}$ is given by

$$\mathbf{\Lambda}^{Ri} = \begin{bmatrix} \dfrac{3-\nu-(1+\nu)\cos(4\theta^{Ri})}{4(1-\nu)} & -\dfrac{(1+\nu)sin(4\theta^{Ri})}{4(1-\nu)} & \dfrac{\cos(2\theta^{Ri})}{2} \\ & \dfrac{3-\nu+(1+\nu)\cos(4\theta^{Ri})}{4(1-\nu)} & \dfrac{sin(2\theta^{Ri})}{2} \\ sym & & \dfrac{1}{2} \end{bmatrix} \tag{38}$$

where (compare with Eq. (35) and Fig. 4), θ^{Ri} is the angle between the global x_1 axis and the unit normal vector n_α^{Ri} to the layering made at the i'th step of the construction of the rank-I microstructure.

Notice that Eq. (36) gives a mathematical expression for the effective stiffness matrix for a rank-I layered 2D microstructure where the microstructure is described in terms of $(2I+1)$ parameters, namely the total density ρ^{RI} of stiff material in the microstructure, the I new density variables p_i, and the I layer orientations in the microstructure given by the angles of rotation θ^{Ri}, $i = 1, \ldots, I$ through Eq. (38). The new density variables p_i, $i = 1, \ldots, I$, of which only $I-1$ are mutually independent in view of Eq. (37), are defined by

$$p_i = \frac{(1-\rho^{R(i-1)})\mu^{Ri}}{\rho^{RI}} \;\; , \quad i = 1, \ldots, I \tag{39}$$

3.2.4 Special Case: Second-Rank 2D Microstructure with Orthogonal Layers

Second-rank layered microstructures with orthogonal layers play an important role in the design parameterization of topology optimization problems for planar disks, because such microstructures are optimum for single load case stiffness design problems for planar disks. In the sequel, we therefore write the analytical expressions for the components of the effective membrane stiffness tensor for second-rank disk microstructure with orthogonal layers. The tensor is expressed relative to a fixed material coordinate system $x_1'x_2'$ with axes parallel to the layering directions of the microstructure, see Fig. 9, and the results will therefore only depend on the two density variables μ^{R1} and μ^{R2}. The dependence

on the layer directions is described via a third design variable θ which gives the overall rotation of the orthotropic microstructure relative to a fixed global coordinate system $x_1 x_2$, see Fig. 9. The dependence of the effective stiffness tensor on this design variable is simply accounted for via transformation formulas for rotation of anisotropic materials, see e.g. [100].

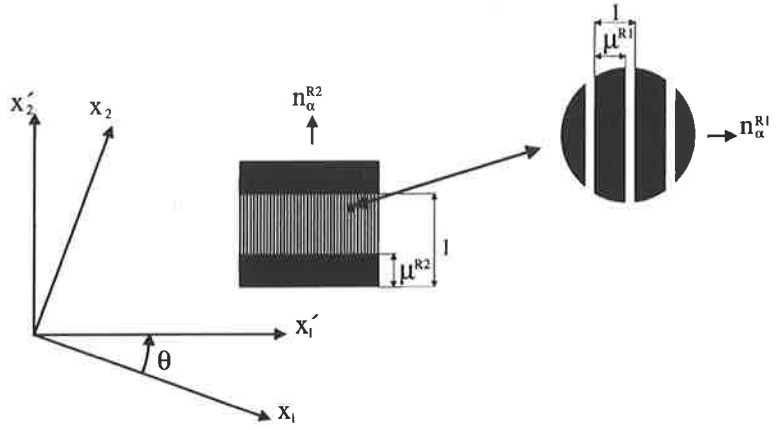


Figure 9: Second-rank microstructure with orthogonal layers.

From Eqs. (36), (37) and (39) we get the following expression for the effective membrane stiffness matrix for a second-rank 2D microstructure with non-orthogonal layers,

$$\mathbf{A}^{R2} = \mathbf{A}^{+} - (1-\mu^{R1})(1-\mu^{R2}) \left[(\mathbf{A}^{+} - \mathbf{A}^{-})^{-1} - \frac{\mu^{R1}\mathbf{\Lambda}^{R1} + (1-\mu^{R1})\mu^{R2}\mathbf{\Lambda}^{R2}}{A_0^{+}} \right]^{-1} \tag{40}$$

The stiffness matrices $\mathbf{A}^{+}$ and $\mathbf{A}^{-}$ which characterize the constituent layers of the microstructure are defined in Eq. (34), while the geometric matrices $\mathbf{\Lambda}^{R1}$ and $\mathbf{\Lambda}^{R2}$ which depend on the normal vectors n_α^{R1} and n_α^{R2} are defined in Eq(38). Taking the latter to be orthogonal and given by the angles $\theta^{R1} = 0°$ and $\theta^{R2} = 90°$ with respect to the x_1' axis we obtain the following expressions for the geometric matrices for a second-rank disk microstructure with orthogonal layers,

$$\mathbf{\Lambda}^{R1} = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{1-\nu} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} \quad ; \quad \mathbf{\Lambda}^{R2} = \begin{bmatrix} \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & \frac{1}{1-\nu} & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} \tag{41}$$

With the definitions in Eq. (34) and Eq. (41) and the expression in Eq. (40), we can now write an analytical expression for the effective stiffness matrix for the microstructure.

By an inverse transformation we have then established the expressions given below for the non-zero components of the effective membrane stiffness tensor for a second-rank 2D microstructure with orthogonal layers, where the density variables μ^{R1} and μ^{R2} have been written by the short-hand symbols μ_1 and μ_2,

$$A^{R2}_{1111} = \frac{A^{R2}_0 + A^{R2}_1 (\mu_2 - \mu_1\mu_2 - \mu_2^2 + 2\mu_1\mu_2^2 - \mu_1^2\mu_2^2)}{A^{R2}_2 + A^{R2}_3} \quad ; \quad A^{R2}_{2222} = \frac{A^{R2}_0 + A^{R2}_1 (\mu_1 - \mu_1^2)}{A^{R2}_2 + A^{R2}_3}$$

$$A^{R2}_{1122} = \frac{\nu A^{R2}_0 + \nu A^{R2}_1 (\mu_1\mu_2 - \mu_1^2\mu_2)}{A^{R2}_2 + A^{R2}_3} \quad ; \quad A^{R2}_{1212} = \frac{(1-\nu) A^-_0 A^+_0}{2A^{R2}_4} \tag{42}$$

with

$$A^{R2}_0 = A^-_0 (A^+_0)^2 \quad ; \quad A^{R2}_1 = (1-\nu^2) \left(A^-_0 - A^+_0\right)^2 A^+_0$$

$$A^{R2}_2 = (A^+_0)^2 + A^+_0 \left(A^-_0 - A^+_0\right)(\mu_1 + \mu_2 - \mu_1\mu_2) \quad ; \quad A^{R2}_3 = (1-\nu^2) \left(A^-_0 - A^+_0\right)^2 (\mu_1\mu_2 - \mu_1^2\mu_2)$$

$$A^{R2}_4 = A^+_0 + \left(A^-_0 - A^+_0\right)(\mu_1 + \mu_2 - \mu_1\mu_2) \tag{43}$$

Corresponding expressions for the effective bending and transverse shear stiffness tensors for a second-rank Mindlin plate microstructure with orthogonal layers are given in [97] and [56].

In topology optimization of disks we may now apply a design parameterization where the material in each domain of the finite element discretized structure is modelled as an orthotropic material with the appropriate stiffness properties of the second-rank microstructure written above. This yields a convenient finite parameter, continuous formulation of the topology optimization problem. At this point we shall not discuss the type of behavioural objective and constraint functions which may be considered using such a design parameterization, but only present the following general formulation of the topology optimization problem:

$$\begin{aligned}
&\textit{Objective}: \quad \underset{\mu_{1e},\mu_{2e},\theta_e}{\textit{min or max}} \left[\, f(\mu_{1e},\mu_{2e},\theta_e) \,\right] , && e = 1,\ldots,N_e \\
&\textit{Subject to}: \quad g_j(\mu_{1e},\mu_{2e},\theta_e) \le 0 , && e = 1,\ldots,N_e \,, - \;\; j = 1,\ldots,p \\
&\qquad\qquad\quad h_j(\mu_{1e},\mu_{2e},\theta_e) = 0 , && e = 1,\ldots,N_e \,, \;\; j = 1,\ldots,q \\
&\qquad\qquad\quad 0 \le \mu_{1e} \le 1 , && e = 1,\ldots,N_e \\
&\qquad\qquad\quad 0 \le \mu_{2e} \le 1 , && e = 1,\ldots,N_e
\end{aligned} \tag{44}$$

Here f, g_j, and h_j are objective and constraint functions for the optimization, and N_e denotes the total number of finite elements of the discretized structure. It is important to

realize that this problem is inherently a very large scale problem; the number of design variables equals three times the number of finite elements of the discretized structure. This naturally imposes some restrictions on the type of objective and constraint functions that may be considered, and the problem generally requires usage of special optimization procedures in order to deal with the many design variables in an efficient way.

3.2.5 The Moment Formulation

While the second-rank 2D microstructure with orthogonal layers considered in the preceding sub-section is optimum for single load case stiffness design problems for disks, a third-rank microstructure with non-orthogonal layers is optimum for the corresponding multiple load case design problem for disks. For other types of planar design problems, where the optimum microstructure is unknown, one should use as general a microstructure as possible, but also here a design parametrization based on a third-rank microstructure with non-orthogonal layers seems reasonable; the full range of effective stiffness properties for all finite-rank 2D microstructures are described by only five independent variables, and the effective stiffness properties may be calculated analytically using the expressions in Eqs. 36, 37 and 38.

However, a design parameterization of stiffnesses directly in terms of layer directions is often associated with problems of nonconvexity as there exist local optimum solutions with respect to the layer directions, see e.g. Pedersen [101], [102], [103], and this implies that usual solution methods based on sensitivity analysis and mathematical programming normally fail.

To illustrate how these problems can be overcome, in the following we therefore consider a special restatement of the expressions for the effective stiffness properties of planar microstructures in terms of so-called *geometric moment variables*. Moment formulations for description of effective stiffness properties of composite laminates and layered microstructures were introduced by Tsai & Pagano [104] and Francfort & Murat [105] and later applied by Miki [106] as an effective parameterization to render laminate optimization problems in a convex formulation, and by Avellaneda & Milton [74] in the derivation of bounds on the elastic stiffness tensors for layered planar microstructures of arbitrary rank. There is an immediate connection between the use of moment variables in topology design and in design of laminates, see, e.g., Fukunaga & Vanderplaats [107], Miki & Sugiama [108], Grenestedt & Gudmondson [109], Hammer *et.al.* [110],[111],[112], Foldager [113], [114] and Foldager *et al.* [115], and references cited therein for optimum design of laminates.

In relation to topology optimization problems, moment formulations have been used by e.g. Lipton [28], Diaz *et al.* [54], Krog [97], Soto [96] and Krogh & Olhoff [56], [98] for the description of effective stiffness properties of microstructurally layered Kirchhoff and Mindlin plates, and by Lipton & Diaz [116] for description of effective stiffness properties of three dimensional microstructures. The moment formulation for layered planar microstructures considered here follows the developments by Diaz *et al.* [54], Soto [96], and Krogh & Olhoff [98], and it is attractive for the following reasons:

- As shown in [54] and [96], the moment formulation yields a full description of the effective stiffness properties for all finite-rank 2D microstructures using only five design variables, see also [28].
- For a fixed strain field, the strain energy density has been found to be concave in the set of variables used to describe the anisotropy of the microstructure, see [54] and [28]. Thus, for stiffness design problems, the possibility of convergence to local extrema is eliminated in the computation of optimum layer orientations.
- For other types of problems than stiffness design problems, application of periodic functions of layering orientations generally is avoided in the expressions for the effective stiffness properties, so difficulties with local extrema with respect to layer orientations also can be expected to be eliminated for those problems.

To establish the moment formulation for a general rank-I layered planar microstructure we now in the set of Eqs. (36), (37) and (38) for the effective stiffness matrix perform a variable substitution in which we condense all information on the layer densities and orientations given by the design variables p_i and θ^{Ri}, $i = 1, \ldots, I$, into four new variables, the moment variables $m_1, \ldots, m_4$, which we define as

$$\begin{aligned} m_1 &= \sum_{i=1}^{I} p_i cos(2\theta^{Ri}) \ , \qquad m_2 = \sum_{i=1}^{I} p_i sin(2\theta^{Ri}) \\ m_3 &= \sum_{i=1}^{I} p_i cos(4\theta^{Ri}) \ , \qquad m_4 = \sum_{i=1}^{I} p_i sin(4\theta^{Ri}) \end{aligned} \tag{45}$$

These moment variables together with the condition $\sum_{i=1}^{I} p_i = 1$ in Eq. (37) and the expressions in Eqs. (38) allow us to define the following matrix

$$\mathbf{M} = \sum_{i=1}^{I} p_i \Lambda^{Ri} = \begin{bmatrix} \frac{3 - \nu - (1+\nu)m_3}{4(1-\nu)} & -\frac{(1+\nu)m_4}{4(1-\nu)} & \frac{m_1}{2} \\ & \frac{3 - \nu + (1+\nu)m_3}{4(1-\nu)} & \frac{m_2}{2} \\ sym & & \frac{1}{2} \end{bmatrix} \tag{46}$$

which we may use to rewrite the expression for the effective stiffness matrix in Eq. (36) in the following convenient form

$$\mathbf{A}^{RI} = \mathbf{A}^{+} - (1 - \rho^{RI}) \left(\left(\mathbf{A}^{+} - \mathbf{A}^{-} \right)^{-1} - \frac{\rho^{RI}}{A_0^{+}} \mathbf{M} \right)^{-1} \tag{47}$$

Hence, by the variable substitution in Eq. (45) we have obtained a very simple expression which gives the full range of effective stiffness properties of all finite-rank planar

microstructures using only five design variables, namely the density ρ^{RI} of layers of stiff material in the microstructure and the four moment variables $m_1 - m_4$ defined in Eqs. (45). It follows from Eqs. (45) that the moment variables, in general, must satisfy the simple side constraints

$$-1 \leq m_i \leq 1 \quad , \quad i = 1, \ldots, 4 \tag{48}$$

The moment variables $m_1, \ldots, m_4$ are obviously not mutually independent due to the trigonometric relations between the functions that define the moments, and there therefore exist some additional conditions which restrict the combinations of the moment variables to some feasible domain. This set of additional constraints on the moment variables was first established by Avellaneda & Milton [74] via the solution to the trigonometric moment problem treated in Krein & Nudelman [117]. Hereby, the following two additional constraint equations for the moment variables were derived:

$$g_1(m_1, m_2) = m_1^2 + m_2^2 \leq 1$$

$$g_2(m_1, m_2, m_3, m_4) = 2m_1^2(1 - m_3) + 2m_2^2(1 + m_3) + (m_3^2 + m_4^2) - 4m_1 m_2 m_4 \leq 1 \tag{49}$$

For topology and layout optimization problems for 2D continuum structures we may now apply a design parametrization where the material in each subdomain of a finite element discretized structure is modelled as an anisotropic material with stiffness properties given by Eqs. (45)-(47). The general topology or layout optimization problem may then be stated as follows,

$$\begin{aligned}
&\textit{Objective}: \quad \underset{\rho_e^{RI}, m_{1e}, m_{2e}, m_{3e}, m_{4e}}{\textit{min or max}} \quad [\, f(\rho_e^{RI}, m_{1e}, m_{2e}, m_{3e}, m_{4e}) \,] \quad , \qquad e = 1, \ldots, N_e \\
&\textit{Subject to}: \quad g_j(\rho_e^{RI}, m_{1e}, m_{2e}, m_{3e}, m_{4e}) \leq 0 \quad , \qquad e = 1, \ldots, N_e \; , \quad j = 1, \ldots, p \\
&\qquad\qquad h_j(\rho_e^{RI}, m_{1e}, m_{2e}, m_{3e}, m_{4e}) = 0 \quad , \qquad e = 1, \ldots, N_e \; , \quad j = 1, \ldots, q \\
&\qquad\qquad g_{1e}(m_{1e}, m_{2e}) \leq 1 \quad , \qquad e = 1, \ldots, N_e \\
&\qquad\qquad g_{2e}(m_{1e}, m_{2e}, m_{3e}, m_{4e}) \leq 1 \quad , \qquad e = 1, \ldots, N_e \\
&\qquad\qquad 0 \leq \rho_e^{RI} \leq 1 \quad , \qquad e = 1, \ldots, N_e \\
&\qquad\qquad -1 \leq m_{ie} \leq 1 \quad , \qquad i = 1, \ldots, 4 \; , \quad e = 1, \ldots, N_e
\end{aligned} \tag{50}$$

where f, g_j, and h_j are the objective and constraint functions for the optimization, g_{1e} and g_{2e} are local element constraints which must be satisfied by the set of moment variables associated with each finite element, and N_e denotes the total number of elements in the finite element discretized structure. As compared with the optimization problem

formulated in Eqs. (44) in Sub-section 3.2.4, the optimization problem in Eqs. (50) is much larger both in terms of the number of design variables and constraint equations. It should be noted that the solution to the optimization problem in Eqs. (50) directly yields the optimum material densities, and hence the global distribution of material over the structure. The solution also contains the optimum values of the moment variables, but these do not directly yield the optimum layer densities and orientations. A method for determining the local microstructure that corresponds to a given combination of the moment variables can be found in e.g. Lipton [28], Diaz *et al.* [54], and Soto [96].

3.3 Layered 3D Microstructures - Homogenization by Quasiconvexification

We now consider the layered 3D microstructure discussed in Section 2.3.3 and shown in Figure 5, and follow the analysis in Gibiansky and Cherkaev [31] in establishing analytical expressions for the complementary energy density E and elastic properties of this 3D microstructure, that can be shown to be optimum in terms of stiffness subject to any given three-dimensional state of stress. This is achieved by construction of an upper and a lower bound $E_l \leq E \leq E_u$ for the complementary elastic energy density E of the optimum microstructure. The upper bound E_u will be established by minimization of the complementary energy density of a bi-material matrix layered composite of any rank, and the lower bound E_l (which is valid for any bi-material microstructure) will be developed by means of quasiconvexification. The optimum characteristics of the microstructure for topology optimization for maximum stiffness (minimum compliance) are derived subsequently by utilizing the fact (Gibianski and Cherkaev [31]) that in the limit where the compliance of one of the two materials tends to infinity in order to mimic void, the above mentioned bounds coalesce with E so that $E_l = E = E_u$, and from this follow the desired analytical results. Now, the composite in Figure 6 is assumed to be made of two materials with 6 x 6 dimensional compliance matrices denoted by $\mathbf{C}_1$ and $\mathbf{C}_2$, respectively. The compliance matrices which are the inverted elasticity matrices are given by

$$\mathbf{C}(\kappa,\mu) = \begin{bmatrix} \frac{3\kappa+\mu}{9\kappa\mu} & \frac{2\mu-3\kappa}{18\kappa\mu} & \frac{2\mu-3\kappa}{18\kappa\mu} & 0 & 0 & 0 \\ \frac{2\mu-3\kappa}{18\kappa\mu} & \frac{3\kappa+\mu}{9\kappa\mu} & \frac{2\mu-3\kappa}{18\kappa\mu} & 0 & 0 & 0 \\ \frac{2\mu-3\kappa}{18\kappa\mu} & \frac{2\mu-3\kappa}{18\kappa\mu} & \frac{3\kappa+\mu}{9\kappa\mu} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2\mu} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2\mu} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2\mu} \end{bmatrix} \tag{51}$$

where κ and μ are the bulk and shear moduli respectively, and where the indices of the two materials have been omitted. The effective elastic properties of the resulting anisotropic

material depend on the material properties of the two constituents given by $\mathbf{C}_1$, $\mathbf{C}_2$, the normals $\mathbf{n}$ to the material interfaces, and the volume fractions m_1 and m_2 ($m_2 = 1 - m_1$) of the two materials.

As a function of the parameter α_i representing the spatial thickness of layer i in the microstructure, the expression for the optimum compliance tensor $\mathbf{C}_M$, for a matrix layered composite of any rank n is given by

$$\mathbf{C}_M = \mathbf{C}_1 + \mathbf{Q}^{-1} \tag{52}$$

where

$$\mathbf{Q} = \frac{1}{m_2}(\mathbf{C}_2 - \mathbf{C}_1)^{-1} + \frac{m_1}{m_2}\sum_{i=1}^{n}\alpha_i \mathbf{N}_i \tag{53}$$

$$\mathbf{N}_i = \mathbf{N}(\mathbf{n}_i) = \mathbf{p}_i(\mathbf{p}_i^T \mathbf{C}_1 \mathbf{p}_i)^{-1}\mathbf{p}_i^T$$

$$\sum_{i=1}^{n}\alpha_i = 1 \quad , \quad \alpha_i \geq 0$$

The projection matrix $\mathbf{p}_i$ controls discontinuous stress components across the material interphases, and $\mathbf{n}_i$ is the normal to the ith layer. Inserting this result in the expression for the complementary energy density, we get the following equation for a minimum value of the upper bound energy density E_u with respect to the variables α_i and $\mathbf{n}_i$

$$E_u = \sigma^T \mathbf{C}_1 \sigma + \underset{\alpha_i,\, \mathbf{n}_i}{min}\ \sigma^T \mathbf{Q}^{-1}\sigma \tag{54}$$

where σ is a 6 x 1 of vector of stress components defined by $\sigma = (\sigma_{11}\ \sigma_{22}\ \sigma_{33}\ \sigma_{12}\ \sigma_{13}\ \sigma_{23})^T$. Assuming that we are dealing with a rank three material that is oriented along the principal stress directions, we can write

$$\sum_{i=1}^{3}\alpha_i \mathbf{N}_i = \frac{2\mu_1}{3\kappa_1 + 4\mu_1}\begin{bmatrix} \mathbf{B}_1 & 0 \\ 0 & \mathbf{B}_2 \end{bmatrix} \tag{55}$$

where the matrices $\mathbf{B}_1$ and $\mathbf{B}_2$ are given by

$$\mathbf{B}_1 = \begin{bmatrix} 2(3\kappa_1 + \mu_1)(\alpha_2 + \alpha_3) & (3\kappa_1 - 2\mu_1)\alpha_3 & (3\kappa_1 - 2\mu_1)\alpha_2 \\ (3\kappa_1 - 2\mu_1)\alpha_3 & 2(3\kappa_1 + \mu_1)(\alpha_1 + \alpha_3) & (3\kappa_1 - 2\mu_1)\alpha_1 \\ (3\kappa_1 - 2\mu_1)\alpha_2 & (3\kappa_1 - 2\mu_1)\alpha_1 & 2(3\kappa_1 + \mu_1)(\alpha_1 + \alpha_2) \end{bmatrix}$$

$$\mathbf{B}_2 = \begin{bmatrix} (3\kappa_1 + 4\mu_1)\alpha_3 & 0 & 0 \\ 0 & (3\kappa_1 + 4\mu_1)\alpha_2 & 0 \\ 0 & 0 & (3\kappa_1 + 4\mu_1)\alpha_1 \end{bmatrix} \tag{56}$$

respectively.

Referring to Gibiansky & Cherkaev [31] and the initial discussion in this Section, the expression for maximum value of the lower bound E_l on the complementary energy density E of any microstructure made of two elastic materials with volume fractions m_1 and m_2 $(= 1 - m_1)$ and compliance matrices $\mathbf{C}_1$ and $\mathbf{C}_2$, is given by

$$E_l \;=\; \max_{a_j} \;\sigma^T \,\beta\,(\mathbf{C}_1\,,\;\mathbf{C}_2\,,\;\mathbf{\Phi}\,(a_j\,))\;\sigma \tag{57}$$

with the compliance matrix $\beta(\mathbf{C}_1, \mathbf{C}_2, \mathbf{\Phi}(a_j))$ defined by

$$\beta \;=\; \Big[\;\big[\,m_1\;(\;\mathbf{C}_1\;-\;\mathbf{\Phi}\;)^{-1}\;+\;m_2\;(\mathbf{C}_2 - \mathbf{\Phi})^{-1}\big]^{-1}\;+\;\mathbf{\Phi}\;\Big] \tag{58}$$

where

$$\mathbf{\Phi} \;=\; \begin{bmatrix} a_1^2 & -a_1\,a_2 & -a_1\,a_3 & 0 & 0 & 0 \\ -a_1\,a_2 & a_2^2 & -a_2\,a_3 & 0 & 0 & 0 \\ -a_1\,a_3 & -a_2\,a_3 & a_3^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_2^2 + a_3^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & a_1^2 + a_3^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & a_1^2 + a_2^2 \end{bmatrix} \tag{59}$$

is the non-convex matrix function that makes E_l quasiconvex. Here, a_j, $j = 1, 2, 3$, are scalar parameters, and the maximization in Eq. 57 is to be carried out with respect to these. The parameters a_1, a_2 and a_3 must satisfy the constraints of positive definiteness of the matrices $\mathbf{C}_i \;-\; \mathbf{\Phi}(a_1, a_2, a_3)$, $i = 1, 2$, which ensure positive definiteness of the compliance matrix β, Eq. 58, and hence positive complementary energy for all σ in Eq. 57.

In the case of topology optimization, where the compliant material is considered as void, the expressions for the effective properties of the microstructure simplify considerably so that the optimum properties can be expressed in explicit analytic form (Gibiansky and Cherkaev [31]). Thus, letting $\mathbf{C}_2 \to \infty$ to mimic void, the term $(\mathbf{C}_2 - \mathbf{\Phi})^{-1}$ of $\beta(\mathbf{C}_1, \mathbf{C}_2, \mathbf{\Phi}(a_j))$ in Eq. (58) vanishes, and the following expression for E_l is obtained

$$E_l \;=\; \max_{a_j} \;\sigma^T \left[\;\left[1 \;+\; \frac{m_2}{m_1}\right]\;\mathbf{C}_1 \;-\; \frac{m_2}{m_1}\;\mathbf{\Phi}\right]\sigma \tag{60}$$

$$=\; \sigma^T\;\mathbf{C}_1\;\sigma \;+\; \frac{m_2}{m_1}\;\max_{a_j}\;\sigma^T\;\mathbf{G}\;\sigma$$

where $\mathbf{G} = (\mathbf{C}_1 - \mathbf{\Phi})$. Assuming that $\mathbf{n}_i$ are oriented along the principal directions of σ, it is only the upper 3×3 blocks of $\mathbf{C}_1$ and $\mathbf{G}$ that have an influence upon the resulting bound. The lower bound E_l for the complementary energy density can then be formulated as

$$E_l \;=\; \hat{\sigma}^T\;\hat{\mathbf{C}}_1\;\hat{\sigma} \;+\; \frac{m_2}{m_1}\;\max_{a_j}\;\hat{\sigma}^T\;\hat{\mathbf{G}}\;\hat{\sigma} \tag{61}$$

where $\hat{\sigma}$ indicates principal stresses, and $\hat{\mathbf{G}}$, $\hat{\mathbf{C}}_1$ indicate the upper-left 3 $\times$ 3 blocks of the 6 $\times$ 6 matrices $\mathbf{G}$ and $\mathbf{C}_1$ respectively. If a_j is chosen optimally we can write

$$E_l = \hat{\sigma}^T \left[\hat{\mathbf{C}}_1 + \frac{m_2}{m_1} \hat{\mathbf{G}} \right] \hat{\sigma} \tag{62}$$

Assuming that σ is oriented along the principal directions such that $\sigma^T = (\sigma_1\, \sigma_2\, \sigma_3\, 0\, 0\, 0)$, where σ_i, $i = 1, 2, 3$, denote the principal stresses, and by defining principal stress ratios ω and η as

$$\omega = \frac{\sigma_1}{\sigma_3} \;, \quad \eta = \frac{\sigma_2}{\sigma_3} \qquad \text{where} \qquad \sigma_3 \geq |\,\sigma_2\,| \quad \text{and} \; \sigma_3 \geq |\,\sigma_1\,| \tag{63}$$

the optimum values of the parameters a_j can be derived in explicit analytical form by assuming that the optimum solution lies on the edges or vertices of the design space. The expressions for the optimum parameters a_j change with changing stresses, and it turns out that the entire stress domain is covered by nine different sets of expressions for the optimum parameters a_j. These expressions can be used to determine the optimum α_i associated with the upper bound (Gibiansky and Cherkaev [31]). Considering the upper

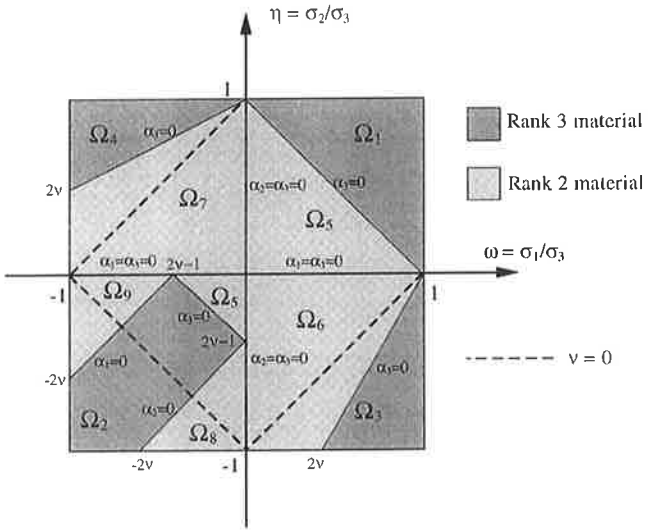


Figure 10: The domain Ω with individual sub-domains Ω_i.

bound E_u in Eq. 54 with $\mathbf{Q}$ given in Eq. 53, and letting $\mathbf{C}_2 \to \infty$, the first term of $\mathbf{Q}$ equals zero. Moreover, assuming that the rank three microstructure is oriented along the principal stress directions, the second term in the expression for $\mathbf{Q}$ simplifies, and the following expression for E_u is obtained

$$E_u = \hat{\sigma}^T \hat{\mathbf{C}}_1 \hat{\sigma} + \min_{\alpha_i} \hat{\sigma}^T \hat{\mathbf{Q}}^{-1} \hat{\sigma} \tag{64}$$

Provided that α_i, $i = 1,2,3$, are chosen optimally we can write

$$E_u = \hat{\sigma}^T \left[\hat{\mathbf{C}}_1 + \hat{\mathbf{Q}}^{-1} \right] \hat{\sigma} = \hat{\sigma}^T \hat{\mathbf{C}}_M \hat{\sigma} \tag{65}$$

where $\hat{\mathbf{C}}_M$ is the optimum effective compliance matrix. In Eqs. 64, 65, the vector $\hat{\sigma}$ and the matrices $\hat{\mathbf{C}}_1$, $\hat{\mathbf{Q}}$ and $\hat{\mathbf{C}}_M$ are all expressed in the basis $\mathbf{n}_i$ oriented along the principal directions of σ. It is shown in Gibiansky and Cherkaev [31] that the two bounds E_u and E_l coalesce for an optimum microstructure which implies that for any admissible a_j the following difference is equal to zero

$$E_u(\sigma) - E_l(\sigma) = \max_{a_j} \min_{\alpha_i} \sigma^T \left(\mathbf{Q}^{-1}(\alpha_i) - \mathbf{G}(a_j) \right) \sigma = 0 \tag{66}$$

which after some algebra can be reduced to

$$\left(\hat{\mathbf{Q}}(\alpha_i)\hat{\mathbf{G}}_i - \mathbf{I} \right) \hat{\sigma} = \mathbf{0} \tag{67}$$

where $\mathbf{I}$ is the unit matrix. In order to solve this problem, we first determine the eigenvalues of the 3×3 matrix $\hat{\mathbf{Q}}(\alpha_i)\,\hat{\mathbf{G}}_i$. It can be shown that for all matrices $\hat{\mathbf{Q}}(\alpha_i)\,\hat{\mathbf{G}}_i$ the corresponding eigenvalues are equal to $(0,0,1)$ and the corresponding eigenvector yields a set of equations which upon solution yields the optimum α_i that satisfy Eq. 67.

The solution of this problem reveals that the relation between the principal stresses and the optimum microstructure changes for different stresses and that the entire stress domain is divided into nine sub-domains (Fig. 6) with different relations. The expressions for the individual domains are listed in Eq. 68 where ν denotes Poisson's ratio.

$$\alpha_1 = \frac{1 - \omega + \eta}{1 + \omega + \eta}, \quad \alpha_2 = \frac{1 + \omega - \eta}{1 + \omega + \eta}, \quad \alpha_3 = \frac{\omega + \eta - 1}{1 + \omega + \eta}, \quad \text{if } \sigma \in \Omega_1$$

$$\alpha_1 = \frac{-\omega + \eta - (1 - 2\nu)}{(1 - 2\nu)(\omega + \eta - (1 + 2\nu))}, \quad \alpha_2 = \frac{\omega - \eta - (1 - 2\nu)}{(1 - 2\nu)(\omega + \eta - (1 + 2\nu))},$$
$$\alpha_3 = \frac{\omega + \eta + (1 - 2\nu)}{\omega + \eta - (1 + 2\nu)}, \quad \text{if } \sigma \in \Omega_2$$

$$\alpha_1 = \frac{1 - \omega - (1 - 2\nu)\,\eta}{(1 - 2\nu)\,(1 + \omega - (1 + 2\nu)\,\eta)}, \quad \alpha_2 = \frac{1 + \omega + (1 - 2\nu)\,\eta}{1 + \omega - (1 + 2\nu)\eta},$$
$$\alpha_3 = \frac{\omega - (1 - 2\nu)\,\eta - 1}{(1 - 2\nu)\,(1 + \omega - (1 + 2\nu)\,\eta)}, \quad \text{if } \sigma \in \Omega_3 \tag{68}$$

$$\alpha_1 = \frac{\eta - 1 - \omega(1 - 2\nu)}{(1 - 2\nu)\,(1 - (1 + 2\nu)\omega + \eta)}, \quad \alpha_2 = \frac{1 - \omega(1 - 2\nu) - \eta}{(1 - 2\nu)\,(1 - (1 + 2\nu)\,\omega + \eta)},$$
$$\alpha_3 = \frac{1 + \omega(1 - 2\nu) + \eta}{(1 - (1 + 2\nu)\,\omega + \eta)}, \quad \text{if } \sigma \in \Omega_4$$

$$\alpha_1 = \frac{\eta}{\omega + \eta}, \quad \alpha_2 = \frac{\omega}{\omega + \eta}, \quad \alpha_3 = 0 \ , \quad \text{if } \sigma \in \Omega_5$$

$$\alpha_1 = -\frac{\eta}{\omega - \eta}, \quad \alpha_2 = \frac{\omega}{\omega - \eta}, \quad \alpha_3 = 0 \ , \quad \text{if } \sigma \in \Omega_6$$

$$\alpha_1 = -\frac{\eta}{\omega - \eta}, \quad \alpha_2 = \frac{\omega}{\omega - \eta}, \quad \alpha_3 = 0 \ , \quad \text{if } \sigma \in \Omega_7$$

$$\alpha_1 = \frac{1}{1 - \omega}, \quad \alpha_2 = 0, \quad \alpha_3 = \frac{\omega}{\omega - 1} \ , \quad \text{if } \sigma \in \Omega_8$$

$$\alpha_1 = 0, \quad \alpha_2 = \frac{1}{1 - \eta}, \quad \alpha_3 = -\frac{\eta}{1 - \eta} \ , \quad \text{if } \sigma \in \Omega_9$$

Eqs. 68 constitute the analytically derived solution to the local microstructure optimization problem. From Figure 10 it is clear that if the size of all principal stresses are in the same range, the resulting microstructure is a rank three laminate, whereas the resulting microstructure is a rank two laminate if one principal stress is considerably less than the two others.

3.3.1 Parametric Study of the Microstructure

Parametric studies of the effective properties of microstructures can be used to reveal valuable information about stability and response of the microstructures under certain conditions. Extensive parametric studies have thus been performed by a number of researchers for the ranked laminates, but these have almost exclusively been devoted to planar rank 2 laminates, see e.g. Bendsøe [2] and Sigmund [43]. The relation between the effective properties of the layered 3D microstructure and some of the used parameters has been studied by Cherkaev and Gibiansky [31], Cherkaev and Palais [32], Jacobsen *et al.* [76], and Olhoff *et al.* [77].

In this context it is important to note ([76],[77]) that the 3D laminate generally does not suffer from the problem of lack of shear stiffness which is the case for the 2D laminate, see e.g. Bendsøe [2]. Considering Eqs. (3) and (55), it is clear that the microstructure posesses a finite stiffness in all tensor directions if the microstructure is of third rank.

Due to the large number of parameters in the 3D case, it is useful to perform a numerical analysis in order to obtain an overall picture of the material properties ([76],[77]). Hence, with the results for the rigorous 3D microstructural model of the preceding section at hand, we may compare the relative energy (stiffness) of this microstructure to that of the non-penalized model used for isotropic materials (i.e., isotropic materials with $p = 1$ in Eq. (15) and Figure 6). A plot of the relative energy $E_{rel} = \frac{E_{iso}}{E_{aniso}}$ vs. the volume density ρ of material is shown in Fig. 11.

The energy density of the anisotropic material is not uniquely defined because it varies with varying stresses, but if we vary the stress field within the admissible limits for a fixed

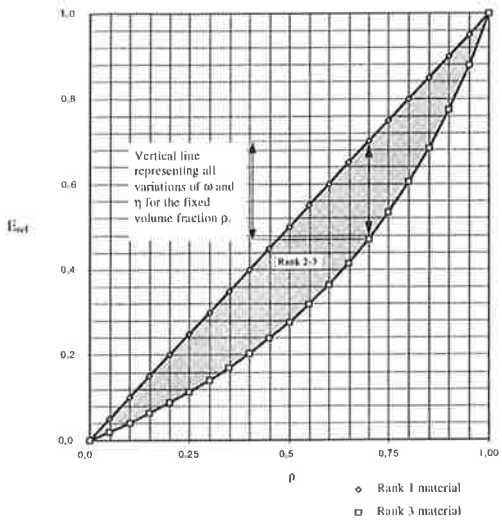


Figure 11: Relative energy $E_{rel} = \frac{E_{iso}}{E_{aniso}}$ vs. the volume density ρ of material for the entire range of principal stress ratios ω and η.

volume density ρ of material, we get the results shown in Figure 11. The upper limit of the grey region corresponds to a rank 1 material while the lower limit corresponds to a rank 3 material. An essential observation is that in all points except of $\rho = 0$ and $\rho = 1$, the energy density of the ranked optimum anisotropic material is always higher than (in the special case of a uniaxial stress, equal to) that of the isotropic material model. This indicates that the commonly used assumption of a linear material density to stiffness relation overestimates the stiffness of a given structure if intermediate densities occur. This is in perfect agreement with results obtained by Sigmund [43] for plane laminates.

3.4 Discussion and Examples

3.4.1 Mesh Independency of Solutions

We now illustrate numerically by way of an example that the layered 2D and 3D microstructures considered in Sections 3.2 and 3.3 provide full relaxation of problems of maximum stiffness topology design for a single case of loading. As discussed in Section 2.2, topology and layout optimization often are not well-posed because the design space is not closed in the appropriate sense. A remedy to ensure closure of the set of feasible designs is then to relax, i.e. regularize, the mathematical formulation of the problem by introducing composites with perforated, periodic microstructures as admissible materials for the structural design [2], [17]–[24]. Numerical indications of the need for regularization of a given problem are lack of convergence and dependence of the designs on the size of

the applied finite element mesh. In particular, if the problem is not properly regularized it is not possible to obtain a limiting, numerically stable design by consecutively decreasing the mesh size, see Cheng and Olhoff [20, 23], Cheng [24] and Olhoff *et al.* [19].

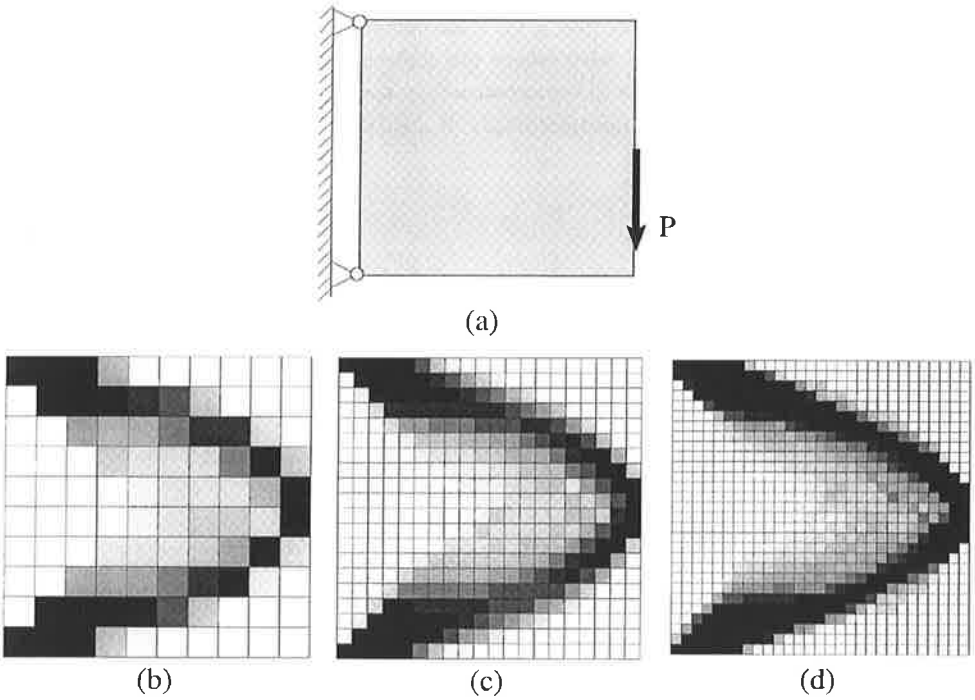


Figure 12: Example of finite element mesh refinement [77]. (a) Admissible design domain, loading and support conditions for example problem. (b, c, d) Topology results obtained for different mesh sizes. It is seen that the topology is consistent.

To illustrate the mesh independency of topologies obtained by using the 3D microstructures of Section 3.3, we consider for convenience a planar design domain with loading and support conditions as depicted in Figure 12a. The problem is modeled by usage of 8-node 3D isoparametric finite elements and solved using three different mesh sizes, whereby solutions as shown in Figs. 12 b, c and d are obtained, see [77]. The results clearly indicate mesh independency of the topology and the existence of a well defined limiting design for any further degree of mesh refinement. This witnesses that the applied microstructures have provided full relaxation of the problem. It is worth noting that the solutions in Figure 12 consist of composite in large sub-domains, and that the designs in Figs. 12 c, d, strongly resemble that of a Michell truss [118]. This is not surprising since the Michell truss is the optimum solution to the problem depicted in Figure 11a provided that the amount of available material is much less than the amount prescribed in the current example.

3.4.2 2D Topology Examples

Consider now a simple maximum stiffness topology design problem for a 2D structure (see [98]). The structure is to be confined within a rectangular design domain and to be subjected to a concentrated load and be supported as shown in Figure 13. The volume fraction of available material relative to the admissible design domain is taken to be 0.45. The design domain is discretized into a 60 x 20 mesh of 8-node 2D isoparametric finite elements, and we consider two cases where the material within each of the elements is modeled by either a second-rank 2D microstructure with orthogonal layers as considered in Section 3.2.4 or by a layered microstructure of arbitrary rank governed by the moment formulation in Section 3.2.5.

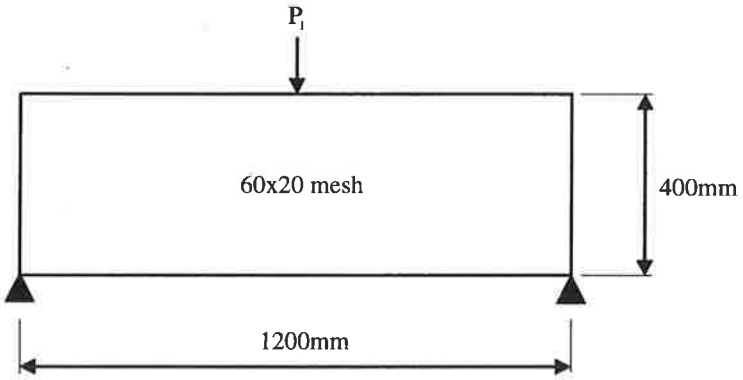


Figure 13: Design domain, load and support conditions for two 2D example problems.

The two optimization problems have been solved using the two iterative approaches outlined in Sections 3.2.4 and 3.2.5, respectively, and Figs. 14 a, b show the optimum topologies obtained. These two optimum solutions are found to have the same stiffness (total elastic energy) and are also seen to be almost identical. This is to be expected as the optimum microstructure for the solution of the problem considered here is a second-rank microstructure (Section 3.2.4), and this microstructure is contained as a special case of the layered microstructure of arbitrary rank covered by the moment formulation in Section 3.2.5.

3.4.3 3D Topology Examples

We now consider two full 3D topology optimization examples [77], where the solution is based on usage of the optimum 3D material microstructure discussed in Section 3.3.

The first example is presented in Figure 15. Here, Figure 15 a shows a cubic design domain which is subjected to four parallel concentrated loads P at the upper surface and equipped with simple supports at the four corners of the lower surface. The volume

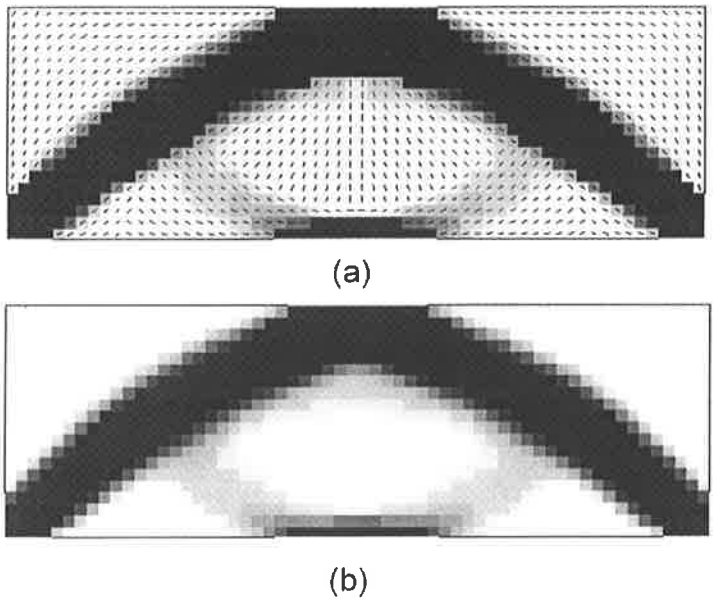


Figure 14: Optimum solutions to example problems [98]. (a) Solution by usage of second-rank 2D microstructure with orthogonal layers. The directions of the principal material stiffnesses are also shown. (b) Solution based on layered microstructure of arbitrary rank governed by the moment formulation.

fraction of available material relative to the design domain is chosen to be 0.08 in this example.

Figure 15 b illustrates the optimum topology solution to the problem which is a quadropod consisting of four legs of quite distinctly solid material, each transferring one of the applied concentrated loads to the nearest simple support, and a substructure made of composite material which interconnects the upper parts of the four legs.

Figure 16 a displays an example [77] of an oblong box-formed design domain for which the available volume fraction of material is taken to be 0.30. The two ends of the domain are clamped, and a concentrated bending moment M acts at the center. Figure 16 b shows the topology solution obtained after usage of a commonly used method of penalization [1, 16] of intermediate densities $0 < \rho < 1$ of material that makes a 0-1 solution more advantageous, but the solution is seen to consist still of a large part of composite material. In Figure 16 c, elements with material densities below 0.5 have been removed from the solution as a means to visualise the topology.

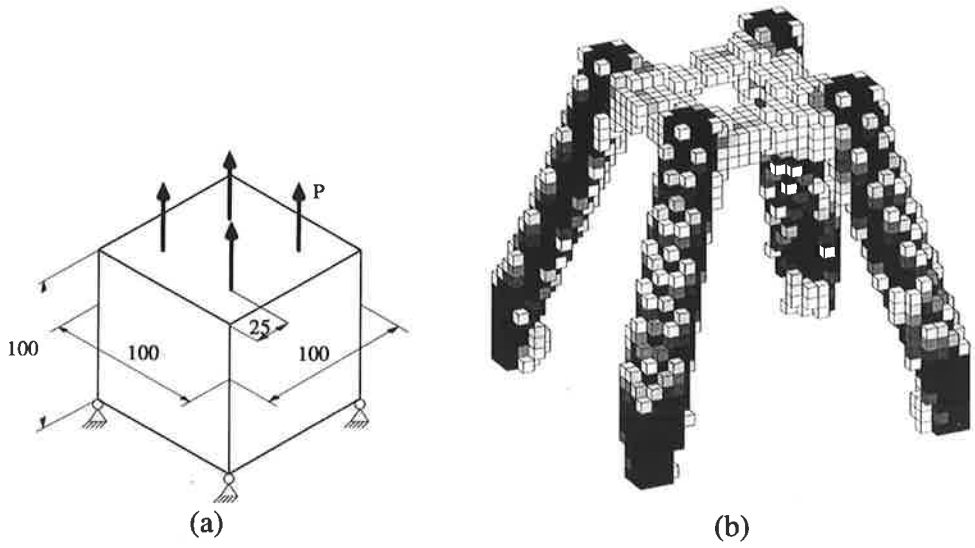


Figure 15: Example problem [77] with cubic design domain subjected to four point loads P. Volume fraction of available material is 0.08. (a) Design domain with loading and support conditions. (b) Quadropod solution to the problem with material densities less than 0.8 removed.

3.4.4 Discussion

As is clearly illustrated by the results in Figs. 12 and 16, it is a characteristic feature that the structures of optimum topology obtained by the approach of relaxation generally consist of composite material in large sub-domains. There are two reasons for this. Firstly, as discussed in Section 3.3.1, depending on the state of stress, composite materials generally are much more efficient that isotropic, solid materials, and therefore largely manifest themselves in optimum solutions. Secondly, it is characteristic that application of the optimum 2D or 3D microstructures considered in Sections 3.2 and 3.3 as a basis for topology optimization actually prompts a larger content of composite material in the resulting solution, as compared with use of non-optimum microstructures.

Hence, optimum microstructures must always be chosen as a basis for topology optimization when it is our desire to obtain or acquire knowledge of the global optimum solution associated with the highest possible value of the performance index. The not surprisingly large structural sub-domains of composite material that are a consequence of such a desire, however, often make it difficult to visualize the overall structural topology of such a solution and to devise simplified, sub-optimum (0-1) designs or designs that are attractive from the point of view of manufacture.

In the subsequent Section 4, we shall consider how the introduction of appropriate re-

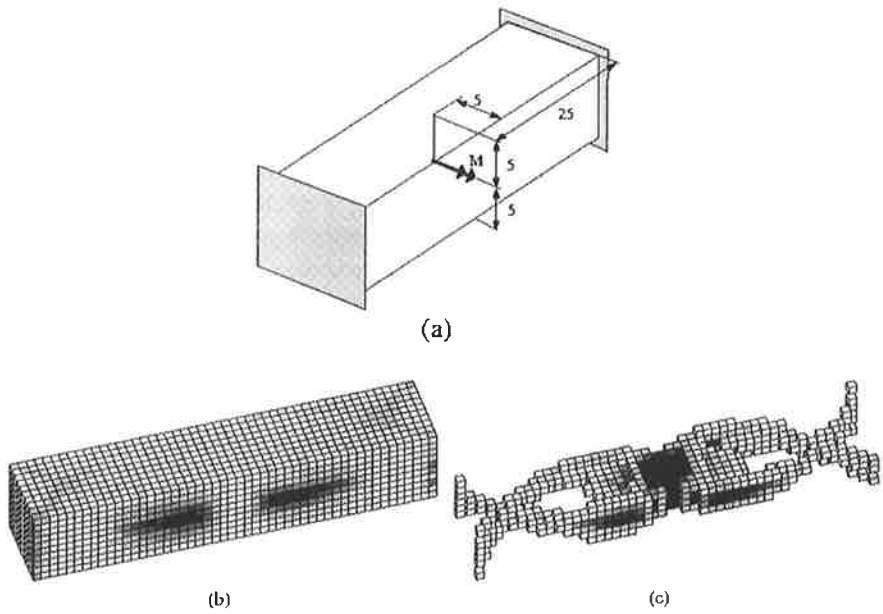


Figure 16: Oblong box-shaped design domain with clamped ends and a concentrated bending moment M acting in the center [77]. The material volume fraction is 0.30. (a) Design domain with loading and support conditions. (b) Topology solution represented with all material densities present. (c) Solution depicted with densities below 0.5 removed.

strictions in the formulation of topology optimization problems may lead to more distinct (0-1) topology designs that are easier to manufacture, but also associated with a lower value of the performance index.

4 The Perimeter Method and Other Approaches to Achieve Well-Posed Problem Formulations

As discussed in the end of Section 2.2, the distributed, discrete valued black-white (0-1) topology optimization problem stated by Eqs. (6) [or (8)], (7) and (10) of determining the stiffest (minimum compliance) structure under a single loading condition by distributing a fixed volume V_{fixed} of homogeneous, isotropic, linearly elastic material within an admissible design domain Ω, is ill-posed. In this formulation Ω_s is the solid part of Ω where material is present, and the design variable χ is a binary indicator function with $\chi = 1$ denoting solid and $\chi = 0$ denoting void. Attempts to solve this problem do not converge to macroscopic patterns of solid and void. Instead, rapid spatial oscillations appear in the

function χ as the compliance is reduced. Although a limiting value of the compliance can be determined, the solution tends toward a design with an infinite number of vanishingly small holes, rather than a finite number of macroscopic holes.

One way of achieving a well-posed problem is to extend the design space (via relaxation) to include materials with perforated microstructure of variable density covering the entire spectrum from solid to void, and to derive their effective mechanical properties by homogenization as considered in Section 3. Numerical methods based on a complete relaxation (i.e. incorporating optimum microstructures as considered in Sections 2.3.2, 2.3.3 and Section 3) are convergent with respect to mesh refinement, but they yield solutions with large sub-domains of perforated material which are undesirable from the point of view of manufacturing, see Section 3.4 and, e.g., Jog *et al.* [119]. Numerical methods based on partial relaxations (c.f. the "Hole-in-Cell" microstructures [15] considered in Section 2.3.1) generate implicit penalties against intermediate densities. They behave similarly to methods with explicit penalties.

This chapter will focus on another way to achieve a well-posed mathematical formulation for topology optimization, namely to restrict the design space such as to remove the possibility of rapid oscillations in the density of material in the structure. Approaches of this kind have been studied in Haber *et al.* [35]–[37], Beckers [71, 72], Duysinx [120] and Beckers and Fleury [73] where a constraint is imposed on the perimeter of the structure (see the following Section 4.1), in Petersson and Sigmund [40] where a pointwise constraint is imposed on the gradient of the density of material (see Section 4.2), and in Sigmund [63, 43], Sigmund and Torquato [86], and Sigmund *et al.* [87], where an image processing inspired filter is used to maintain a certain minimum length scale in the structure (Section 4.3).

4.1 Perimeter Method

Haber *et al.* introduced in [36] (see also [35, 37]) a new method that achieves a well-posed continuum optimization problem by restricting the design space to exclude rapidly oscillating designs. These authors defined the perimeter of a design as the measure of the boundary of the solid region: $|\partial\Omega_s|$. Designs with fewer, larger holes have lower perimeter measures than designs of equal volume with more numerous smaller holes [35, 36]. Chattering designs are characterized by unbounded perimeter measures, so an upper-bound constraint on the perimeter effectively excludes chattering solutions from the feasible design space. Furthermore, a perimeter bound controls the number and sizes of holes in a macroscopic design without otherwise restricting the shape or layout of the holes. Thus, Haber *et al.* [35]–[37] appended an upper bound constraint on the perimeter $|\partial\Omega_s| \leq \bar{P}$ (where $\bar{P}$ is a designer-specified value) to the abovementioned problem stated by Eqs. (6) [or (8)], (7) and (10). An existence proof for a topology design problem similar to this perimeter constrained problem is presented by Ambrosio & Buttazzo [38].

Finite element discretizations of the discrete valued, perimeter constrained topology optimization problem generate integer programming problems that were until very recently

considered too large for direct solution. However, Beckers [71] recently overcame this difficulty for problems with absence of local constraints by developing a very efficient discrete mathematical programming method working in the dual space, and she has subsequently published several solutions to these large-scale, discrete valued problems [71, 72, 73] with global design objectives and constraints.

Haber *et al.* [35]–[37] used a continuos approximation to the problem by replacing the indicator function χ, see Eq. 10, with a distributed interpolation parameter ρ , $0 \leq \rho_{min} \leq \rho \leq 1$, for the volume density of material, and introduced a penalty on intermediate values of ρ to suppress transitional material in the final design. The small lower bound ρ_{min} on ρ ensures that the energy is positive definite, and the volume of structural material is simply given by $\int_\Omega \rho d\Omega$.

As a representation of the perimeter measure $|\partial\Omega_s|$ that is compatible with the continuous interpolation model, Haber *et al.* [35]–[37] adopted the *total variation* $TV(\rho)$ of ρ, as this is a suitable measure which approaches the perimeter, in the limit, as the amount of transitional material is forced to zero [121]. In anticipation of a piecewise-continuous finite element model for ρ , Haber *et al.* partitioned the admissible design domain Ω into a set of open, disjoint regions Ω_α so that $\Omega = \cup_\alpha \Omega_\alpha$, and accounted for the possibility that ρ may be discontinuous across sets of measure zero (e.g. the finite element boundaries). The total variation of the scalar function ρ can then be expressed [122] as

$$TV(\rho) \;=\; \int_{\Omega\setminus\Gamma_J} |\nabla\rho|\; d\Omega \;+\; \int_{\Gamma_J} |<\rho>|\; d\Gamma \tag{69}$$

where $\Gamma_J = \Omega\setminus \cup_\alpha \Omega_\alpha$ is the jump set of ρ , and $<\rho>$ is the jump in ρ across Γ_J .

The total variation in Eq. (69) is used to express the 'perimeter' P of the material volume density function ρ via the formula

$$P \;\equiv\; \int_{\Omega\setminus\Gamma_J} g_h(\nabla\rho,\xi)\; d\Omega \;+\; \int_{\Gamma_J} j(<\rho>,\xi)\; d\Gamma \tag{70}$$

where the functions g_h and j are defined by

$$g_h(\mathbf{w},\xi) \;\equiv\; \left[(1+2\xi)\mathbf{w}^T\mathbf{w} + \frac{\xi^2}{h^2}\right]^{1/2} \;-\; \frac{\xi}{h} \;,\quad j(r,\xi) \;\equiv\; \left[(1+2\xi)r^2+\xi^2\right]^{1/2} \;-\; \xi \tag{71}$$

and are smooth approximations to $|\mathbf{w}|$ and $|r|$, respectively. In Eq. (71), h is a characteristic mesh dimension (e.g., the size of a finite element). The smoothing, based on the parameter ξ in Eq. (71), circumvents numerical problems associated with the non-differentiability of the magnitude and absolute value operators appearing in the expression for the total variation in Eq. (69). Note that $\lim_{\xi\to 0}\; g_h(\nabla\rho,\xi) = |\nabla\rho|$ and $\lim_{\xi\to 0}\; j(<\rho>,\xi) = |<\rho>|$. Further, the approximations are exact in the limit of a discrete approximation to an integer design, even for $\xi > 0$. That is, for all $\xi \geq 0$, $g_h(\nabla\rho,\xi) \to |\nabla\rho|$ as $|\nabla\rho| \to 0$ or $\frac{1}{h}$, and $j(<\rho>,\xi) \to |<\rho>|$ as $|<\rho>| \to 0$ or 1. Haber *et al.* [35]–[37] use an optimally

criterion method to solve the design problem. The constraint $|\partial\Omega_s| \leq \bar{P}$ is treated with an interior penalty method, and the prescribed volume constraint is enforced by means of a Lagrangian multiplier. The modified problem considered has the form

$$\sup_{\rho} \inf_{\mathbf{u}} \Pi + \alpha S_1 + \beta S_2 - \lambda \left[\int_\Omega \rho d\Omega - V_{fixed}\right] \tag{72}$$

in which Π is the total potential energy, $\mathbf{u}$ the displacement field, α and β are positive scalars, S_1 is a function that penalizes intermediate values of ρ, S_2 is an interior penalty function for the constraint $|\partial\Omega_s| \leq \bar{P}$, and λ is the Lagrange multiplier associated with the volume constraint.

Fig. 17 illustrates the use of the perimeter method to control the optimum topology and to achieve mesh independent solutions [35]. The results are obtained for the MBB beam example first considered in [6]: a simply supported beam with a midspan load on the top surface. Solid material is prescribed around the periphery, and $V_{fixed} = 0.60|\Omega|$. The results in Figs. 17a and b are based on a perimeter bound $\bar{P} = 30L$. While the coarse mesh in Fig. 17a is not able to completely resolve the design, the optimum topology is essentially the same as the one obtained with the refined mesh in Fig. 17b. The characteristic Michell truss layout is evident. A simpler optimum topology is obtained by reducing the perimeter bound to $\bar{P} = 24L$. The same optimum topology is clearly evident in Figs. 17c and 17d, demonstrating the ability of the perimeter method to generate mesh independent solutions. An even simpler design is obtained by further reducing the perimeter bound to $\bar{P} = 22L$ in Figs. 17e and f. Fig. 17g shows a very simple topology obtained by setting $\bar{P} = 18L$.

The compliance values obtained on different meshes cannot be directly compared, because the discretization error associated with a coarser mesh artificially reduces the compliance. However, the trade-off between stiffness (low compliance) and perimeter is clearly evident when designs obtained on the same mesh are compared (Figs. 17a, c, e and g for the coarse mesh and Figs. 17b, d and f for the refined mesh). Improved stiffness can be achieved at the cost of design complexity; the compliance increases as the perimeter bound is decreased to achieve a simpler design. On the other hand, the results show that designs with simple, practical topologies can be achieved with a relatively small increase in compliance.

It may be concluded that the perimeter method yields high-quality numerical solutions and has the following distinct advantages:

Firstly, solutions are convergent with respect to mesh refinement; finer meshes offer improved resolution without altering the design. Next, the SIMP material model (see Section 2.3.4) or any other reasonable interpolation can be used to model material properties for intermediate values of ρ. The complexities and limitations of the homogenization method are circumvented; so the perimeter method is easy to implement and is extensible to general constraint and objective functions. Finally, the perimeter method allows the designer to control the number and scale of the holes, and the designer can vary $\bar{P}$ to explore alternative designs.

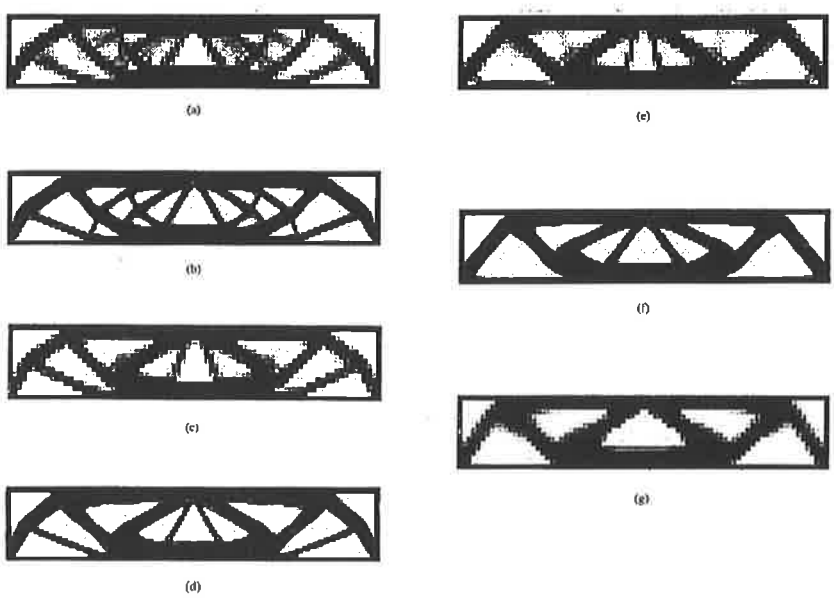


Figure 17: Solutions for the MBB beam problem [6] with perimeter control (from [35]): (a) $\bar{P} = 30$L, coarse mesh; (b) $\bar{P} = 30$L, fine mesh; (c) $\bar{P} = 24$L, coarse mesh; (d) $\bar{P} = 24$L, fine mesh; (e) $\bar{P} = 22$L, coarse mesh; (f) $\bar{P} = 22$L, fine mesh; (g) $\bar{P} = 18$L, coarse mesh.

4.2 Local Constraint on Gradient of Material Density

Introduction of a local gradient constraint on thickness variation of plates was first done by Niordson [123] in 1983. The following approach with a local constraint on the gradient of the volume density ρ of material in a two-dimensional topology optimization problem, see Petersson and Sigmund [40], presupposes that ρ is sufficiently smooth and defined for intermediate values. This is the case when, e.g., the SIMP material model in Section 2.3.4 is applied to the ill-posed topology optimization problem stated by Eqs. (6) [or (8)], (7) and (10), and for this type of problem proof of existence, mesh independency and numerical implementation of a scheme introducing local gradient constraint on density variation was given by Petersson and Sigmund [40].

The constraint on the local density variation is written as the following pointwise constraint on the derivatives of the function ρ:

$$\left| \frac{\partial \rho}{\partial x_i} \right| \leq \mathbf{c} \ , \quad i = 1, 2 \tag{73}$$

The convergence proof implies that checkerboards (see Sub-section 6.1.3) and other numerical anomalies will be eliminated, or at least, they can be made arbitrarily weak using this

scheme. However, implementation of the scheme results in up to 2N additional constraints in the topology optimization problem, and the method must therefore be considered to be too slow for practical design problems [40].

4.3 Filtering Techniques

Based on filtering techniques from image processing, Sigmund [43] in 1994 suggested a filter that prevented creation of checkerboards in numerical solutions to topology optimization problems (see Sub-section 6.1.3) by modifying the design sensitivities used in each iteration of the algorithm solving the discretized problem. This filter makes the design sensitivity of a given element depend on a weighted avarage over the element itself and its eight neighbours.

As an extension of the checkerboard filter just mentioned, Sigmund [63, 43] developed a method of mesh independent filtering which modifies the design sensitivity of any specific element based on a weighted average of the element sensitivities in a fixed neighbourhood. It should be mentioned that this filter is purely heuristic, but it produces results very similar to local gradient constrained results [40] (see the preceding section), requires only very little extra cpu-time, and is very simple to implement compared to the other approaches. Similar ideas of weighted averages have been used to ensure mesh independency in bone mechanics simulation see, e.g., Mullender *et al.* [124].

The mesh independence scheme works by modifying the element sensitivities as follows

$$\widehat{\frac{\partial f}{\partial \rho_k}} = (\rho_k)^{-1} \frac{1}{\sum_{i=1}^{N} \hat{H}_i} \sum_{i=1}^{N} \hat{H}_i \rho_i \frac{\partial f}{\partial \rho_i} \tag{74}$$

Here the convolution operator (weight factor) $\hat{H}_i$ is written as

$$\hat{H}_i = r_{min} - dist(k,i) \quad , \quad \{i \in N \mid dist(k,i) \leq r_{min}\} \quad , \quad k = 1, \ldots, N \tag{75}$$

where the operator $dist(k,i)$ is defined as the distance between the center of element k and the center of element i. The convolution operator $\hat{H}_i$ is zero outside the filter area. The convolution operator for element i is seen to decay linearly with the distance from element k.

This means that instead of applying the real sensitivities (i.e. $\frac{\partial f}{\partial \rho_k}$), the filtered sensitivities computed by Eq. (74) are used. It is worthwhile noting that the filtered sensitivity in Eq. (74) converges to the original sensitivity when r_{min} approaches zero, and that all sensitivities will be equal (resulting in an even distribution of material) when r_{min} approaches infinity.

The existence issue for the mesh independence filter has yet to be proved, but applications in several papers on material design and mechanism design (see e.g. Sigmund [62]–[65], [86, 87]) show that the method in practice produces mesh independent designs.

4.4 Discussion of Methods

As discussed in Sigmund and Petersson [125], the perimeter, local gradient and mesh independence filter methods produce very similar designs. However, there are some notable differences between the methods which will be briefly discussed below.

The perimeter method [35]–[37], [71]–[73] in Section 4.1 involve a global constraint and will allow the formation of locally very thin structural members. The local gradient [40, 125] and filtering [62]–[65],[43, 86, 87] approaches (Section 4.2 and 4.4) involves local constraints and will generally remove such thin structural members.

Predicting the value $\bar{P}$ of the perimeter constraint for a new design problem must be determined by experiments, since there is no direct relation between local scale in the structure and the perimeter bound. If the perimeter bound is too tight, there may be no solution to the optimization problem. This problem is particularly difficult for three-dimensional problems. In contrast, the gradient and filtering schemes define a local length scale under which structural variation is filtered out. This local length scale corresponds to a lower limit on widths of structural members and can be easily defined when, e.g., manufacturing is taken into consideration.

Another important difference is the implementation aspect. The perimeter method implies that an additional constraint is included in the formulation of the optimization problem. Although the addition of one extra constraint to the optimization problem should not be a problem for advanced large-scale mathematical programming algorithms, practice has shown that implementation of the constraint can give rise to some convergence problems [120]. The local gradient constraint approach is considered impractical due to the addition of 2N extra constraints to the optimization problem [125]. The great advantage of the filtering technique is that it requires no extra constraints in the optimization problem. Furthermore, it is very easy to implement, and experience shows that its implementation even stabilizes convergence [62]–[65],[86, 87]. The only disadvantage of the filtering method is that it is based on heuristics.

The perimeter and mesh independence filter methods have both been successfully applied to 3D problems (see Beckers [71, 72] and Sigmund *et al.* [87], respectively), and the comparisons above also hold true for 3D problems.

5 A Macro-Approach to Topology Optimization

The topology optimization techniques considered in the foregoing all use a fixed finite element mesh, and as they are all parameterized based on the application of a microstructured material model[3], see Section 2.3, they may be referred to as 'micro-approaches'. However, there exists a conceptually very different approach to topology design, a so-called 'macro-approach' which will be discussed in this Chapter.

[3]Note that in the light of recent work by Bendsøe & Sigmund [89], also the SIMP model can be characterized as a microstructured composite (see Section 2.3.4).

In contrast to the 'micro-approaches', the 'macro-approach' is based on using during the entire design procedure the constitutive law for the solid material of the structure to be obtained. Moreover, instead of applying microstructural parameterizations and a fixed finite element mesh, the 'macro-approach' is based on variable meshing and parameterizations governing the variation of the shapes of the exterior boundary and the boundaries of interior holes, just as in a conventional shape optimization problem. In the so-called 'Bubble Method' (Eschenauer *et al.* [44], [45], [46], [47] and Schumacher [48]), this shape design technique is augmented to one of simultaneous topology design by considering additional design variables that govern the number and positions of holes (called bubbles), and where in each step of an iterative hierarchical procedure the introduction and optimum positioning of a new hole is followed by a conventional fixed-topology shape optimization of the structure.

5.1 'Bubble Method'

As already mentioned, the 'Bubble Method' (Eschenauer *et al.* [44],[45], [46],[47], Schumacher [48]) covers a hierarchical technique of simultaneous topology and shape design. The basic idea behind the method is an iterative positioning of new holes (so-called 'bubbles') into the structure to be designed, and it is a principal task of the method to determine a problem-dependent optimum position vector for an initially small hole of given shape (e.g. circular, elliptical, triangular) to be inserted into the structure. The criteria required for this purpose are derived from a general formulation of optimality. For a structure governed by a mixed thermomechanical boundary value problem with prescribed boundary stresses, displacements, temperatures or heat flux, the following optimization problem is formulated:

> Minimize the structural volume subject to prescribed integral behavioural constraints (e.g. compliance) and local constraints in the form of the equilibrium conditions, the strain-displacement relationships, and the constitutive thermomechanical conditions.

This problem can be described completely by means of a Lagrangian function. In order to establish the first variation of the Lagrangian, the problem is divided into sub-problems concerning the variation of the stress boundaries δJ_σ, the boundaries of displacement δJ_v, and the stress-free boundaries δJ_0:

$$\delta J \;=\; \delta J_v + \delta J_\sigma + \delta J_0 = 0 \tag{76}$$

The hole can be positioned by means of the variation over the unloaded boundaries since the boundary of the hole is assumed to be traction-free,

$$T^j(\Gamma_\gamma) \;=\; (\sigma^{ij}\, n_i)_{\Gamma_\gamma} \;=\; 0 \tag{77}$$

The stationarity of the Lagrange function, J in Eq. 76, now yields a set of necessary conditions which leads to a so-called "characteristic function" ϕ, and the point in the

structure where this function attains its minimum value, constitutes the optimum position for the hole [44].

As examples, let us consider the expressions derived for the "characteristic function" ϕ for the cases of introducing (i) a circular hole in a 2D structure subjected to in-plane loading, (ii) a circular hole in a Kirchhoff plate with transverse loading, and (iii) a spherical hole in a 3D structure subjected to general loading. In all the cases the design objective is assumed to be minimization of the structural volume subject to given compliance.

(i) Introduction of small circular hole in a 2D structure subjected to in-plane mechanical loading and thermal load.

The characteristic function ϕ derived for this problem reads as follows in terms of the principal stresses σ_1, σ_2 and the temperature T:

$$\phi(\sigma_1, \sigma_2, T) = \left| \frac{1}{2E}\left[(\sigma_1+\sigma_2)^2 + 2(\sigma_1-\sigma_2)^2\right] + \alpha_T((\sigma_1+\sigma_2))(T-T_0) - \rho c T_0 \left[\frac{T}{T_0}\ln\left(\frac{T}{T_0}\right) - \frac{T}{T_0} + 1\right]\right|, \tag{78}$$

Here, T_0 is the reference temperature, α_T the thermal expansion coefficient, c the specific heat capacity, and ρ is the mass density of the structural material. The characteristic function ϕ has to be evaluated for each point of the structure and as mentioned above, the hole is positioned optimally at that point of the structure where the function attains a minimum value.

(ii) Introduction of small circular hole in a Kirchhoff plate with transverse loading.

The characteristic function for this problem has the form

$$\phi(\sigma_{1,top}, \sigma_{2,top}) = \frac{1}{6E}\left[(\sigma_{1,top}+\sigma_{2,top})^2 + \left(\frac{2(1+\nu)}{3+\nu}\right)^2 (\sigma_{1,top}-\sigma_{2,top})^2\right], \tag{79}$$

where $\sigma_{1,top}$, $\sigma_{2,top}$ denote principal stresses at the top surface of the plate.

(iii) Introduction of small spherical hole in a 3D structure subjected to general loading.

For this problem the characteristic function reads as

$$\phi(\sigma_1, \sigma_2, \sigma_3) = \frac{1}{2E(14-10\nu)}\left[C_a(\sigma_1^2+\sigma_2^2+\sigma_3^2) + C_b(\sigma_1\sigma_2+\sigma_1\sigma_3+\sigma_2\sigma_3)\right] \tag{80}$$

with the coefficients C_a and C_b given by

$$\begin{aligned} C_a &= 1920\pi(1-\nu^2) - 160\pi c(1-\nu^2) + 8\pi c^2(1-\nu) \\ C_b &= 1440\pi(1-3\nu^2) + 160\pi c(2\nu^2+\nu-1) + 8\pi c^2(1-2\nu) \end{aligned} \tag{81}$$

with $c = 3 + 15\nu$, and σ_1, σ_2 and σ_3 denoting principal stresses.

Characteristic functions ϕ for other types of constraint functions (e.g., displacements at certain points of the structure, and also local failure constraints) are available in Schumacher [48].

The optimization process of the Bubble Method consists of the following steps (see Fig. 18) within a given iteration:

Step 1 For a fixed topology, a shape optimization is carried out for the given objective and constraint functions. In [44], [45], [46], [47], [48], the shape optimization problem is parameterized in terms of positions of control points of NURBS functions (Non-Uniform Rational B-Splines) [126], and the geometry is processed by means of a self-developed computer program which supports the pre-processor of the finite element program. After the shape optimization, the structure cannot be improved further within the fixed topology class considered.

Step 2 By inserting a hole (thereby changing the topology class) it is attempted to achieve improved results. The coordinates of the optimum position for the new, initially small hole are determined by minimizing the characteristic function ϕ as described above.

Step 3 After the positioning, a fixed-topology shape optimization (as described in Step 1) is carried out in order to find the optimum (macroscopic) shape of the new bubble, the bubbles introduced earlier, and of the outer boundary of the structure.

After these steps, the next iteration is commenced by introducing a further bubble, and Steps 1-3 are repeated. A sequence of iterations is then carried out until the objective function has attained a stationary level, or until further design refinement is not required.

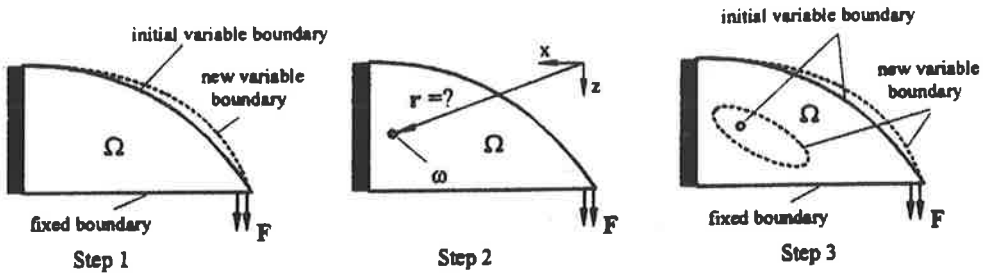


Figure 18: The three steps within an iteration of the Bubble Method (here the first iteration)

The Bubble Method has been implemented numerically using the optimization procedure SAPOP (Structural Analysis Program and Optimization Procedure) [127] which has resulted in the development of a program system that can automaticly determine the topology design but also allow for interactive communication during the optimization process. Strategies for geometry processing have been introduced which allow for successive refinement on the basis of standardized approximation functions. The sensitivities required for the shape optimization are derived by variational analysis [128].

The Bubble Method may be characterized by the following features. Firstly, the method yields pure black-white designs which is attractive from the point of view of manufacture. Secondly, due to the successive introduction of new holes in the structure, the degree of complexity of the topology is directly controlled, and less practice relevant designs can be discarded. Next, since smooth boundaries are obtained by the method, no post-smoothing is required. Finally, the smooth boundaries facilitate the inclusion of local behavioural constraints (e.g. stress and failure constraints) in the design process.

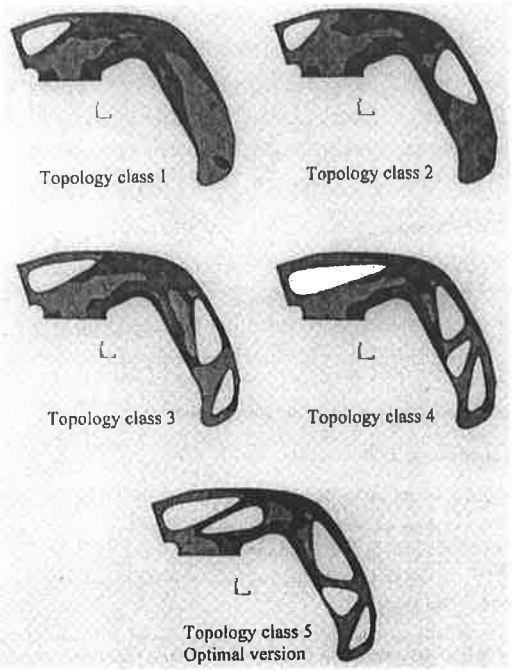


Figure 19: Step-wise hole positioning and fixed-topology shape optimizations of the grip of a handsaw. The grey scale indicates von Mises stress levels in the solid part of the grip. (From [45],[48].)

Figure 19 shows as an example an application of the Bubble Method to the design of the grip of a handsaw. The grip is rigidly fixed to the motor unit (not shown) at its upper left-hand (rear) end. This mimics the primary loading condition for the design of the grip, which is a 2m drop test where the grip is turned 90^0 clockwise and hits the floor by its rear end. In this example, the shape of the outer boundary is held fixed, but Fig. 19 clearly illustrates the stepwise positioning and shape optimization of the holes in the process of designing the grip of the handsaw. Further details concerning the example are found in [45] and [48].

6 Recent Developments and New Applications of Topology Optimization

While it has been the aim of the preceding chapters to give a picture of the current state of knowledge in the field of topology optimization of continuum structures, we in Section 6.1 of this chapter give a brief survey of literature that reflects recent developments, and in Section 6.2 some new avenues of application of topology optimization are presented.

6.1 Recent Developments

This section presents a brief survey of literature in sub-areas of topology optimization with rapid recent development, and most of the literature has appeared since the comprehensive reviews in Rozvany *et al.* [1] and Bendsøe [2] of literature published up to 1995. The following literature survey is not claimed to be exhaustive, and for broad coverage the reader is referred to the proceedings of the world congresses on structural and multidisciplinary optimization held in Goslar in 1995 [10], Zakopane in 1977 [11], Buffalo in 1999 [12], and literature cited therein.

Traditionally, most studies in topology optimization have been devoted to compliance minimization of linearly elastic 2D structures subject to a single case of in-plane loading, but much effort has been invested in extending the methodology. The reader is referred to Diaz & Bendsøe [49] and Bendsøe [2] for the extension to multiple cases of loading.

6.1.1 New Types of Structures and Materials

Plate and Shell Structures

Topology and layout design problems have been extended to plate and shell bending problems by Suzuki & Kikuchi [51], Soto & Diaz [52],[53],[55], Diaz *et al.* [54], Krog [97] and Krog & Olhoff [56], [129],[98].

3D Structures

In very recent years, topology optimization problems for 3D continuum structures have been solved by Cherkaev & Palais [32], Allaire *et al.* [30], Diaz & Lipton [75], Jacobsen *et al.* [76], Olhoff *et al.* [77], Beckers [71],[72], and Beckers & Fleury [73].

Multiple Materials

Topology design of structures with multiple materials was first undertaken by Olhoff *et al.* [50], and later by Krog [97], Krog & Olhoff [129],[98], and Burns & Cherkaev [78]. All these studies were based on usage of microstructured material models. The problem has also recently been treated by Sigmund and Torquato [86], who use an adaptation of the SIMP model (see Section 2.3.4) to interpolate between multiple material properties, an issue that has very recently been studied very thoroughly by Bendsøe & Sigmund [89].

Nonlinearly Elastic Materials

Comparatively little work has been performed in the area of nonlinear elasticity, presumably due to the difficulty to model the nonlinear behaviour of materials at intermediate

density. For work in nonlinear elasticity with non-quadratic potentials, see Jog [130]; in plasticity, see Yuge & Kikuchi [131], Maute *et al.* [68], and Swan & Arora [132]; and in damage and impact, see Bendsøe & Diaz [82] and Min *et al.* [133].

Remodeling Materials in Biomechanics

Biomechanics is a field of considerable current interest and research activity, cf. Pedersen & Bendsøe [134]. Topology design in this field has recently been addressed in Folgado & Rodrigues [135]. The use of special material models for topology design combined with bone remodelling schemes is, e.g., discussed in Tanaka *et al.* [136] and Bagge [137], see also Bendsøe [2].

6.1.2 Design Objectives and Constraints

Eigenvalues

Topology optimization with respect to eigenfrequencies of structural vibration has been studied by Diaz & Kikuchi [138], and more recently by Soto & Diaz [52],[53], Diaz *et al.* [54], Ma *et al.* [139],[140],[141], Krog [97], and Krog & Olhoff [56],[129],[98]. Buckling eigenvalue topology optimization has recently been undertaken by Folgado *et al.* [57].

Stress Constraints

Topology optimization of continuum structures with consideration of stress constraints requires that one can devise a yield criterion for material at intermediate density. It is another complexity of the problem that a computational problem appears in the form of the so-called singularity phenomenon of topology design which is well understood for truss topology problems, see Cheng & Jiang [58], Cheng [142], Cheng & Guo [59], and references cited therein. For continuum structures, stress constraints are treated in Yang & Chen [88], Duysinx & Bendsøe [60], and Duysinx & Sigmund [61].

Damage and Impact

Recent work on topology design with respect to damage and impact has been recently reported in Bendsøe & Diaz [82] and Min *et al.* [133].

Controlled Structures

Constraints related to the simultaneous optimum design of structure and controls for a controlled structure has been recently studied by Ou & Kikuchi [143],[144].

Structures with Design Dependent Loading

Topology optimization with respect to objectives pertaining to, e.g., wind and snow loading, hydrostatic pressure or fluid flow, i.e. loading that changes with the structural design, has very recently been considered in Hammer & Olhoff [145].

Multiobjective Formulations

Topology optimization problems with multiple objectives have been considered in, a.o., Diaz & Bendsøe [49], Diaz *et al.* [54] and Soto [96] for compliance problems and in Soto & Diaz [52],[53], Diaz el.al. [54], Ma *et al.* [139],[140],[141] for vibration problems, by treating the multiobjective criterion in the form of a weighted scalar function of the objectives. In

Krog [97], and Krog & Olhoff [56],[129],[98] and Folgado *et al.* [57], the multiple objectives of such problems are scalarized by using a weighted min-max formulation which is more relevant and of direct practical significance as it is a worst-case type of design objective.

6.1.3 Computational Issues

Optimization Algorithms

Much work in topology design of continuum structures has relied on the applicability of optimality criteria methods for the mostly global constraints which have been treated. To gain further versatility of the topology design methods there is now a tendency to move towards the use of more standardized mathematical programming methods. See, for example, Duysinx [120] for use of CONLIN, and Sigmund & Torquato [86], Krog & Olhoff [56],[129],[98] for use of SLP methods. See also Bendsøe [2].

Checkerboard Control

In topology optimization of continuum structures one often sees that the numerical results are polluted' by so-called checkerboard patterns. This has been identified as a numerical problem related to the choice of interpolation spaces and physically speaking certain interpolation spaces make checkerboard patterns seem artificially stiff (see Diaz & Sigmund [146], Jog & Haber [147], Petersson [148]). Various techniques have been proposed to circumvent this problem (see Bendsøe [2] and the references just cited). These techniques usually imply an increase in computer time, as the basic problem is that displacements are required to be approximated by higher order elements. The perimeter constraint (see Section 4.1 and Duysinx [120]) and filter techniques (see Section 4.3 and Sigmund and Torquato [86]) are reported to be very efficient remedies for suppressing checkerboards.

Adaptive Methods

As mentioned earlier, topology design of continuum structures has traditionally been implemented with reference to a fixed finite element mesh. However, the use of adaptive finite element methods does allow for an increased resolution of the structure. The idea is to start with a rough mesh and only refine where there is borders between material and void to be identified. Such an approach is described in Maute & Ramm [66],[67] and Maute *et al.* [68].

6.2 New Applications of Topology Optimization

This section presents some particularly interesting and rapidly developing new areas of application of methods of optimum topology design, namely in Sub-section 6.2.1 the design of microstructures of materials with prescribed macroscopic material properties, and in Sub-section 6.2.2 the design of compliant mechanisms.

6.2.1 Design of Materials for Prescribed Macroscopic Properties

The problem of tailoring materials for specified elastic properties has attracted considerable interest over recent years (see, e.g., [2], [149], [150], [43], [62], [42], [151], [152]).

Milton and Cherkaev [153] developed a mathematical method for the design of materials which have an elasticity tensor matching an arbitrarily prescribed positive semi-definite fourth order elasticity tensor compatible with thermodynamics. A material with Poisson's ratio close to 1 was constructed by Milton [154]. These developments were based on the use of layered material microstructures constructed from a rigid and an infinitely weak material. The materials obtained this way serve to substantiate that prescribed as well as extreme elastic properties can be achieved, but they are merely mathematical tools than practical composites due to their widely differing length scales, cf. Lakes [155].

Sigmund [43, 62, 42, 151], on the other hand, has treated the material design problem as a topology optimization problem for a periodic microstructure represented by a base cell that only involves a single length scale. The topology optimization problem is formulated as minimization of the density of material in the base cell, subject to equality constraints that express the specification of the elastic properties. Some of the results will be presented in the sequel. In Sigmund [62],[42],[151], the microstructure is discretized by continuum-type finite elements where the material densities in the individual elements are taken as design variables. The effective elastic properties of the discretized microstructure are found by homogenization which leads to the definition of the material design problem as an inverse homogenization problem.

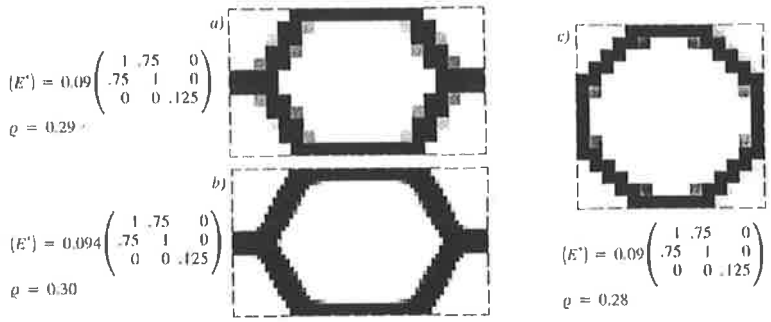


Figure 20: Honeycomb microstructures with Poisson's ratio equal to 0.75. Three different base cells, with (a) 21 × 12 elements, (b) 63 × 36 elements, and (c) 15 × 15 elements, are used. (From [151].)

The optimization problem can be stated as

$$\begin{aligned} &\text{Minimize} \quad \rho_{cell} = \int_\Omega \rho^p \, d\omega \\ &\text{Subject to} \quad E^H_{ijkl} - E^*_{ijkl} = 0 \\ &\qquad\qquad 0 < \rho_{min} \leq \rho \leq \rho_{max} \end{aligned} \tag{82}$$

where ρ_{cell} is the density of material in the base cell, the local material densities ρ are the design variables constrained by ρ_{min} and ρ_{max}, Ω denotes the design domain (base cell),

and p is a penalization power used to ensure that each element in the discretized cell will represent either solid or void. E^H_{ijkl} denotes the effective elasticity tensor determined by the homogenization procedure as briefly described in Section 3.1, and E^*_{ijkl} is the prescribed elasticity tensor.

The microstructures presented in the following two examples [151] are all obtained from rectangular base cells discretized by four node bilinear finite elements, and the base material for all examples has Young's modulus 0.91 and Poisson's ratio 0.3 such that a purely solid base cell will have $E_{1111} = 1.0$ (assuming plane stress).

As a first example, the optimization algorithm was used to 'reinvent' the 'perfect' honeycomb. Prescribing the elastic properties of a material with $E_{1111} = 0.09$ and Poisson's ratio 0.75 in a rectangular and quadratic domain for the base cell, the results in Figure 20 were obtained. The optimized microstructures in Figs. 20a,b are seen to display great similarity with the 'perfect' honeycomb. If the rectangular base cell domain in Figs. 20a,b is changed to a quadratic one discretized by 15×15 elements, the optimization yields the 'octagonal honeycomb' shown in Figure 20c. As the density of the material is very nearly the same for the 'perfect' and the 'octagonal' honeycomb, preference may be made from manufacturability considerations.

As a second example, we consider the design of materials with a negative Poisson's ratio, i.e., materials that expand transversely when subjected to elongation. It is well known that Poisson's ratios is confined to the interval from 0 to 1/2 for solid isotropic materials, but for porous (e.g., cellular) materials Poisson's ratio may approach the value -1 without violation of positive semi-definiteness of the elasticity tensor. Negative Poisson's ratio materials have attracted considerable interest in recent years, and use of such materials may, for example, be advantageous for hydrophones and mechanical fasteners (e.g., 'Rawlplugs'). However, numerical experiments in [151] have shown that only low overall stiffness can be obtained if extreme elastic properties are prescribed. The reader is referred to [43] for potential applications and for a survey of recent research. Figure 21 (left) displays topologies of base cells obtained by Sigmund [151] subject to specification of elastic properties of materials with a negative Poisson's ratio. The illustrations on the right hand side of Figure 21 are obtained by repeating the base cell periodically, and offer a clear picture of the aggregate microstructures; it is easily seen that horizontal elongation of the microstructures will give rise to expansion in the vertical direction, and thereby a negative value of the Poisson's ratio. An experimental project devoted to the manufacturing of the negative Poisson's ratio materials in real micro-scale (size of base cell: 30 μm) and subsequent testing of their elastic properties, has been successfully caried out at the Microelectronics Center in Lyngby, Denmark, see [156].

It will be finally illustrated via an example from Sigmund and Torquato [86] that the material topology optimization algorithm can be easily modified to be used for design of materials with prescribed thermoelastic properties. The design of a material with a prescribed thermal expansion tensor α^*_{ij} cannot be performed alone by designing the microstructure based on one material - it is necessary to consider a mixture of at least two materials with different thermal expansion coefficients α^1_{ij} and α^2_{ij}.

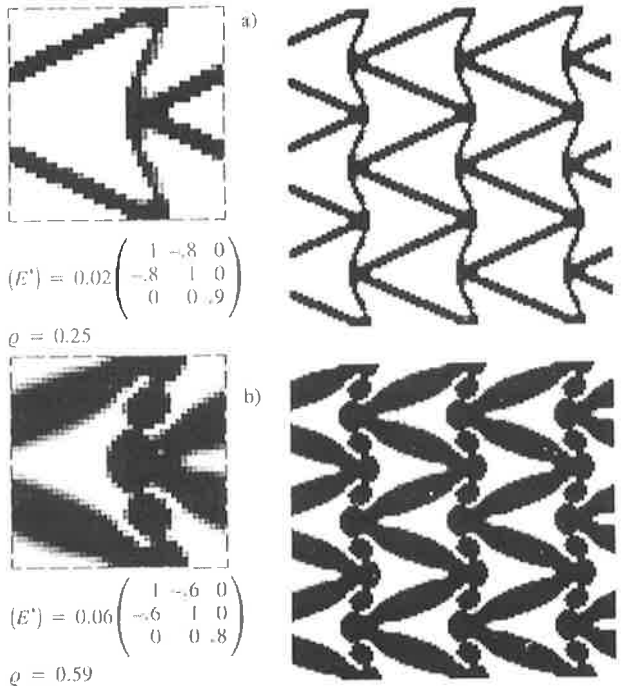


Figure 21: Negative Poisson's ratio materials [(a): -0.8; (b): -0.6] obtained from base cells with 40 × 40 elements and enforcement of vertical symmetry. (From [151].)

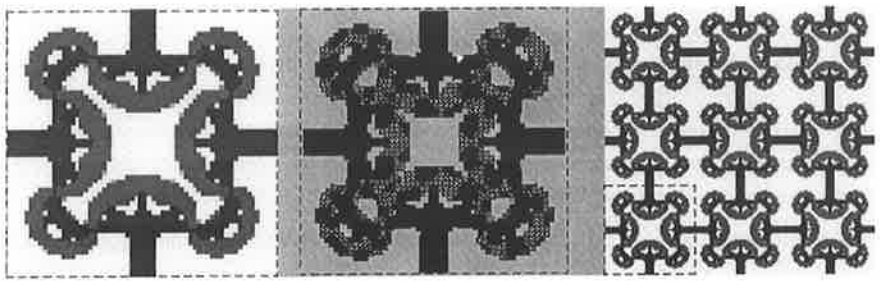


Figure 22: Design of a material with negative thermal expansion coefficient. Left: base cell (design domain) with topology optimized bi-material microstructure that contracts when heated. The grey and the black material phases have a high and a low positive thermal coefficient, respectively, and the white sub-domains are void. Middle: thermal displacement of microstructure when heated. Right: Aggregate negative thermal expansion material. (By courtesy of O. Sigmund.)

Consider now as an example (Sigmund & Torquato [86], see also Sigmund [157]) the problem of designing a material with a negative (isotropic) thermal expansion coefficient $\alpha^* < 0$ by including two isotropic material phases 1 and 2 with different, but both positive thermal expansion coefficients $\alpha^1 > \alpha^2 > 0$, in the base cell as shown in Figure 22. The grey phase in Fig. 22 has the high thermal expansion coefficient α^1 and the black phase has the low (but still positive) thermal expansion coefficient α^2. By performing the topology optimization on the base cell, one obtains the result shown in Fig. 22 (left). The resulting microstructure has a negative thermal expansion coefficient, cf. the deformation indicated in Fig. 22 (middle) of the microstructure when heated. Fig. 22 (right) shows the periodic material composed of repeated base cells. Studying the optimal microstructure, one notices that it consists of several small bi-material beams that in an intricate way make the periodic structure contract when heated, even though the materials it is built from expand when heated.

6.2.2 Topology Optimization of Compliant Mechanisms

The design (synthesis) of compliant mechanisms is a very important and rapidly increasing new field for the application of structural topology design methods. This chapter will present a few illustrative examples from the field.

A compliant mechanism gains its mobility from the flexibility of some or all of its members, as opposed to a conventional rigid-body mechanism. Advantages of compliant mechanisms are that they require fewer parts; are easy to fabricate; have less wear, friction, and backlash; have no need for lubrication; and have built-in restoring force. The concept of compliant or flexible mechanisms is not new (see, e.g., Burns & Crossley [158]) and is by no means restricted to small-scale mechanisms, but it has recently received increased attention because of the introduction of materials with superior properties and the rapidly expanding field of MicroElectroMechanical Systems (MEMS). The systems are built in sub-millimeter scale, are integrated with electronic circuits, and are manufactured using etching techniques from the semi-conductor industry. Such systems are already widely used for integrated sensor applications, and have potential applications for in-body surgery, health monitoring, micromanipulation, and nano-fabrication, see Kim *et al.* [159]. Compliant mechanisms are well-suited for MEMS because of the problems with friction and wear that prohibit use of conventional rigid-body mechanisms.

The optimum design of a compliant mechanism depends on, e.g., the manipulation task it should perform, the material of which it is to be made, the available design space, the locations of the input and output ports, and the available input force or actuator. Possible manipulation tasks can be crunching or clamping of workpieces, path generation, prescribed input/output force/deflection relationships, desired energy storage, and many others, see Sigmund [63, 157].

For some of the types of mechanisms just mentioned, the major design goal is maximization of the output force F_{out} (the force on the workpiece) for a given input force F_{in}. The ratio between the output and input forces $M = F_{out}/F_{in}$ is called the *Mechanical Advantage*. Three constraints are important in the synthesis of compliant mechanisms [63]:

(i) a constraint on the maximum stress in the mechanism (to hinder fatigue or failure), (ii) a constraint on deflection at the input port, and (iii) a constraint on volume to save material and cost. Here, the stress levels (i) in compliant mechanisms can be controlled indirectly by constraining the displacement at the input port [63].

Using the the methodology of topology optimization (Sigmund [63, 157]) the above design problem can be defined as the problem of finding the optimum mechanism topology within a given design domain Ω that satisfies the above-mentioned goals and constraints, i.e.,

$$\begin{array}{lll} \text{Maximize:} & \text{Mechanical Advantage} & M \\ \text{Subject to:} & \text{Volume constraint} & V \leq V^* \\ & \text{Input displacement} & \Delta_{in} \leq \Delta_{in}^* \end{array} \tag{83}$$

Since a typical structure for micromanipulation purposes is a gripping device, let us consider the design example from Sigmund [63] shown in Figure 23 (left). A horizontal input force at the mid-left side of the design domain should be converted to a closing of the jaws at the right-hand side of the design domain. The microgripper is constructed in Silicon, with E = 180 GPa and yield strength $\sigma_Y = 7$ GPa. For manufacturing reasons, the thickness is only $t = 7$ μm. To allow for underetching of the structure, it must be thin; i.e. have no wide elements. This can be controlled by allowing a low volume fraction of 20 %. The input force is $F_{in} = 1000$ μN, and the workpiece (which should be a blood cell) has stiffness K_s =1 N/mm.

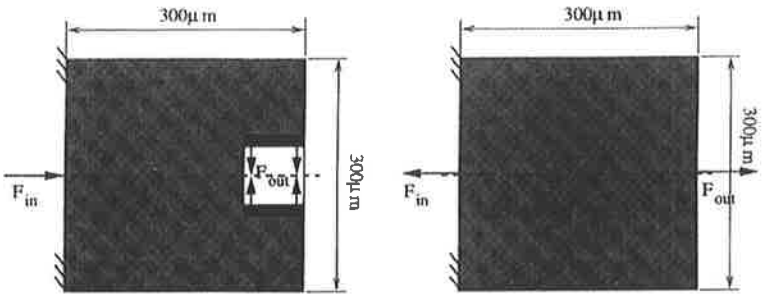


Figure 23: Design domains for a micro gripper (left) and a micro displacement inverter and amplifier (right). From Sigmund [63].

For an input displacement of $\Delta_{in}^* = 2$ μm the optimum mechanism topology is shown in Fig. 24 (top left). The resulting output displacement is $\Delta_{out}^* = 1.3$ μm, and the mechanical advantage is $M = 1.3$.

The output jaws of the gripping mechanism in Fig. 24 (top left) open and close in a crocodile-way, as seen in Fig. 24 (bottom left), i.e., the jaws do not move in parallel. If interest is in a gripping mechanism with parallel moving jaws, two output ports must

be considered. This is done by defining load conditions at the second output port (at the internal part of the jaws) and defining an extra mechanical advantage as the force relation between the extra output port and the input force. The objective function for the optimization problem in Eq. 83 then becomes

$$\phi = M_1 + M_2 - r(M_2 - M_1)^2 \tag{84}$$

where r is a penalty parameter. Formulating the objective function in this way, the two mechanical advantages are constrained to be equal (by least square error), and they are at the same time maximized. Furthermore, an extra displacement constraint $\Delta_{in(2)} \leq \Delta_{in}^*$ is added to the optimization problem.

Repeating the optimization problem for two output ports and workpiece stiffness K_S = 0.5 N/μm, the mechanism topology in Fig. 24 (top right) is obtained. From the displaced mechanism in Fig. 24 (bottom right), it can be see that the output jaws now move in parallel. The two resulting mechanical advantages are $M_1 = M_2 = 0.9$ and $\Delta_{out} = 1.8$ μm.

This example shows that a mechanism with complex output behavior can be designed by the topology optimization technique.

The following examples from Sigmund [63] demonstrate the design of a simple force inverter and a displacement amplifier and discuss the energy conversion for such mechanisms.

The dimensions are the same as for the microgripper, and the design domain is sketched in Fig. 22 (right). We first consider the design of a displacement (and force inverter). The input force is $F_{in} = 1000$ μN, the workpiece stiffness is $K_S = 0.2$ N/mm, and the allowed input displacement is $\Delta_{in}^* = 5.0$ μm. The resulting mechanism is shown in Fig. 24 (left). The resulting output displacement is $\Delta_{out} = 4.8$ μm, and the mechanical advantage is M = 0.97. For this mechanism, the work done by the input force is calculated as $W_{in} = 5$ nJ, and the work done by the output force is calculated as $W_{out} = 4.7$ nJ, which means that 7 % of the input work was stored as elastic energy in the mechanism. Note that if this had been a rigid-body mechanism, there would have been no 'loss' of work between the input and the output ports.

Constraining the input displacement to be $\Delta_{in}^* = 2.0$ μm, and letting the stiffness of the workpiece be $K_S = 0.03$ N/mm, the resulting output displacement is $\Delta_{out} = 7.8$ μm, and the mechanical advantage is $M = 0.23$. The mechanism can be said to be a 1 : -3.9 displacement amplifier. The optimum topology for the mechanism is shown in Fig. 25 (right). For this mechanism, the input and output works can be found to be $W_{in} = 2.0$ nJ and $W_{out} = 1.8$ nJ, respectively, which means that 10 % of the work was stored as elastic energy in the mechanism. If a high value of displacement amplification is needed, one could mount several of the 1 : -3.9 displacement amplifiers in series. However, this will result in poor work transmission, since 10 % of the input work is lost for each amplifier.

The assumption of linear displacements is close to being violated for the last two design examples. If the displacement amplifier in Fig. 25 (right) is studied, it is clear that the

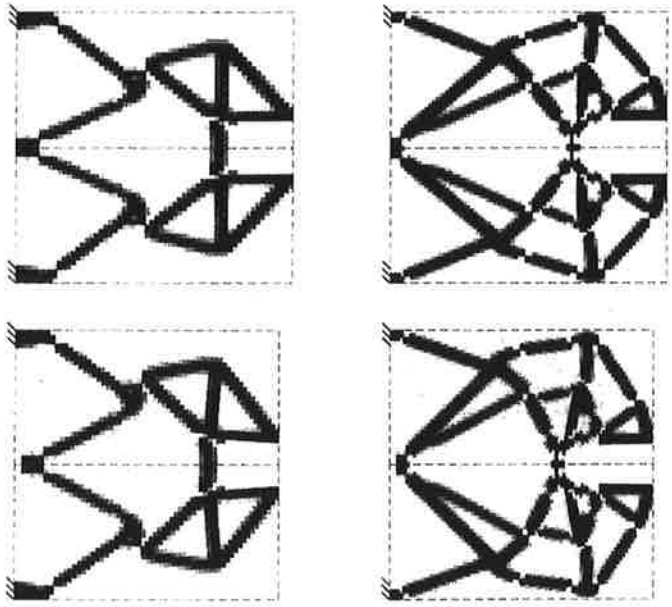


Figure 24: Micro gripping mechanisms. Top left: optimum microgripper topology for one output port and its displacement pattern (bottom left). Top right: optimum microgripper topology for two output ports and its displacement pattern (bottom left). From Sigmund [63]

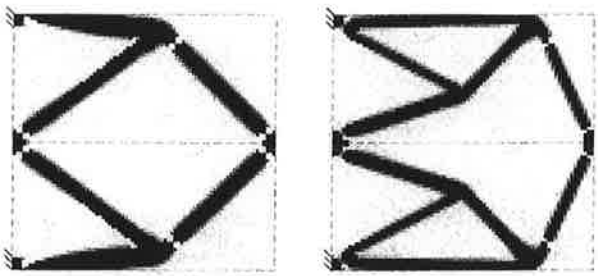


Figure 25: Micro displacement inverters. From Sigmund [63]

output displacement is limited by the point at which the two bars that lead to the output port become vertical. A linear displacement finite element model would not predict this locking. It is therefore clear that non-linear displacement theory must be considered, in order to generalize the design method to the design of large displacement mechanisms, and this has been very recently done by Pedersen *et al.* [160].

7 Conclusions

Basic concepts and knowledge concerning optimum topology design of continuum structures were discussed in this paper, and it has been attempted to give an overview of the current status of the field.

Problem formulations based on solid, isotropic material models are generally not well-posed since the design space is not closed in an appropriate sense, and a regularization of the formulation is then required. The regularization can be achieved by either extending (relaxing) the design space such as to include solutions with microstructure in the problem formulation, or one can restrict the space of the admissible designs.

If it is desired to obtain or acquire knowledge of the global optimum solution associated with the highest possible value of the performance index (for example in order to know the upper-bound benchmark for other designs), it is necessary to extend the design space by including optimum microstructures in the problem formulation, and to determine their macroscopic material properties by a homogenization technique. The layered 2D and 3D microstructures considered in this paper are optimum for integral stiffness (minimum compliance) design under a single case of loading, but optimum microstructures are as of yet unknown for most other design objectives.

Global optimum designs of the type just mentioned, generally consist of composite material in large sub-domains. This simply reflects the fact that composite materials are in general much more efficient than isotropic, solid materials. However, the large sub-domains with composite often make it difficult to visualize the overall structural topology, and are less attractive from the point of view of manufacture. In order to reduce the composite regions, one may apply a non-optimum microstructure like the "hole-in-cell" microstructure considered in the paper, or introduce schemes for penalization of composite in the formulation or solution procedure for the problem.

A direct way to obtain distinct black-white designs within the realm of a well-posed problem, is to restrict the design space by including in the problem formulation an upper bound constraint on the perimeter or surface area of the structure, or, what is simpler and used more and more often, to apply an appropriate filtering technique, which will also have the effect of restricting rapid variations in the structural design. Approaches of this kind are advantageous because they admit application of simple, material models like the SIMP model (which by itself penalizes intermediate material densities). This model is very popular in commercial topology optimization codes, and by now also very often used in research directed toward extending the scope of topology design, since the simplicity of the model facilitates research into other aspects of the problems.

It is obvious that the creation of quite distinct black-white topologies by the approaches just mentioned, greatly facilitates the manufacture of the designs obtained, but it should be also borne in mind that there is a price to be paid for this, namely that the solution will be associated with a lower value of the performance index as compared to the corresponding global optimum solution obtained by usage of an optimum microstructure.

On the basis of the account of the present state of knowledge presented in the main part of this paper, the last chapter gives a brief survey of literature that reflects recent developments in topology optimization of continuum structures. Furthermore, new avenues of application of topology design are illustrated by way of very recent results in areas of tailoring materials for prescribed mechanical properties and design of compliant mechanisms with focus on MicroElectroMechanical Systems (MEMS).

8 Acknowledgement

The first author gratefully acknowledges the sponsorship from the Max Planck Research Award program and the collaboration with the second author funded by the German Federal Ministry for Education, Science, Research and Technology.

References

[1] G. I. N. Rozvany, M. P. Bendsøe, and U. Kirsch. Layout Optimization of Structures. *Applied Mechanics Reviews*, 48:41–119, 1995.

[2] M. P. Bendsøe. *Optimization of Structural Topology, Shape, and Material.* Springer-Verlag, Heidelberg, Germany, 1995.

[3] N. Olhoff and J. E. Taylor. On structural optimization. *J. Appl. Mech.*, 50:1134–1151, 1983.

[4] N. Olhoff. On optimum design of structures and materials. *Meccanica*, 31:143–161, 1996.

[5] M. P. Bendsøe. Structural optimization of and with advanced materials. In T. Tatsumi, E. Watanabe, and T. Kambe, editors, *Theoretical and Applied Mechanics*, pages 269–284. Elsevier Science B.V., Amsterdam, 1997.

[6] N. Olhoff, M. P. Bendsøe, and J. Rasmussen. On CAD-integrated Structural Topology and Design Optimization. *Comp. Meths. Appl. Mech. Engrg.*, 89:259–279, 1991.

[7] M. Bremicker, M. Chirehdast, N. Kikuchi, and P. Y. Papalambros. Integrated topology and shape optimization in structural design. *Mechanics of Structures and Machines*, 19(4):551–587, 1991.

[8] M. P. Bendsøe and C. A. Mota Soares, editors. *Topology Design of Structures.* NATO ASI, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1993.

[9] P. Pedersen, editor. *Optimal Design with Advanced Materials.* Elsevier, Amsterdam, The Netherlands, 1993.

[10] N. Olhoff and G. I. N. Rozvany, editors. *Structural and Multidisciplinary Optimization*, WCSMO-1, Goslar, Germany, 1995. Pergamon, Oxford, UK.

[11] W. Gutkowski and Z. Mróz, editors. *Structural and Multidisciplinary Optimization*, WCSMO-2, Warsaw, Poland, 1997. Institute of Fundamental Technological Research.

[12] C. L. Bloebaum, editor. *WCSMO 3 - Third World Congress of Structural and Multidisciplinary Optimization*, Buffalo, New York, USA, 1999 (to appear).

[13] M. P. Bendsøe. Recent development in topology design of continuum structures. In *Paper presented at the NAFEMS Seminar on Topology Optimization*, Aalen, Germany, Sept. 23, 1997.

[14] B. Hassani and E. Hinton. A review of homogenization and topology optimization; part i: homogenization theory for media with periodic structure; part ii: analytical and numerical solution of homogenization equations; part iii: topology optimization using optimality criteria. *Computers and Structures*, 1999.

[15] M. P. Bendsøe and N. Kikuchi. Generating Optimal Topologies in Structural Design using a Homogenizaiton Method. *Comp. Meth. Appl. Mechs. Engrg.*, 71:197–224, 1988.

[16] M. P. Bendsøe. Optimal Shape Design as a Material Distribution Problem. *Struct. Optim.*, 1:193–202, 1989.

[17] K. A. Lurie. *Applied Optimal Control Theory of Distributed Systems.* Plenum Press, New York ,USA, 1993.

[18] K. A. Lurie, A. V. Cherkaev, and A. Fedorov. Regularization of Optimal Design Problems for Bars and Plates. *JOTA*, pages 37:499–522(Part I), 37:523–543(Part II), 42:247–282(Part III), 1982.

[19] N. Olhoff, K. A. Lurie, A. V. Cherkaev, and A. Fedorov. Sliding Regimes of Anisotropy in Optimal Design of Vibrating Axisymetric Plates. *Int. J. Solids Struct.*, 17(10):931–948, 1981.

[20] G. Cheng and N. Olhoff. An Investigation Concerning Optimal Design of Solid Elastic Plates. *Int. J. Solids Struct.*, 17:305–323, 1981.

[21] R. V. Kohn and G. Strang. Optimal Design and Relaxation of Variational Problems. *Comm. Pure Appl. Math.*, 39:1–25(part I), 139–182(partII) and 353–357(partIII), 1986.

[22] R. V. Kohn and G. Strang. Optimal Design in Elasticity and Plasticity. *Num. Meths. Engrg.*, 22:183–188, 1986.

[23] G. Cheng and N. Olhoff. Regularized Formulation for Optimal Design of Axisymetric Plates. *Int. J. Solids Struct.*, 18(2):153–169, 1982.

[24] G. Cheng. On non-smoothness in optimal design of solid, elastic plates. *Int. J. Solids Struct.*, 17:795–810, 1981.

[25] F. Niordson. Some new results regarding optimal design of elastic plates. In Eschenuer H. and Olhoff N., editors, *Optimization Methods in Structural Design*, Euromech-Colloquium 164, pages 380–386. B.I.-Wissenschaftsverlag, Wien, 1982.

[26] M. P. Bendsøe. G-closure and homogenization problems arising in plate optimization. In Eschenuer H. and Olhoff N., editors, *Optimization Methods in Structural Design*, Euromech-Colloquium 164, pages 270–275. B.I.-Wissenschaftsverlag, Wien, 1982.

[27] M. Avellaneda. Optimal Bounds and Microgeometries for Elastic Two-Phase Composites. *SIAM J. Appl. Math.*, 47(6):1216–1228, 1987.

[28] R. Lipton. On Optimal Reinforcement of Plates and Choice of Design Parameters. *Control and Cybernetics*, 23(3):481–493, 1994.

[29] J. Thomsen. Optimization of Properties of Anisotropic Materials and Topologies of Structures (in Danish). Ph.D thesis, Institute of Mechanical Engineering, Aalborg University, Denmark, 1992.

[30] G. Allaire, E. Bonnetier, G. Francfort, and F. Jouve. Shape optimization by the homogenization method. Report, Lab. d'Analyse Numerique, Universite Paris 6, 1996.

[31] L. V. Gibiansky and A. V. Cherkaev. Microstructures of composites of extremal rigidity and exact estimates of provided energy density (in russian). Technical Report 1155, A. F. Ioffe Physical-Technical Institute, Academy of Sciences of the USSR, Leningrad, 1987. English translation in [152].

[32] A. V. Cherkaev and R. Palais. Optimal design of three-dimensional axisymmetric elastic structures. *Structural Optimization*, 12:35–45, 1996.

[33] G. Buttazzo and G. Dal Maso. An existence result for a class of shape optimization problems. *Arch. Ratonal Mech. Anal.*, 122:183–195, 1993.

[34] G. Allaire. Explicit lamination parameters for three-dimensional shape optimization. *Control & Cybernetics*, 23:311–326, 1994.

[35] R. B. Haber, M. P. Bendsøe, and C. S. Jog. A new approach to variable-topology shape design using a constraint on the perimeter. *Structural Optimization*, 11:1–12, 1996.

[36] R. B. Haber, C. S. Jog, and M. P. Bendsøe. Variable-topology shape optimization with a control on perimeter. In *Advances in Design Automation*, volume 69, pages 261–272. ASME, 1994.

[37] R. B. Haber, C. S. Jog, and M. P. Bendsøe. The perimeter method – a new approach to variable-topology shape optimization. In N. Olhoff and G.I.N. Rozvany, editors, *Structural and Multidisciplinary Optimization*, pages 153–160. WCSMO-1, Elsevier Science, Oxford, U.K., 1995.

[38] L. Ambrosio and G. Buttazzo. An optimal design problem with perimeter penalization. *Calc. Var.*, 1:55–69, 1993.

[39] P. R. Fernandes, H. Rodrigues, and J. M. Guedes. Topology optimization of 3d linear elastic structures with a constraint on "perimeter". IDMEC/CPM 96/004, Instituto Superior Technico, Lisboa, Portugal, 1996.

[40] J. Petersson and O. Sigmund. Slope constrained topology optimization. *Int. J. Num. Meths. Engng.*, 41:1417–1434, 1998.

[41] J. Petersson. Some convergence results in perimeter-controlled topology optimization. Technical Report LiTH-IKP-R-995, Inst. of Technology, Dept. Mech. Engng., Linkoping University, 1998.

[42] O. Sigmund. Tailoring materials with prescribed elastic properties. *Mech. Mat.*, 20:351–368, 1995.

[43] O. Sigmund. Design of material structures using topology optimization, DCAMM report S 69. Ph.d. thesis, Department of Solid Mechanics, Technical University of Denmark, 1994.

[44] H. A. Eschenauer, V. Kobelev, and A. Schumacher. Bubble method of topology and shape optimization of structures. *Structural Optimization*, 8:42–51, 1994.

[45] H. A. Eschenauer and A. Schumacher. Topology and shape optimization procedures using hole positioning criteria – theory and applications. In G.I.N. Rozvany, editor, *Topology Optimization in Structural Mechanics*, volume 374 of *CISM Courses and Lectures*, pages 135–196. Springer, Vienna, 1997.

[46] H. Eschenauer, A. Schumacher, and T. Vietor. Decision makings for initial designs made of advanced materials. In M. P. Bendsøe and C . A. Mota Soares, editors, *Topology Design of Structures*, pages 469–480. Kluwer Academic Publishers, 1993.

[47] H. Eschenauer and A. Schumacher. Simultaneous topology and shape optimization of structures. In N. Olhoff and G.I.N. Rozvany, editors, *WCSMO 1 - First World Congress of Structural and Multidisciplinary Optimization*, pages 177–184. Pergamon Press, Oxford, U.K., 1995.

[48] A. Schumacher. Topologieoptimierung von bauteilstrukturen unter verwendung von lochpositionierungskriterien. Diss., FOMAAS, TIM-Bericht T09-11.95, University of Siegen, University of Siegen, 1995.

[49] A. Diaz and M. P. Bendsøe. Shape Optimization of Structures for Multiple Loading Situations using a Homogenization Method. *Struct. Optim.*, 4:17–22, 1992.

[50] N. Olhoff, J. Thomsen, and J. Rasmussen. Topology Optimization of Bi-Material Structures. In P Pedersen, editor, *Optimal Design with Advanced Materials*, pages 191–206, Lyngby, Denmark, 1993. Elsevier, Amsterdam, The Netherlands, 1993.

[51] K. Suzuki and N. Kikuchi. Layout Optimization using the Homogenization Method: Generalized Layout Design of Three-Dimensional Shells for Car Bodies. In G. I. N. Rozvany, editor, *Optimization of Large Structural Systems*, pages 110–126, Berchtesgaden, Germany, 1991. Nato-ASI, Lecture Notes, 1991.

[52] C. A. Soto and A. Diaz. Optimum Layout of Plate Structures using Homogenization. In M. P. Bendsøe and C. A. Mota Soares, editors, *Topology Design of Structures*, 1993.

[53] C. A. Soto and A. Diaz. A Model for Layout Optimization of Plate Structures. In P Pedersen, editor, *Optimal Design with Advanced Materials*, pages 337–350, Lyngby, Denmark, 1993. Elsevier, Amsterdam, The Netherlands, 1993.

[54] A. Diaz, R. Lipton, and C. A. Soto. A New Formulation of the Problem of Optimum Reinforcement of Reissner-Mindlin Plates. *Comp. Meth. Appl. Mechs. Engrg*, 1994.

[55] C. A. Soto and A. Diaz. Layout of Plate Structures for Improved Dynamic Response using a Homogenization Method. In Gilmore et. al., editor, *Advances in Design Automatization*, pages 667–674, Albuquerqe, New-Mexico, USA, 1993. ASME, 1993.

[56] L. A. Krog and N. Olhoff. Topology Optimization of Integral Rib Reinforcement of Plate and Shell Structures with Weighted-Sum and Max-Min Objectives. In D Bestle and W Schiehlen, editors, *Optimization of Mechanical Systems.*, pages 171–179, Stuttgart, Germany, 1995. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1996.

[57] J. Folgado, H. Rodriques, and J. M. Guedes. Layout Design of Plate Reinforcements with a Buckling Load Criterion. In N Olhoff and GIN Rozvany, editors, *Structural and Multidisciplinary Optimization*, pages 659–666, Goslar, Germany, 1995. ISSMO, Pergamon, Oxford, UK, 1995.

[58] G. Cheng and Z. Jiang. Study on topology optimization with stress constraints. *Engineering Optimization*, 20:129–148, 1992.

[59] G. Cheng and X. Guo. ε-related approach in structural topology optimization. *Structural Optimization*, 13(4):258–266, 1997.

[60] P. Duysinx and M. P. Bendsøe. Topology optimization of continuum structures with local stress constraints. *Int. J. Num. Meths. Engng.*, (to appear).

[61] P. Duysinx and O. Sigmund. New developments in handling stress constraints in optimal material distribution. In *7'th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, volume 1, pages 1501–1509. AIAA, AIAA, USA, 1998.

[62] O. Sigmund. Materials with prescribed constitutive parameters: An inverse homogenization problem. *Int. J. Solids Structures*, 31:2313–2329, 1994.

[63] O. Sigmund. On the design of compliant mechanisms using topology optimization. *Mech. Struct. Mach.*, 25:493–524, 1997.

[64] O. Sigmund. Systematic design of electrothermomechanical microactuators using topology optimization. In *Modelling and Simulation of Microsystems, Semiconductors, Sensors, and Actuators*, pages 1492–1500, Santa Clara, California, USA, 1998. MSM98.

[65] O. Sigmund. Topology optimization in multiphysics problems. In *7'th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, volume 1, pages 1492–1500. AIAA, AIAA, USA, 1998.

[66] K. Maute and E. Ramm. Adptive techniques in topology optimization. In *5'th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, volume 1, pages 121–131. AIAA, AIAA, USA, 1994.

[67] K. Maute and E. Ramm. Adaptive topology optimization. *Structural Optimization*, 10:100–112, 1995.

[68] K. Maute, S. Schwarz, and E. Ramm. Adptive topology optimization of elastoplastic structures. *Structural Optimization*, 15(81-91), 1998.

[69] K. Maute. *Topologie- und Formoptimierung von dünnwandigen Tragwerken.* Bericht nr. 25, Institut für Baustatik, Universität Stuttgart, Stuttgart, Germany, 1998.

[70] G. Allaire and R. V. Kohn. Optimal design for minimum weight and compliance in plane stress using extremal microstructures. *European J. Mech. A.*, 12:839–878, 1993.

[71] M. Beckers. Optimisation de structures en variables discretes. These de Doctorat 181, Universite de Liege, Belgium, Liege, Belgium, 1997.

[72] M. Beckers. Topology optimization using a dual method with discrete variables. *Struct. Optim.*, 17:14–24, 1999.

[73] M. Beckers and C. Fleury. Topology optimization involving discrete variables. In W. Gutkowski and Z. Mróz, editors, *Second World Congress of Structural and Multidisciplinary Optimization*, pages 533–538. WCSMO-2, Inst. Fundamental Technological Research, Warsaw, 1997.

[74] M. Avellaneda and G. W. Milton. Bounds on the Effective Elastic Tensor of Composites based on Two-Point Correlations. In D Hui and TJ Kozik, editors, *Composite Material Technology*, pages 89–93. ASME, 1989.

[75] A. Diaz and R. Lipton. Optimal material layout for 3D elastic structures. *Structural Optimization*, 13:60–64, 1997.

[76] J. B. Jacobsen, N. Olhoff, and E. Rønholt. Generalized shape optimization of three-dimensional structures using materials with optimum microstructures. *Mech. Mat.*, 28:207–225, 1997.

[77] N. Olhoff, E. Rønholt, and J. Scheel. Topology optimization of three-dimensional structures using optimum microstructures. *Structural Optimization*, 16:1–18, 1998.

[78] T. Burns and A. Cherkaev. Optimal distribution of multimaterial composites for torsional beams. *Structural Optimization*, 13:4–11, 1997.

[79] M. Zhou and GIN Rozvany. The COC algorithm, Part ii: topological, geometrical and generalized shape optimization. *Appl. Mech. Engr.*, 89:309–336, 1991.

[80] H. Mlejnek and R. Schirrmacher. An engineering approach to optimal material distribution and shape finding. *Comp. Meth. Appl. Mechs. Engng*, 106(1–26), 1993.

[81] G. I. N. Rozvany, M. Zhou, T. Birker, and O. Sigmund. Topology optimization using iterative contimuum-type optimality criteria (COC) methods for discretized systems. In M. P. Bendsøe and C. A. Mota Soares, editors, *Topology Design of Structures*, 1993.

[82] M. P. Bendsøe and A. Diaz. A method for treating damage related criteria in optimal topology design of continuum structures. *Structural Optimization*, (to appear).

[83] E. Hinton, J. Sienz, S. Bulman, and B. Hassani. Fully integrated design optimization of engineering structures using adaptive finite element models. In W. Gutkowski and Z. Mróz, editors, *WCSMO 2 - Second World Congress of Structural and Multidisciplinary Optimization*, pages 73–78. Inst. of Fundamental Technological Research, Warsaw, 1997.

[84] R. Magister and P.U. Post. Use of structural optimization in the development of pnenmatic components. In N. Olhoff and G. I. N. Rozvany, editors, *First World Congress of Structural and Multidisciplinary Optimization*, pages 445–452, Goslar, Germany, 1995. Pergamon, Oxford, UK, 1995.

[85] K. Maute and E. Ramm. Adaptive topology optimization of shell structures. *AIAA J.*, 35:1767–1773, 1997.

[86] O. Sigmund and S. Torquato. Design of materials with extreme thermal expansion using a three-phase topology optimization method. *J. Mech. Phys. Solids*, 45:1037–1067, 1997.

[87] O. Sigmund, S. Torquato, and I. A. Aksay. On the design of 1-3 piezocomposites using topology optimization. *J. Mat. Res.*, 13:1038–1048, 1998.

[88] R. J. Yang and C. J. Chen. Stress-based topology optimizaton. *Struct. Optim.*, 12:98–105, 1998.

[89] M. P. Bendsøe and O. Sigmund. Material interpolation in topology optimization. *Arch. Appl. Mech.*, (to appear).

[90] Z. Hashin and S. Shtrikman. A variational approach to the theory of the elastic behaviour of multiphase materials. *J. Mech. Phys. Solids*, 11:127–140, 1963.

[91] A. Bensoussan, J. L. Lions, and G. Papanicolaou. Asymptotic analysis for periodic structures. Technical report, North-Holland, Amsterdam, 1978.

[92] E. Sanchez-Palencia. Non-homogeneous media and vibration theory. Lecture Notes in Physics 127, Springer Verlag, Berlin, 1980.

[93] J. F. Bourgat. Numerical experiments of the homogenization method for operators with periodic coefficients. Lecture Notes in Mathematics 740: 330-356, Springer-Verlag, Berlin, 1977.

[94] J. M. Guedes and N. Kikuchi. Pre- and postprocessing for materials based on the homogenization method with adaptive element methods. *Comput. Meth. Appl. Mech. Engrg.*, 83:143–198, 1990.

[95] C. A. Soto and A. Diaz. On Modelling of Ribbed Plates for Shape Optimization. *Stuct. Optim.*, 6:175–188, 1993.

[96] C. A. Soto. Shape Optimization of Plate Structures using Plate Homogenization with Applications in Mechanical Design. Ph.D thesis, Department of Mechanical Engineering, Michigan State University, USA, 1993.

[97] L. A. Krog. Layout Optimization of Disk, Plate, and Shell Structures. Ph.D thesis, Special Report No. 27, Institute of Mechanical Engineering, Aalborg University, Denmark, 1996.

[98] L. A. Krog and N. Olhoff. Optimum topology and reinforcement design of disk and plate structures with multiple stiffness and eigenfrequency objectives. *Computers and Stuctures*, 72:535–563, 1999.

[99] R. Lipton. Optimal Design and Relaxation for Reinforced Plates Subject to Random Transverse Loads. *J. Probabilistic Engineering Mechanics*, 9:167–177, 1994.

[100] J. R. Vinson and R. L. Sierakowski. *The Behaviour of Structures Composed of Composite Materials.* Martinus Nijhoff Publishers, Dordrecht, The Netherlands, 1987.

[101] P. Pedersen. On Optimal Orientation of Orthotropic Materials. *Struct. Optim.*, 1:101–106, 1989.

[102] P. Pedersen. Bounds on Elastic Energy in Solids of Orthotropic Materials. *Struct. Optim.*, 2:55–63, 1990.

[103] P. Pedersen. On Thickness and Orientational Design with Orthotropic Materials. *Struct. Optim.*, 3:69–78, 1991.

[104] S. W. Tsai and N. J. Pagano. Invariant properties of composite materials. In S. W. Tsai, J. C. Halpin, and N. J. Pagano, editors, *Composite Materials Workshop*, pages 233–253. Technomic Publishing Co., Stamfort Connecticut, 1968.

[105] G. A. Francfort and F. Murat. Homogenization and Optimal Bounds in Linear Elasticity. *Arch. Rat. Mech. Anal.*, 94:307–334, 1986.

[106] M. Miki. Material design of composite laminates with required in-plane elastic properties. In T. Hayashi, K. Kawata, and M. Umekawa, editors, *Progress in Science and Engineering of Composites*, pages 1725–1731. ICCM-IV, Tokyo, 1982.

[107] H. Fukunaga and G. N. Vanderplaats. Stiffness optimization of orthotropic laminated composites using lamination parameters. *AAIA Journal*, 29:641–646, 1991.

[108] M. Miki and Y. Sugiama. Optimal design of laminated composite plates using lamination parameters. *AAIA Journal*, pages 275–283, 1991.

[109] J. L. Grenestedt and P. Gudmondson. Layup optimization of composite material structures. In P. Pedersen, editor, *IUTAM Symp. Optimal Design with advanced materials*, pages 311–336. Elsevier Science Publ. B.V., 1993.

[110] V. B. Hammer, M. P. Bendsøe, R. Lipton, and P. Pedersen. Parametrization in laminate design for optimal compliance. *International Journal of Solids and Structures*, 34(4):415–434, 1997.

[111] V. B. Hammer. Design of composite laminates with optimized stiffness, strenght and damage properties., DCAMM report S 72. Ph.d. thesis, Departement for Solid Mechanics, Technical University of Denmark, 1997.

[112] V. B. Hammer. Optimal laminate design subject to single membrane loads. *Structural optimization*, 17:65–74, 1999.

[113] J. Foldager. Design optimization of laminated composite plates divided into rectangular patches with use of lamination parameters. In W. Gutkowski and Z. Mróz, editors, *WCSMO 2 - Second World Congress of Structural and Multidisciplinary Optimization*, pages 669–675. Inst. of Fundamental Technological Research, Warsaw, 1997.

[114] J. P. Foldager. Design of composite structures. Ph.D thesis, Institute of Mechanical Engineering, Aalborg University, Denmark, 1999.

[115] J. P. Foldager, N. Olhoff, and J. S. Hansen. A general procedure forcing convexity in ply angle optimization in composite laminates. *Structural Optimization*, 16(2/3):201–212, 1998.

[116] R. Lipton and A. Diaz. Moment Formulations for Optimum Layout in 3D Elasticity. In N Olhoff and GIN Rozvany, editors, *Structural and Multidisciplinary Optimization*, pages 161–168, Goslar, Germany, 1995. Pergamon, Oxford, UK, 1995.

[117] M. G. Krein and A. A. Nudelman. *The Markov Moment Problem and Extremal Problems.* Translation of Mathematical Monographs, 50. American Mathematical Society, 1977.

[118] A. G. M. Michell. The limits of economy of material in frame-structures. *Phil. Mag.*, 8(589–597), 1904.

[119] C. Jog, R. B. Haber, and M. P. Bendsøe. Topology Design with Optimized Self-Addaptive Materials. *Int. J. Num. Meth. Engng.*, 37:1323–1350, 1994.

[120] P. Duysinx. Layout optimization: A mathematical programming approach. DCAMM Report 540, Technical University of Denmark, Technical University of Denmark, 1997.

[121] L. C. Evans and R. F. Gariepy. *Measure theory and fine properties of functions.* Boca Ratan, CRC Press, 1992.

[122] R. L. Wheedon and A. Zygmund. Measure and integral: an introduction to real analysis. volume 43 of *Monographs in Pure and Applied Math.* New York, 1977.

[123] F.I. Niordson. Optimal design of plates with a constraint on the slope of the thickness function. *Int. J. Solids Structures*, 19(141–151), 1983.

[124] M. G. Mullender, R. Huiskes, and H. Weihnans. A physiological approach to the simulation of bone remodelling as a self-organizational control process. *J. Biomechanics*, 11(1389–1394), 1994.

[125] O. Sigmund and J. Petersson. Numerical instabilities in topology optimization. *Struct. Optim.*, 16:68–75, 1998.

[126] G. Farin. Splines in cad/cam. In *In Survey on math. for industry*, pages 39–73. Springer-Verlag, Vienna, Austria, 1991.

[127] H. Eschenauer, J. Geilen, and H. J. Wahl. SAPOP - an optimization procedure for multicriteria structural design. In H. Hrnlein and K. Schittkowski, editors, *Numerical Methods in FE-based Structural Optimization Systems, Int. Series of Num. Math.* Birhhuser-Verlag, Germany, 1993.

[128] R. Korycki. Multiparametric shape optimal design of disks. *Structural Optimization*, 9:25–32, 1995.

[129] L. A. Krog and N. Olhoff. Topology and reinforcement layout optimization of disk, plate, and shell structures. In G. I. N. Rozvany, editor, *Topology Optimization in Structural Mechanics, Volume 374 of CISM Courses and Lectures*, pages 237–322. Springer, Vienna, 1997.

[130] C. Jog. Distributed-parameter optimization and topology design for nonlinear thermoelasticity. *Comp. Meth. Appl. Mech. Engrg.*, 132:117–134, 1996.

[131] K. Yuge and N. Kikuchi. Optimization of a frame structure subjected to a plastic deformation. *Struct. Optim.*, 10:197–208, 1995.

[132] C. C. Swan and J. S. Arora. Topology design of material layout in structured composites of high stiffness and strength. *Struct. Optim.*, 13:45–59, 1997.

[133] S. Min, N. Kikuchi, Y. C. Park, S. Kim, and S. Chang. Optimal reinforcement design of structures under impact loads. In W. Gutkowski and Z. Mróz, editors, *WCSMO 2 - Second World Congress of Structural and Multidisciplinary Optimization*, pages 583–588. Inst. of Fundamental Technological Research, Warsaw, 1997.

[134] In P. Pedersen and M. P. Bendsøe, editors, *IUTAM Symp. Synthesis in Bio Solid Mechanics, Copenhagen, Denmark, May 24-27 1998.* Kluwer, Dordrecht, 1999.

[135] J. Folgado and H. Rodriques. Topology optimization methods applied to the design of orthopaedics implants. In W. Gutkowski and Z. Mróz, editors, *WCSMO 2 - Second World Congress of Structural and Multidisciplinary Optimization*, pages 563–568. Inst. of Fundamental Technological Research, Warsaw, 1997.

[136] M. Tanaka, T. Adachi, and Y. Tomita. Optimum design of lattice continuum material suggested by mechanical adaptation of cancellous bone. In N. Olhoff and G. I. N. Rozvany, editors, *WCSMO 1 - First World Congress of Structural and Multidisciplinary Optimization*, pages 185–192. Pergamon Press, Oxford, U.K., 1995.

[137] M. Bagge. Remodeling of bone structures, DCAMM report S 84. Ph.d. thesis, Departement for Solid Mechanics, Technical University of Denmark, 1999.

[138] A. Diaz and N. Kikuchi. Solution to Shape and Topology Eigenvalue Optimization Problems using a Homogenization Method. *Int. J. Num. Meth. in Engng.*, 35:1487–1502, 1992.

[139] Z-D Ma, Cheng H-C, and N. Kikuchi. Structural design for obtaining desired eigenfrequencies by using the topology and shape optimization method. *Computing Systems in Engng.*, 5:77–89, 1994.

[140] Z-D Ma, , N. Kikuchi, and Cheng H-C. Topological design for vibrating structures. *Comp. Meth. Appl. Mech. Engrg.*, 121:259–280, 1995.

[141] Z-D Ma, , N. Kikuchi, Cheng H-C, and I. Hagiwara. Topological optimization technique for free vibration problems. *J. Appl. Mech.*, March 1995.

[142] G. Cheng. Some development of structural topology optimization. In T. Tatsumi, E. Watanabe, and K. Kambe, editors, *Theoretical and Applied Mechanics 1996*, pages 379–394. Elsevier Sciences B.V., Amsterdam, 1997.

[143] J. S. Ou and N. Kikuchi. Optimal design of controlled structures. *Structural Optimization*, 11:19–28, 1996.

[144] J. S. Ou and N. Kikuchi. Integrated optimal structural and vibration control design. *Structural Optimization*, 12:209–216, 1996.

[145] V. B. Hammer and N. Olhoff. Topology optimization with design dependent loads. In C. L. Bloebaum, editor, *WCSMO 3 - Third World Congress of Structural and Multidisciplinary Optimization*, Buffalo, New York, USA, (to appear).

[146] A. Diaz and O. Sigmund. Checkerboard patterns in layout optimization. *Structural Optimization*, 10:40–45, 1995.

[147] C. Jog and R. B. Haber. Stability of finite element models for distributed-parameter optimization and topology design. *Comp. Meth. Appl. Mech. Engng.*, 130:203–226, 1996.

[148] J. Petersson. A finite element analysis of optimal variable thickness sheets. Mat. Report No. 1996-44, Dept. of Mathematics, Technical University of Denmark, 1996.

[149] R. F. Almgren. An isotropic three-dimensional structure with Poisson's ratio = -1. *J. Elasticity*, 15:427–430, 1985.

[150] M. P. Bendsøe, A. R. Diaz, R. Lipton, and J. E. Taylor. Optimal design of material properties and material distribution for multiple loading conditions. *Int. J. of Num. Meth. Engng.*, 38, 1994.

[151] O. Sigmund. Design of material structures using topology optimization. In N. Olhoff and G. I. N. Rozvany, editors, *WCSMO 1 - First World Congress of Structural and Multidisciplinary Optimization.* Pergamon, Oxford, 1995.

[152] R. V. Kohn and A. V. Cherkaev, editors. *Topics in the Mathematical Modelling of Composite Materials.* Birkhauser, New York, 1997.

[153] G. V. Milton and A. V. Cherkaev. Materials with elastic tensors that range over the entire set compatible with thermodynamics. In *Presented at the joint ASCE-ASME-SES MettN*, University of Virginia, Charlottesville, Virginia, USA, 1993.

[154] G. V. Milton. Composite materials with Poisson's ratio close to -1. *J. Mech. Phys. Solids*, 40(5):1105–1137, 1992.

[155] R. Lakes. Design considerations for materials with negative Poisson's ratio. *J. Mech. Design*, 115:696–700, 1993.

[156] U. D. Larsen, O. Sigmund, and S. Bouwstra. Fabrication of compliant micromechanisms and materials with negative Poisson's ratio. Manuscript, 1995.

[157] O. Sigmund. Topology optimization: a tool for the tailoring of structures and materials. *Philosophical Transactions of the Royal Society: Science into the Next Millennium*, (to appear).

[158] R. H. Burns and F. R. E. Crossley. Kinetostatic synthesis of flexible link mechanisms. *ASME-Paper*, 68(36), 1964.

[159] C. J. Kim, A. P. Pisano, and R. S. Muller. Silicon-processed overhanging microgripper. *Journal of Microelectromechanical Systems*, 1:31–36, 1992.

[160] C. B. W. Pedersen, T. Buhl, and O. Sigmund. Topology synthesis of large-displacement compliant machanisms. In *1999 ASME Design Engineering Technical Conference*, number DETC99/DAC-8554, Las Vegas, November, 1999 (to appear).

SOME ASPECTS OF MATHEMATICAL MODELLING OF CONTACT WITH FRICTION

MARTTI MIKKOLA
Laboratory of Theoretical and Applied Mechanics
Helsinki University of Technology
P.O.Box 1100
FIN-02150 HUT, FINLAND

ABSTRACT

The paper deals with the contact problem with Coulomb friction. In particular, the problem of stability and uniqueness of solutions is considered. The phenomenon of stick-slip is demonstrated using a simple model of mass and springs. A dynamic contact with moving obstacle is also discussed.

1. INTRODUCTION

Contact and friction occur in many mechanical systems. The term 'tribology', in Greek 'τριβοσ λογοσ', is used for the science and technology which deals with the phenomena of surfaces of solid bodies in contact.. It includes besides friction also lubrication, wear and adhesion. In addition to machines, frictional contact is also important e.g. in soil mechanics and seismology. The materials involved are metals, polymers, ceramics, composites, rubber, minerals etc.

Surfaces in contact are more or less rough (asperities of order micrometer) in microscopic view. Friction, i.e. the resistance in sliding, is due to roughness of and also to adhesive forces between atoms on contacting surfaces. Since the highest asperities contact first, the real contact area depends on the normal force between surfaces. Some models of friction try to take into account the real microscopic behaviour of contact. The simplest model of friction is usually credited to Coulomb in 1785 [1], although it was already presented by Amontons in 1699 [2]. It states that the frictional force depends linearly on the normal force between solids in contact. More regular or realistic friction laws have been suggested, for instance the normal compliance law by Oden and Martins [3].

In this paper the conditions of unilateral contact and the formulation of Coulomb friction are first recalled. Then some results by Nguyen and Chateau [4] dealing with the solution

of frictional contact problems are stated, in particular propositions concerning stability and bifurcation. A simple example due to Klarbring [5] is used to demonstrate the phenomena occurring in contact problems.

2. CONDITIONS OF CONTACT

A solid body is in contact with a rigid obstacle which occupies the region $h(\mathbf{x})<0$ (Figure 1). On the part S_u of the boundary of the body the displacements and on the part S_t the tractions are prescribed. On the part S_c the contact of the body and the obstacle can occur. The non-penetration condition implies that

$$h(\mathbf{x}+\mathbf{u}(\mathbf{x})) \geq 0 \quad \forall \mathbf{x} \in S_c \tag{1}$$

The displacement $\mathbf{u}$ and the reaction force $\mathbf{F}$ on the boundary S_c are decomposed into normal and tangent components

$$\begin{aligned} \mathbf{u} &= u_N \mathbf{n} + \mathbf{u}_T \\ \mathbf{F} &= N\mathbf{n} + \mathbf{T} \end{aligned} \tag{2}$$

$\mathbf{n}$ is the outward unit normal of the obstacle at contact point, N the normal reaction, $\mathbf{T}$ the tangent reaction and u_N and $\mathbf{u}_T$ the normal and tangent displacements. The conditions of the unilateral contact are

$$u_N \geq 0, N \geq 0, N u_N = 0 \tag{3}$$

The first condition means that there is no penetration of the body into the obstacle, the second implies that the normal reaction is always compressive, and the third is the complementarity condition indicating that in active contact $N>0$ the normal displacement has to be zero while in separation $u_N > 0$ the reaction must be zero. The classical contact problem for elastic bodies without friction was first formulated by Signorini 1933 along the lines presented above.

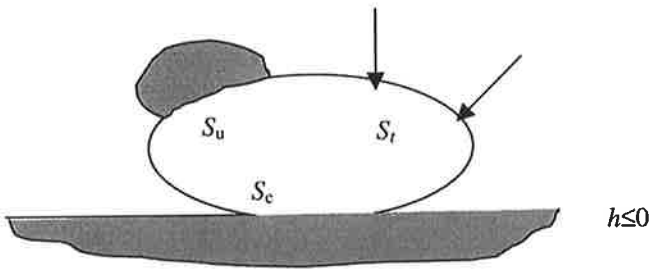


Figure 1. Contact of a deformable body with a rigid obstacle.

3. COULOMB LAW OF DRY FRICTION

At the contact point of the body with the obstacle the relative velocity $\mathbf{v} = \dot{\mathbf{u}}^{body} - \dot{\mathbf{u}}^{obst}$ can also be decomposed into normal and tangential components

$$\mathbf{v} = v_N \mathbf{n} + \mathbf{v}_T \tag{4}$$

Coulomb's law of dry friction can be formulated in several ways (see Nguyen [6]). The traditional way is to write

$$\Phi(\mathbf{T}, N) = |\mathbf{T}| - fN \leq 0 \quad \text{with} \quad \mathbf{v}_T = -\nu \mathbf{T} \quad \text{and} \quad \Phi\nu = 0 \tag{5}$$

f is the friction coefficient. Another way is to introduce the dissipation potential

$$D(\mathbf{v}_T, N) = fN|\mathbf{v}_T| \tag{6}$$

The friction force is then

$$-\mathbf{T} = \frac{\partial D}{\partial \mathbf{v}_T} \tag{7}$$

where the derivative has to be understood as subgradient for the value $\mathbf{v}_T = 0$.

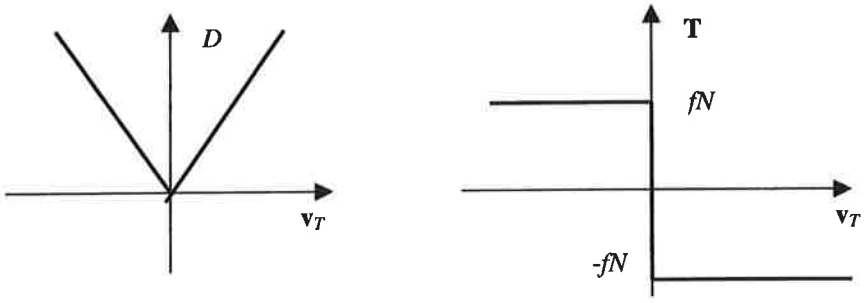


Figure 2. Coulomb friction: dissipation potential and its subdifferential.

4. FORMULATION OF EQUILIBRIUM AND RATE PROBLEMS

The total potential energy of the system considered is

$$E(\mathbf{u}, \lambda) = U(\mathbf{u}) + V(\mathbf{u}, \lambda) \tag{8}$$

where U is the strain energy of the body and V the potential of external forces. λ is the control parameter of implied loads or displacements. Further, the Lagrangian $L(\mathbf{u},N,\lambda)$ is defined

$$L(\mathbf{u},N,\lambda)=E(\mathbf{u},\lambda)-\int_{S_c} Nu_N \mathrm{d}S \tag{9}$$

According to [4] the equilibrium state has to satisfy the following equation

$$L_{,\mathbf{u}}(\mathbf{u},N,\lambda)\bullet\delta\mathbf{u}+\int_{S_c}-\mathbf{T}\bullet\delta\mathbf{u}\mathrm{d}S=0 \quad \forall\delta\mathbf{u}\in U_0 \tag{10}$$

where U_0 is the set of virtual displacements satisfying $\delta\mathbf{u}=0$ on S_u. The normal and tangential reactions on the contact boundary have to satisfy the conditions of Coulomb friction. However, due to the incremental nature of Coulomb's law this equation does not define completely the solution as in the problem of contact without friction (Signorini's problem). The rate problem has to be considered. It is formulated as follows [4]:

The forward rate $\dot{\mathbf{u}}\in V_{1c}=\{\mathbf{v}|\mathbf{v}\in V_1,\mathbf{v}_T \text{ admissible slip}\}$ *is a solution of the variational inequality*

$$(\mathbf{v}-\dot{\mathbf{u}})\bullet(L_{,\mathbf{uu}}\bullet\dot{\mathbf{u}}+L_{,\mathbf{u}\lambda}\dot{\lambda})+\int_{S_c} f(N_{,\mathbf{u}}\bullet\dot{\mathbf{u}}+N_{,\lambda}\dot{\lambda})(|\mathbf{v}_T|-|\dot{\mathbf{u}}_T|)\mathrm{d}S\geq 0 \quad \forall\mathbf{v}\in V_{1c} \tag{11}$$

In the same paper conditions of non-bifurcation and of stability are formulated:

Non-bifurcation: *The present equilibrium is not a critical state of angular bifurcation if the following positivity condition is fulfilled*

$$(\mathbf{v}-\mathbf{v}')\bullet L_{,\mathbf{uu}}\bullet(\mathbf{v}-\mathbf{v}')+\int_{S_c} f(\mathbf{v}-\mathbf{v}')\bullet N_{,\mathbf{u}}\bullet(|\mathbf{v}_T|-|\dot{\mathbf{u}}_T|)\mathrm{d}S>0 \quad \forall\mathbf{v},\mathbf{v}'\in V_{1c} \tag{12}$$

Stability: *The present equilibrium is stable if the following positivity condition is satisfied:*

$$\mathbf{v}\bullet L_{,\mathbf{uu}}\bullet\mathbf{v}+\int_{S_c} f\mathbf{v}\bullet N_{,\mathbf{u}}\bullet|\mathbf{v}_T|\mathrm{d}S>0 \quad \forall\mathbf{v}\in V_{1c} \tag{13}$$

5. STUDY OF A SIMPLE MODEL OF KLARBRING

Consider a simple mechanical model, which has been previously studied by Klarbring [5] and Martins et al.[7]. The model represents a mass supported by springs in unilateral frictional contact with a horizontal rigid base (Figure 3). Klarbring and Martins et al.

investigated the existence and uniqueness of the rate problem. The purpose of this study is to consider the rate problem from the point of view of stability and bifurcation.

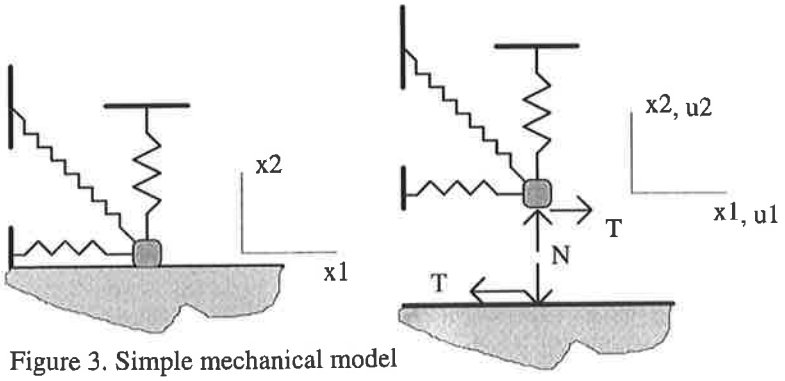


Figure 3. Simple mechanical model

5.1. Definition of the problem

The system has two degrees of freedom. The rigid obstacle is the half-plane $x_2 \leq 0$. The elastic properties of the system are given by the stiffness matrix

$$\mathbf{K} = \begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{pmatrix} \tag{14}$$

Thus, the strain energy is

$$U(\mathbf{u}) = \frac{1}{2}\mathbf{u}^{\mathrm{T}}\mathbf{K}\mathbf{u} \tag{15}$$

where $\mathbf{u}$ is the displacement vector

$$\mathbf{u} = (u_1 \; u_2)^{\mathrm{T}} \tag{16}$$

The total potential energy is

$$E(\mathbf{u}, \lambda) = U(\mathbf{u}) - \lambda \, \mathbf{f}_0^{\mathrm{T}}\mathbf{u} \tag{17}$$

where $\mathbf{f}_0$ is the reference load.

The unilateral contact is defined by conditions

$$u_2 \geq 0, N \geq 0, Nu_2 = 0 \tag{18}$$

N is the reaction perpendicular from the obstacle. The Coulomb law of dry friction is assumed

$$|T| \leq fN \tag{19}$$

The dissipation potential is given by

$$D = -T\dot{u}_1 = fN|\dot{u}_1| \tag{20}$$

The subgradient of D is then

$$D_{,\dot{u}_1} = -T \quad \text{with } |T| < fN \quad \text{for } \dot{u}_1 = 0 \text{ and } T = -fN\,\mathrm{sgn}(\dot{u}_1) \text{ for } \dot{u}_1 \neq 0 \tag{21}$$

Using the Lagrangian, which takes into account the unilateral contact,

$$L(\mathbf{u}, \lambda, \mathrm{N}) = \mathrm{E}(\mathbf{u}, \lambda) - \mathrm{Nu}_2 \tag{22}$$

the equations of the equilibrium can be obtained as follows

$$L_{,\mathbf{u}}(\mathbf{u}, \lambda, \mathrm{N}).\delta\mathbf{u} + \mathrm{D}_{,\dot{\mathbf{u}}}.\delta\mathbf{u} = 0 \tag{23}$$

or written explicitely

$$\begin{aligned} K_{11}u_1 + K_{12}u_2 - \lambda f_{10} - T &= 0 \\ K_{21}u_1 + K_{22}u_2 - \lambda f_{20} - N &= 0 \end{aligned} \tag{24}$$

In solution of these equations, the conditions of unilateral contact (18) and of Coulomb friction (19) have to be taken into account.

5.2. Rate problem

The rate problem follows directly from eq.(22) by differentiation with respect to time

$$L_{,\mathbf{uu}}(\mathbf{u}, \lambda, N)\bullet\mathbf{u} + L_{,\mathbf{u}\lambda}(\mathbf{u}, \lambda, N)\dot{\lambda} + L_{,\mathbf{u}N}(\mathbf{u}, \lambda, N)\dot{N} + D_{,\dot{\mathbf{u}}T}\dot{T} = 0 \tag{25}$$

The rate equations corresponding to eqs.(24) are

$$\begin{aligned} K_{11}\dot{u}_1 + K_{12}\dot{u}_2 - \dot{\lambda} f_{10} - \dot{T} &= 0 \\ K_{21}\dot{u}_1 + K_{22}\dot{u}_2 - \dot{\lambda} f_{20} - \dot{N} &= 0 \end{aligned} \tag{26}$$

As to the unilateral contact and friction conditions, following situations have to be distinguished (superscript e refers to equilibrium solution):

(a) loose contact, no reaction, both slip and separation possible: $u_2^e = N^e = T^e = 0$. Then $\dot{u}_2 \geq 0,\ \dot{N} \geq 0,\ \dot{N}\dot{u}_2 = 0,\ |\dot{T}| \leq f\dot{N}$ for $|\dot{u}_1| = 0,\ \dot{T} = -f\dot{N}\,\mathrm{sgn}(\dot{u}_1)$ for $|\dot{u}_1| \neq 0$

(b) active contact with possible slip: $u_2^e = 0, N^e > 0, |T^e| = fN^e$. Then

$\dot{u}_2 = 0; \dot{u}_1 = -\nu\,\mathrm{sgn}(T^e), \nu \geq 0, \dot{T} \leq f\dot{N}\,\mathrm{sgn}(T^e)$ for $\dot{u}_1 = 0, \dot{T} = f\dot{N}\,\mathrm{sgn}(T^e)$ for $\dot{u}_1 \neq 0$

(c) active contact with stick: $u_2^e = 0, N^e > 0, |T^e| < fN^e$. Then $\dot{u}_2 = \dot{u}_1 = 0$.

(d) no contact: $u_2^e > 0,\ N^e = T^e = 0$.

5.3. Stability and bifurcation of solutions of the rate problem

Application of criteria of non-bifurcation eq.(12) and of stability eq. (13) for the buckling of elastic structures in unilateral contact with friction under constant load to the present case leads to the conditions

$$K_{11} - f|K_{21}| > 0 \text{ in case (a) } N = 0 \text{ and } K_{11} - \mathrm{sgn}(T) f K_{12} > 0 \text{ in case (b) } N > 0 \tag{27}$$

for both non-bifurcation and stability. The following analysis shows that if these criteria are satisfied, the rate problem has a unique solution for all possible load directions. If they are not satisfied, there can be stable bifurcated solutions for certain load directions in case of zero reaction (case (a)) and unstable solutions in case of non-zero reaction (case (b)).

Case (a): If $\dot{u}_2 > 0$, we have $\dot{N} = \dot{T} = 0$, and the solution represents the separation

$$\begin{pmatrix} \dot{u}_1 \\ \dot{u}_2 \end{pmatrix} = \dot{\lambda}(Det(K))^{-1} \begin{pmatrix} K_{22} & -K_{21} \\ -K_{12} & K_{11} \end{pmatrix} \begin{pmatrix} f_{10} \\ f_{20} \end{pmatrix} \tag{28}$$

The condition $-K_{12}f_{10} + K_{11}f_{20} > 0$ corresponding to $\dot{u}_2 > 0$ gives the load directions for which the separation can occur.

If $\dot{u}_2 = 0$, contact prevails and

$$\begin{aligned} \dot{N} &= K_{21}\dot{u}_1 - \dot{\lambda} f_{20} \geq 0 \\ \dot{T} &= K_{11}\dot{u}_1 - \dot{\lambda} f_{10} \end{aligned} \tag{29}$$

Assuming that slip occurs leads to equation

$$(K_{11} + \mathrm{sgn}(\dot{u}_1) f K_{21})\dot{u}_1 = \dot{\lambda}(f_{10} + \mathrm{sgn}(\dot{u}_1) f f_{20}) \tag{30}$$

If the uniqueness criterion (27) is satisfied, eq.(30) gives well-defined solutions for appropriate load directions, i.e.

$$\dot{u}_1 > 0 \text{ for } f_{10} + f f_{20} > 0 \text{ and } \dot{u}_1 < 0 \text{ for } f_{10} - f f_{20} < 0 \tag{31}$$

For $f_{10}+ff_{20}<0$ and $f_{10}-ff_{20}>0$ stick takes place, i.e. $\dot{u}_1=0$. The rate of load $\dot{\lambda}$ is, of course, positive. These results can be conveniently presented for different load directions as shown in Figure 4. It corresponds exactly the previous results [5] and [7].

Next, the case is considered in which the uniqueness criterion (27) is not satisfied. We assume that $K_{21}>0$ and $fK_{21}\geq K_{11}$. (The case $K_{21}<0, f|K_{21}|>K_{11}$ can be handled in analogous way). For positive slip the solution is unique as can be seen from eqs.(30) and (31). Uniqueness is lost for negative slip. Let us take first the case $fK_{21}>K_{11}$. Eq.(30) yields

$$\dot{u}_1=\dot{\lambda}(K_{11}-fK_{21})^{-1}(f_{10}-ff_{20}) \tag{32}$$

which leads to the rate of reaction

$$\dot{N}=\dot{\lambda}(K_{11}-fK_{21})^{-1}(K_{21}f_{10}-K_{11}f_{20}) \tag{33}$$

This solution is valid for $f_{10}-ff_{20}>0$, $K_{21}f_{10}-K_{11}f_{20}<0$, $\dot{\lambda}>0$. But the first inequality indicates that the stick solution is possible, too, and the second means that the separation can also occur. This result, illustrated by Figure 5, coincides with the previous findings [5] and [7].

In the case $fK_{21}=K_{11}$ eq.(30) indicates that the right hand side must be $\dot{\lambda}(f_{10}-ff_{20})=0$ and that the slip is indetermined (but negative). Assumption of $\dot{\lambda}=0$ and $\dot{u}_1<0$ leads to negative rate of reaction, which is not possible.The possibility of $\dot{u}_1=0$ still exists. The study of second order equations (which is not included here) leads to the result $\ddot{\lambda}=\ddot{u}_1=0$. Thus, without going to higher order equations it is concluded that for $f_{10}-ff_{20}<0$ only the separation solution is valid, while for $f_{10}-ff_{20}>0$ the stick solution remains. In case $f_{10}-ff_{20}=0$, second order rate equations lead to the conclusion is that $\dot{u}_1=0$, i.e. stick occurs. The load rate remains indetermined. The result is shown in Figure 6.

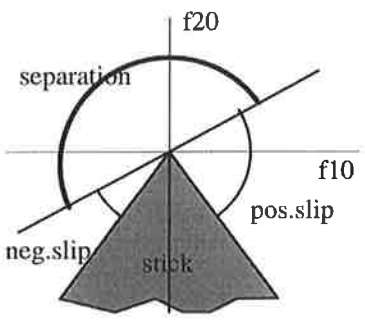


Figure 4. Regular solutions for N=0.

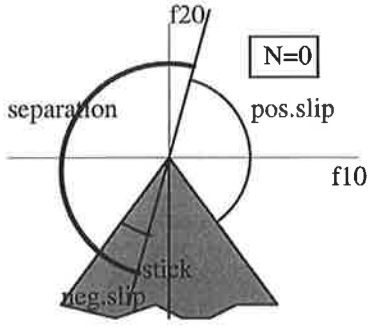


Figure 5. Solutions for case $fK_{21}>K_{11}$ with N=0.

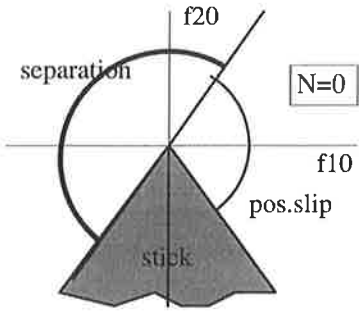


Figure 6. Solutions for case $fK_{21}=K_{11}$ with N=0.

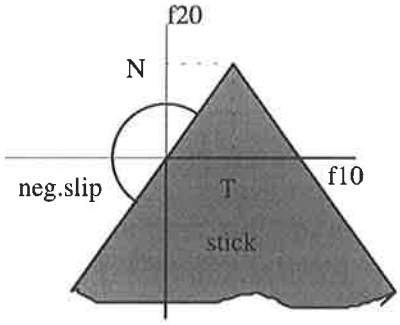


Figure 8. Regular solutions for N>0.

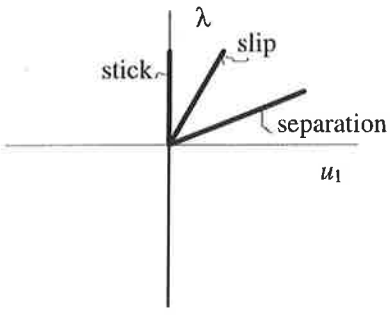


Figure 7. Bifurcation solutions.

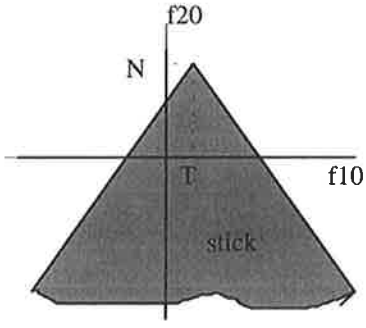


Figure 9. Situation for N>0, $T<fN$.

The nonunique solutions for the case $fK_{21} > K_{11}$ found above are all stable in the sense of Hill, since positive work must be injected to the system in order to perturb it. As a matter of fact, the stability is already indicated by the positivity of $\dot{\lambda}$, which means that a solution exists. Indeed, computing the mechanical work required to perturb the system in the case of three possible solutions, i.e. for load $\mathbf{f}(\tau)=\dot{\lambda}\tau\mathbf{f_0}$, $f_{10}=af_{20}<0$, $K_{11}/K_{21}<a<f$ and displacement $\mathbf{u}(\tau)=\tau\dot{\mathbf{u}}$, $\tau\in[0,t]$ yields:

$$W_d(t)=\int_0^t \mathbf{f}(\tau).\dot{\mathbf{u}}d\tau = E(\mathbf{u}(t),\lambda(t))+\int_0^t fN(\tau)|\dot{u}_1|d\tau \qquad (34)$$

The amounts of work are

for separation: $W_{d,se}=(\dot{\lambda}t)^2(\mathrm{Det}(K))^{-1}(K_{22}a^2-2K_{21}a+K_{11})f_{20}{}^2/2$

for slip: $W_{d,sl}=(\dot{\lambda}t)^2(K_{11}-fK_{21})^{-1}(a-f)af_{20}{}^2/2$ (35)

for stick: $W_{d,st}=0$

It can be seen that $W_{d,se} > W_{d,sl} > W_{d,st}$, so that separation is the mode of bifurcation to be expected to occur. The slopes of the load-displacement curves are schematically shown in Figure 7.

Case (b): Eqs. (29) and (30) are still valid, except that the rate of reaction can be also negative. The solution is unique, if the uniqueness criterion (27) holds. To be specific, assume that $T^e > 0$. Hence, the possible slip is in negative direction. The solution is

$$\dot{u}_1 = 0 \text{ for } f_{10} - ff_{20} > 0,\ \dot{\lambda} > 0$$
$$\dot{u}_1 < 0 \text{ for } f_{10} - ff_{20} < 0,\ \dot{\lambda} > 0 \qquad (36)$$

The result can again be illustrated in the load space as shown in Figure 8.

Consider then the case in which the uniqueness criterion (27) does not hold, $fK_{21} > K_{11}$. As before, $T^e > 0$ and slip occurs in negative direction. For $f_{10} - ff_{20} > 0,\ \dot{\lambda} > 0$ the stick solution $\dot{u}_1 = 0$ is in force. Negative slip occurs for $f_{10} - ff_{20} < 0$, but the rate of load has to be negative $\dot{\lambda} < 0$. This indicates unstable behaviour. If the amount of mechanical work is again computed for the load defined by $f_{10} = af_{20} < 0,\ a > f$ the expression $(35)_2$ is obtained, but now it is negative. Physically this result means that the strain energy capacity of the system is inadequate to absorb the work done by external load when at the same time the reaction and the frictional force decrease due to relieving influence of the system. For constant load this would mean snap-through phenomenon, i.e. dynamic behaviour.

Still, the case $fK_{21} = K_{11}$ is to be investigated. As in case (a), using the second order rate equations one can conclude that for $f_{10} - ff_{20} = 0$ stick occurs. For $\dot{\lambda}$ equal to zero the relationship between load and displacement increments is quadratic like in limit point behaviour of stability problems. The study by use of second order equations reveals that the slip is unstable for $f_{10} - ff_{20} < 0$.

Case (c): The situation can be conveniently illustrated in load space, Figure 9. For proportional (radial) loading paths emanating from origin, the stick state prevails as long as the load point remains inside the cone. The slip can occur when the load point reaches the boundary, in stable, unique manner if the uniqueness criterion (27) is satisfied and in unstable manner if it is not satisfied.

5.4. Concluding remarks

The results obtained above considering the simple mechanical model suggest that in unilateral contact problems distinction should be made between the stability of equilibrium state and of equilibrium path (process). It has been shown that stable loading paths exist although the criterion of stability of equilibrium (under constant load) proposed in [4] is violated. Further, in solution of the rate problem, the dependence of the reaction on the load has to be taken into account so that the Proposition 8 of Ref.[4] should read:

The forward derivative $\dot{\mathbf{u}}$ *is a solution of the variational inequality*:

$$\dot{\mathbf{u}} \in V_{1c} = \left\{ \mathbf{v} \middle| \mathbf{v} \in V_1,\ \mathbf{v}_T \in N_{C(m)}(-\mathbf{T}) \right\} \text{ and satisfies}$$

$$(\mathbf{v}-\dot{\mathbf{u}}).(L_{,uu}.\dot{\mathbf{u}} + L_{,u\lambda}.\dot{\lambda}) + \int_{S_c} (m_{,u}.\dot{\mathbf{u}} + m_{,\lambda}.\dot{\lambda}) \mathrm{tg}\varphi(|v_T| - |\dot{u}_T|) da \geq 0 \quad \forall \mathbf{v} \in V_{1c} \tag{37}$$

where the dependence $m(\mathbf{u}, \lambda)$ is given by eq.(24) and S_c denotes the portion of S with effective contact (h=0).

7. DYNAMIC CONTACT WITH MOVING OBSTACLE

Consider the same simple model as previously but now in contact with a moving obstacle. Thus, the model consists of a mass m supported by springs in unilateral frictional contact with a horizontal rigid half-plane $x_2 \leq 0$ moving in the direction of negative x_1-axis.

The system has two degrees of freedom. The mass is initially resting on the plane so that the initial conditions are

$$u_1(0) = u_2(0) = 0, \dot{u}_1(0) = -v_0, \dot{u}_2(0) = 0 \tag{38}$$

The equations of motion are

$$\begin{aligned} m\ddot{u}_1 &= -K_{11}u_1 - K_{12}u_2 + T \\ m\ddot{u}_2 &= -K_{21}u_1 - K_{22}u_2 - mg + N \end{aligned} \tag{39}$$

In addition, the conditions of unilateral contact (18) and of Coulomb friction (19) have to be taken into account, noting that in this case the direction of the friction force depends on the relative velocity $\dot{u}_1 - (-v_0)$.

As long as the contact prevails, the solution is $u_1 = -v_0 t, u_2 = 0$. Slip starts when

$$T = -K_{11}u_1 = -fN = -f(K_{21}u_1 + mg) \tag{40}$$

which yields

$$t = t_1 = \frac{fmg}{(K_{11} + fK_{21})v_0} \tag{41}$$

provided that $K_{11} + fK_{21} > 0$. In next stage slip continues with initial conditions $u_1(0) = -v_0 t_1, u_2(0) = 0, \dot{u}_1(0) = -v_0, \dot{u}_2(0) = 0$ taking t'=t-t_1. The solution is

$$u_1(t') = -\frac{v_0}{\omega}\sin\omega t' \quad \text{with } \omega = \sqrt{(K_{11} + fK_{21})/m} \tag{42}$$

which is valid as long as the relative velocity $\dot{u}_1 + v_0 = v_0(1-\cos\omega t')$ remains positive. This is true if the coefficient of friction is smaller than a critical value

$$f < f_c = \frac{K_{11}}{K_{21}}\left(\frac{mg^2K_{11}}{v_0^2K_{21}^2}-1\right) \quad \text{for } K_{21}>0 \tag{43}$$

or larger than a critical value

$$f > f_c = \frac{K_{11}}{|K_{21}|}\left(1-\frac{mg^2K_{11}}{v_0^2K_{21}^2}\right) \quad \text{for } K_{21}<0 \tag{44}$$

Indeed, looking at the value of normal reaction

$$N = K_{21}u_1 + mg = \frac{K_{11}g}{\omega^2} - \frac{K_{21}v_0}{\omega}\sin\omega t' \tag{45}$$

and its minimum value

$$N_{\min} = \frac{K_{11}g}{\omega^2} - \frac{|K_{21}|v_0}{\omega} \geq 0 \tag{46}$$

which should be positive for active contact, the critical values of friction can be found. If the friction coefficient is larger (case $K_{21}>0$) (or smaller, case $K_{21}<0$) than the critical value, the normal reaction tends to zero which eventually means separation of the mass from the obstacle, i.e. the mass will bounce at the instant $t_2' = \arcsin(gK_{11}/\omega v_o K_{21})$. The initial conditions are (taking $t'' = t - t_1 - t_2'$)

$$u_1(0) = -v_0t_1 - \frac{v_0}{\omega}\sin\omega t_2' = -\frac{mg}{K_{21}}, u_2(0) = 0, \dot{u}_1(0) = -v_0\cos\omega t_2', \dot{u}_2(0) = 0 \tag{47}$$

Equations of motion (39) with T and N equal to zero are now valid. After the jump the mass will again become in contact with the obstacle and bounce a second time with conditions depending on the properties of impact.

In case $K_{11} + fK_{21} < 0$ the mass will stick to the obstacle. Eqs.(39) yield

$$T = K_{11}u_1 = -K_{11}v_0t, N = mg + K_{21}u_2 = mg - K_{21}v_0t \tag{48}$$

and $-T<fN$ for all $t>0$. Large displacements should be considered to solve the problem.

8. CONCLUSIONS

The problem of frictional contact of solids includes many interesting and complicated features. Complications are partly due to the incremental nature of Coulomb's friction law. Some of the features could be demonstrated using a simple mechanical model. Computational procedure for systems of multiple degree of freedom would involve rather intriguous bookkeeping of nodes in contact and slipping or sticking. Also, the dynamic phenomena, like bouncing and wave propagation, occurring at slip would represent topics of further research.

REFERENCES

1. Coulomb C.A., Théorie de machines simples. Mem. de Mathématique et de Physique de l'Acad. Royale, 10, p. 161-342, 1785.
2. Amontons G., Sur l'origine de la résistance dans les machines. Mem. Acad.Royale, p.206-222, 1699.
3. Oden J.T. and Martins J.A.C., Models and computational methods for dynamic friction phenomena. Comp.Meth.Appl.Mech.Eng. 52(1985) 527-634:
4. Klarbring A., (Non-)uniqueness and (non-)existence of frictional rate problems. Euromech Colloquium 273 on Unilateral Contact and Dry Friction, La Grande Motte, France, May 1990. P.89-92.
5. Martins J.A.C., Monteiro Marques, Manuel D.P., and Gastaldi Fabio, On an example of non-existence of solution to a quasistatic contact problem with friction. Relatório CMEST AI 1/92, Lisboa, Fevereiro de 1992. 98 p.
6. Chateau X. and Nguyen Q.S., Buckling of elastic structures in unilateral contact with or without friction. Eur.J.Mech.,A/Solids,**10**, no1, 71-89, 1991.
7. Nguyen Q.S., Stability, Bifurcation and Nonlinear Analysis in Solids. Manuscript, to be published by J.Wiley&Sons, Inc.

WOODEN DOME MADE OF LAMINATED BIRCH PLYWOOD

Ari Aalto, Tampere University of Technology, Structrural Mechanics, Tampere Finland
Markku Heinisuo, KPM-Engineering Ltd., Tampere Finland
Mauri Mäki-Marttunen, Architect company Mauri Mäki-Marttunen, Helsinki Finland
Markku Tuomala, University of Technology, Structrural Mechanics, Tampere Finland

ABSTRACT

In this paper the development of a wooden glued spherical dome is described. The dome itself is constructed from 12 prefabricated elements or segments. In the final form the dome consist of a net of longitudinal and circumferential arches made of birch plywood, so that the circumferential arches are joined on top of the longitudinal ones. A plywood cover is nailed to the circumferential arches, and the roof is finished with a mineral wool layer and coating. The functional goals of the dome were a good architectural impression by using the supporting system as an integral part of the restaurant interior and also good acoustic characteristics.

INTRODUCTION

The foundation of the Linnanmäki Amusement Park wanted to establish a fast food restaurant for children and adults alike, the form of which would reflect the shape of one the most popular products on their menu. An architect company, Mauri Mäki-Marttunen, came up with an idea of a round two-storey wooden building with a diameter of about 17 m, and which is reminiscent of - a hamburger. It was decided that the restaurant hall on the second floor should be entirely an open space and that the form of the ceiling should be convex outwards. In this context an idea emerged that the roof could be a dome with the load bearing members being left exposed. The restaurant building was constructed during the winter 1999 and it was opened for the public at the start of the season on the first of May.

THE STRUCTURE

The dome as a structural form is not a new invention, but from an architectural point of view its strictly symmetric geometry is worth exploring.

The objectives of the development work in the project were to minimize the usage of material and the structural dimensions by utilizing the excellent strength properties of birch and that all load bearing members would be left exposed. In addition the structure wanted to be constructed of ready-made elements.

The roof of the building is a spherical wooden dome consisting of a net of longitudinal (meridional) arches and circumferential rings both made of laminated birch plywood. The circumferential rings are joined by gluing on top of the meridional arches. A plywood cover with a thickness 12 mm is nailed to the rings, and the roof is finished with a mineral wool layer and coating (Fig. 1)

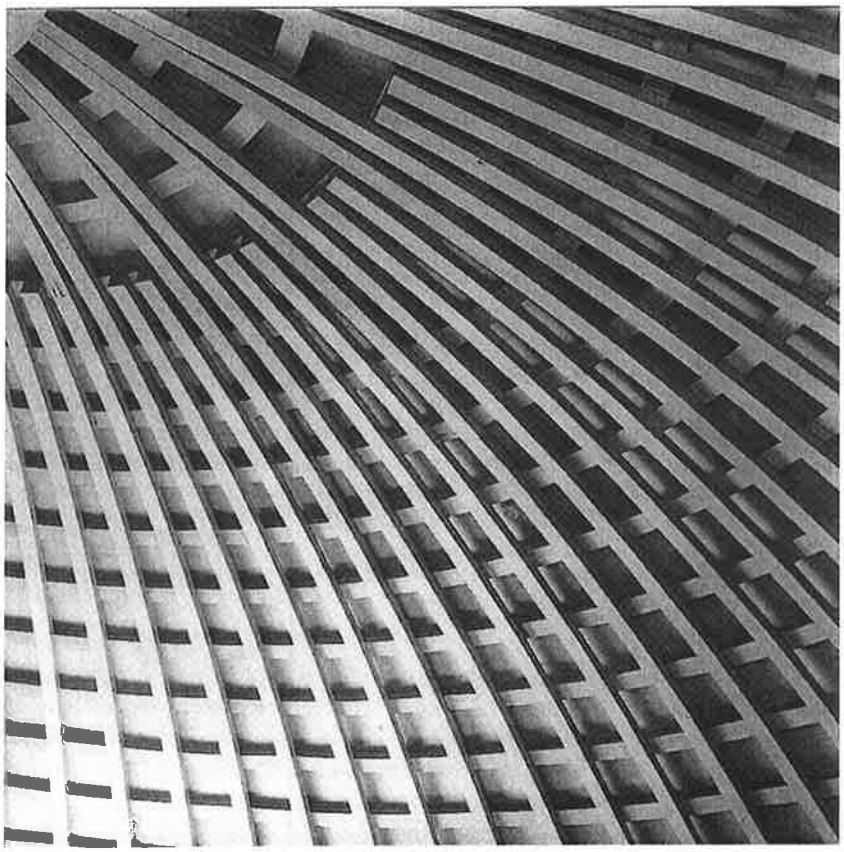


Fig. 1: Inner view of the dome.

The dome net is supported by a central ring with a diameter of 2 m made of laminated plywood and by a boundary ring with a diameter of 17 m made of glued laminated wood elements. The outer ring is supported by 11 stocky round wooden pillars. This type of dome roof can be constructed of prefabricated elements which are assembled on the site. In this case the number of elements is 22.

The meridional arches and segments of the circumferential rings are manufactured by gluing grain-direction oriented birch plywood sheets, with a thickness of 6.5 mm and cut to board width strips, to form rails. In one plywood sheet the grains of four veneer layers were axial oriented when the middle layer was transversely oriented. In this way manufactured

the small shear modulus of plywood can be increased enough. The total height of the arch and ring cross section is 80 mm and the width varies between 120 and 140 mm. The orthogonal radial rib net of arches and rings with the covering plywood sheets forms the inner surface of the ceiling.

The assembled dome consists of 44 full length (running from the lower supporting ring to the upper one) and of 44 three-quarter-length arches and of 14 rings. The diameter of the dome is 17 m and it is three meters high. The radius of curvature of the dome is 13 m. The total thickness of the load bearing structure is 160 mm.

The glued joint between the longitudinal arches and the circumferential rings was tested under cyclically varying moisture conditions by applying a standard ASTM D 3434. The glue type used in the construction was selected by comparing the behavior of four different glues.

ACOUSTICS

An aim of the development work was to use the waffle like shell inner face as an acoustic surface. On the other hand, the net like structure enables an easy addition of sound absorbing material.

The measured reverberation times are appropriate for the needs. With furniture the reverberation time is two seconds according to the measurements, and it is estimated to shorten to about 1.2 seconds at actual use.

ASSEMBLY

Each segment of the dome consists of two full length boundary arches and of two shorter intermediate arches. The parts of the circular rings belonging to one element are joined on top of the arches. The average distance between the rings is 560 mm and the distance between the arches at the boundary supporting ring is 590 mm (Fig. 2).

Due to the doubly curved structure of the arch and ring network the geometry, especially with respect to the rings, is demanding. Because of the relatively small size of the dome it could be assembled of similar ready-made segments running from the lower boundary ring to the upper one surrounding a light opening on top of the dome. The elements were joined together so that the steel parts of the joints were hidden.

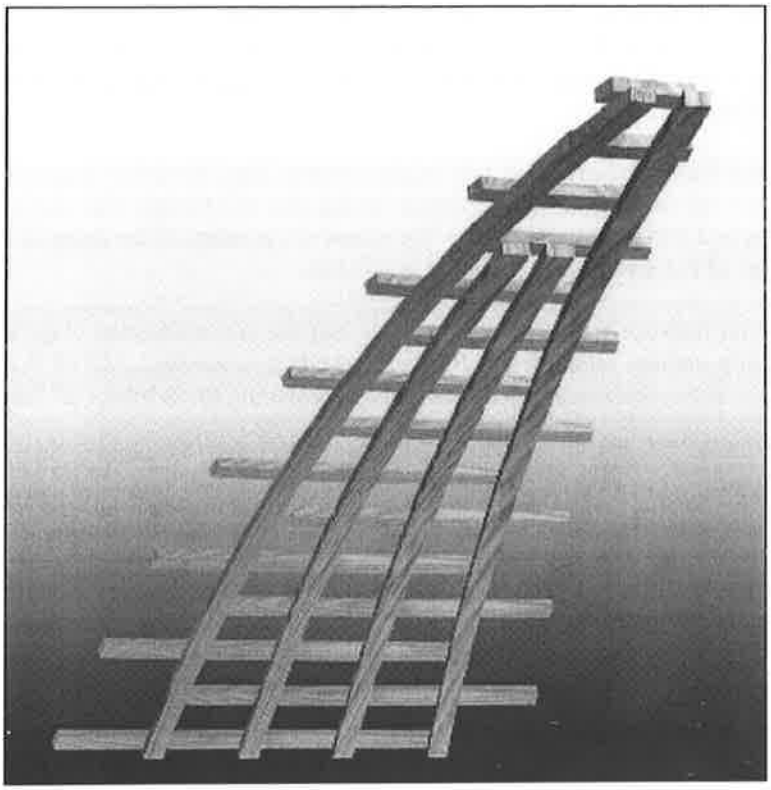


Fig. 2: Prefabricated dome segment.

SPHERICAL DOME, AN EFFICIENT AND A SIMPLE STRUCTURAL FORM

Shells are slender and effective load bearing structures. Their aesthetic character is based on the structural form. In the present case the load bearing structure is an integral part of the restaurant interior. A dome roof as a doubly curved shell is a very efficient load bearing structure which can be made thin.

The spherical shape has an advantage over other doubly curved shell forms that it is in the bending state less sensitive to the spatial variation of snow and wind loads. Besides, the spherical structural form is simple.

The most common shell roof structural type consists of arches and rings (radial rib dome), although the greatest spherical shell roofs are geodetic domes consisting of equilateral triangles of beams. If the diameter of the dome is small enough, then the arches can be made continuous between the central and boundary supporting rings and the arches and rings can be placed on their own levels (surfaces).

The radial rib dome can be constructed by prefabricated elements. In the geodetic dome the bars are on the same level with numerous joints and prefabricated segments cannot be used.

In both structural systems the cover of the roof distributes the load to the arches or beams and it is not fully utilized as a load bearing part of the structure.

The efficiency of shells is based on the developing membrane state of stresses. However, with respect to the loss of stability many shell forms are sensitive to structural imperfections and the post buckling load bearing capacity may be much less than the linear buckling load estimate implies. Therefore, in order to assess reliably the behavior of shells a geometrically nonlinear analysis method with the ability to solve nonlinear eigenvalue problems and to follow the nonlinear load-deflection paths in the post critical state is required.

If the loading creates only small bending moments in the shell, then the imperfection in the nonlinear analysis can be given as a linear combination of the lowest buckling modes. On the other hand, if the loading gives rise only to a local bending moment field, then the form of initial imperfection is not so important.

In Finland the own weight of the shell roof is small when compared with the snow load. The membrane forces of a spherical dome due to a uniform snow load are related to the load value and to the radius of curvature. Therefore a rule is obtained: *If the dimensions of the arches and rings forming the spherical dome and their spacing is increased by the same scale factor as the diameter and height then an equivalent dome is obtained.*

SNOW LOAD

Various design specifications give slightly variable assumptions for the drifting of snow. The codes generally assume that snow is drifting on the leeward side of the prevalent wind direction, but they differ with respect to the shape of the snow drift and with respect to the snow load intensity.

Because for domes no rules for snow drifting distribution are given the designers customarily assume that the snow load distribution is similar in shape to the one given for singly curved or cylindrical roofs and the snow load distribution in the circumferential direction is based merely on taste.

In Finland the snow load distribution and density on a spherical dome, built in a high snow load area, was monitored during several years. The measured results were in agreement with a meridional distribution given in a US-code from the year 1982. According to observations the snow load distribution on the central dome surface is quite uniform over one third of the roof and descends to zero rather steeply at the boundary zone (spanning an angle of 15 degrees in the meridional direction).

The snow load distribution of the present dome is based on these observations, and it is given in Fig. 3.

MATERIAL PROPERTIES AND LOAD BEARING CAPACITY

An example dome with a diameter of 28 m and with a curvature radius of 19.8 m was used for studying the effects of material properties on the load bearing capacity. The shell roof was analysed as geometrically nonlinear by the finite element code ABAQUS. The number of meridional and circumferential arches were 128 and 29, respectively, and the arches were located on the same level. Each side of the curved arch lattice was modelled by one 3D-beam element. The used beam elements of type B31 included the effects of transverse shear deformation. Shell elements of type S4R5 with negligible bending stiffness were utilized in distributing the load to the beam elements.

The arches were rigidly joined to the supporting rings. The modulus of elasticity in the axial direction of the beam elements was 5985 MPa, and for the shear modulus the values 95, 190, 380 and 3800 MPa were adopted. The snow load distribution was assumed in accordance with Fig. 3. It corresponds to the snow load in central Finland.

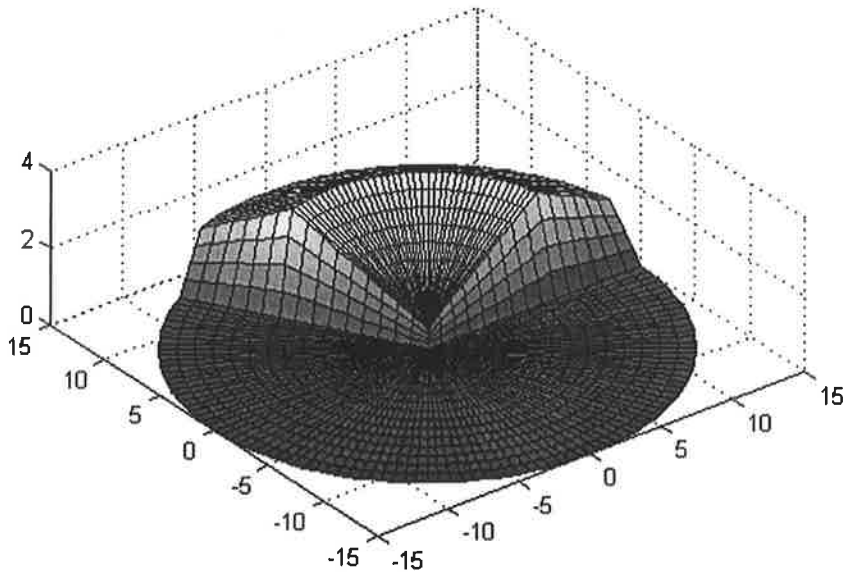


Fig. 3: Snow load distribution.

The results were rather surprising. The shear stiffness of the bar elements had a great influence on the limit load, i.e. the load value at which the load parameter starts to diminish on the load deflection curve in a path following analysis. Also the buckling mode shape changed with shear modulus.

The load bearing capacity of the model with the smallest shear modulus was only about one seventh of the corresponding value obtained by the largest shear modulus by which value the shear deformation is negligible as in the Euler-Bernoulli theory.

In Fig. 4 the transverse displacement of node 521 is depicted as a function of load factor from various analyses.

As a reference structure a geodetic dome built in Oulu was considered. The Oulu-dome has a diameter of 115 m and a height of 21 m. The snow load was assumed in accordance with the Linnanmäki-dome.

Material values of Eurocode for a geometrically nonlinear case were adopted. By using shear deformable 3D-beam elements for modelling of the dome, in linear analysis the compression strength of the beams was exceeded. In a geometrically nonlinear analysis the load bearing capacity of the dome was 0.70 times the given load.

By using a high value for the shear modulus in the transverse direction, corresponding to the Euler-Bernoulli model, in a linear analysis the compression strength of the beams was still exceeded, and in a geometrically nonlinear analysis the ultimate load factor was 0.67.

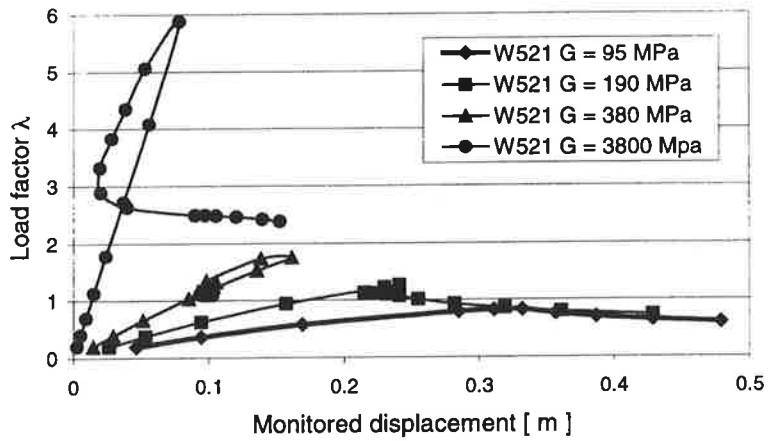


Fig. 4: Influence of shear modulus G on the load bearing capacity of an example dome.

EXPERIMENTAL LOADING OF THE CONSTRUCTED DOME

The stiffnesses of the constructed dome and of one of the prefabricated segments were studied experimentally. The bending stiffness of the shell roof was tested by applying a central point load on an element at the carpentry shop. The tested element was lying on roller supports.

The segment was analysed also numerically by ABAQUS. The arches were modelled by beam elements of type B31 and the covering plywood deck by shell elements of type S4R5. The nail connections between the deck and the rings and the glued connections secured with bolts were modelled by short beam elements.

The material properties are given in a coordinate system (x,y,z) where x is the beam axis and z points to the direction of the shell normal. The following mean values based on Eurocode, EC5, for the moduli of elasticity, shear moduli and Poisson's ratios were adopted:

$$E_x = 13280 \text{ MPa}, \quad G_{xy} = 600 \text{ MPa}, \quad \nu_{xy} = 0.2,$$
$$E_y = 1080 \text{ MPa}, \quad G_{yz} = 600 \text{ MPa}, \quad \nu_{yz} = 0.2,$$
$$E_z = 3320 \text{ MPa}, \quad G_{xz} = 600 \text{ MPa}, \quad \nu_{xz} = 0.2.$$

For the plywood deck $E = 8300$ MPa and $\nu = 0.1$. The stiffness properties of the connecting elements were calculated also in accordance with EC5 design specification.

During the test the horizontal displacement of the narrow end of the segment and the deflections of the rings 4, 7, 9 and 11, counted beginning from the top of the segment, were monitored. The point load was applied at ring 10.

Parametric studies by varying the degree of fixity at the segment boundaries were also made. The overall correspondence between the measurements and the analytical results was good.

The constructed dome was loaded incrementally by lifting concrete blocks on top of the supporting upper ring surrounding the light opening. The total load summed up to 105.3 kN. This kind of symmetrical load is mainly carried by membrane forces and due to the high membrane stiffness of the dome the maximum deflection will, in this case, be of the order of one centimetre only.

The complete shell roof was analysed also by ABAQUS. In the finite element model the arches and rings were located on their own levels or surfaces. Each side of the orthogonal arch-ring net was modelled by one shear deformable 3D-beam element of type B31, (two noded Timoshenko beam element). The plywood deck was discretized by shell elements of type S4R5. At the joints between the arches and rings small elements modelling the glued joints were placed.

The same material properties as given above for the model of a segment were used for the roof model, too. The lower supporting ring is made of glued laminated wood of grade GL32 and its modulus of elasticity and shear modulus are E_x=14500 MPa and G_{xz}=840 MPa, respectively.

In modelling of the joints between the arches and the upper ring short beam elements of type B33 were applied. Due to geometrical discrepancies some filling wooden blocks had to be used and their effect was taken into account by reducing the axial stiffness of the parts of the arches between the last ring and the supporting ring at the lower end of the shell.

The displacement measurements were carried out by a theodolite. The accuracy of the measurements, or maximum error is 0.1 mm. Nine measurement points equipped with a prism were selected on two arches on the opposite sides of the dome. Three prisms were attached to the upper supporting ring for deflection measurements.

Five different analyses for the shell roof were carried out. In the first three the boundary conditions were varied at the lower and upper ends. For the fixed-fixed, fixed-pinned and pinned-pinned boundary conditions the maximum displacements were -7.15 mm, -8.87 mm and -8.94 mm, respectively.

Next, the analysis model was modified to correspond to the built dome. The joint at the upper end was modelled by a bolted connection and the effect of the filling wooden blocks at the lower end were taken into account by reducing the axial stiffness of the arch parts near the lower support as described above.

For central loading the downward displacement of the upper ring was 11.38 mm. In the analysis case 5 an eccentric loading with an eccentricity of 300 mm was applied. In this case the maximum and minimum deflections were 6.76 mm and -36.72 mm, respectively.

In Fig. 5 the calculated displacements at measurement points are compared with the experimentally observed values. The correspondence is quite good, and it may be concluded that the material stiffness values given in EC5 are appropriate for the design purposes of this kind of structure.

CONCLUSION

A novel type of wooden dome which fulfils the structural, architectural and other goals and needs set by the client has been developed and constructured. Similar structural systems, especially the utilization of prefabricated elements, for larger spans will be studied.

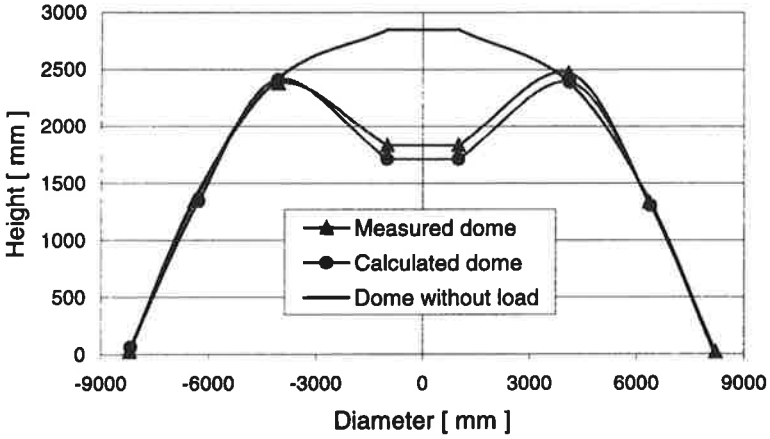


Fig. 5: Comparison on measured and calculated deflected shapes of the dome at Linnanmäki, F = 105 kN (deflections are multiplied by 100).

NEW STRUCTURAL ANALYSIS MODEL FOR NAIL PLATE STRUCTURES IN AN INTEGRATED CAD SYSTEM

M. HEINISUO, R. KYTÖMÄKI, M. LAURILA
KPM-Engineering Oy
Yliopistonkatu 60 C
FIN-33100 Tampere, FINLAND

ABSTRACT

The paper deals with a new structural analysis model implemented to an integrated CAD system for the design of wooden nail plate structures. A new analysis model follows strictly the new Finnish design codes for these structures. The new analysis model includes very careful consideration of the geometrical model of joints and their analysis models. In practice this means that the analysis model will include much more finite elements than that used before. The analysis model is linear with the respect of geometrical and material entities. The local joint flexibilities are taken into account in the model. Some comparisons of the results given by the new structural model and the old one, used over 15 years in the program, will be given in the paper. The old model was quite simplified compared to the new model and it is believed that the new model will give more economical designs than the old system.

1. INTRODUCTION

Over 50% of the sorted timber in Finland is used to fabricate nail plate trusses. Due to developed design and fabrication methods nail plate structures include into the main systems to build roofs to detached houses. Nail plate structures are used for roof structures in blocks of flats, too. Other applications are scaffoldings of concrete cast for bridges etc. It can be said, that the nail plate structure business is large in the field of timber engineering.

The CAD systems used for the design of nail plate structures are highly integrated involving the integration of geometrical model and structural analysis model and strength check of wooden parts and joints. The information of the designed nail plate structure is transferred forward to fabrication systems, such as minimising the loss of timber, material control databases, control of cutting machines and other applications in the production environment such e.g. as a database application for reuse of designs. The last application is very important when offering the structures. One factory can make many thousands new structures in a year and then it is good to have some database to seek for old designs to make a good offer. In KPM-Engineering Oy about 15000 new designs are made in every year. Some designs are done for foreign fabricators and part of production is going abroad.

When writing a program for the design of this kind of structures there are many principles, which must be kept in mind. When considering the analysis of these structures then one principle is very important:

- Program must start from geometrical modelling, after that follows the structural analysis and strength checks and other tasks (production of fabrication data etc)

This principle is of the main importance when creating the real working software for the design purposes. The data transfer between tasks (such as geometrical and structural modelling) is rather easy to perform in the software under consideration, because all the tasks are programmed to the same program. The starting point of the design process must be the geometrical model in order to get a reasonable analysis model especially for the joints. This can be seen in Fig. 1, which illustrates geometrical and structural models for one nail plate joint.

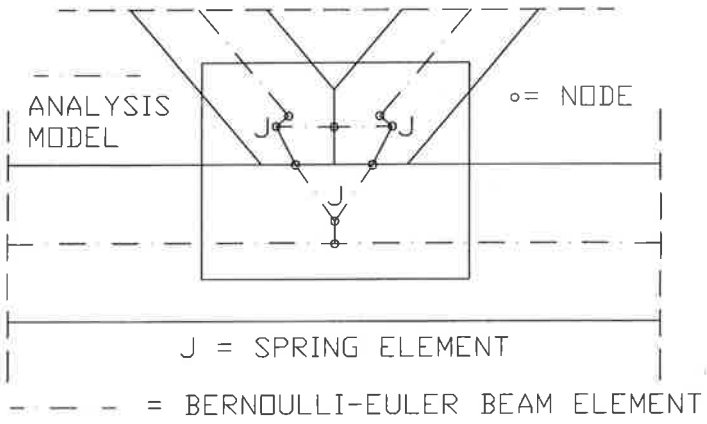


Fig. 1. Typical nail plate joint model

It is impossible to generate right analysis model for the joint of Fig. 1 without starting from the geometrical model. The details of Fig. 1 will be described later.

The paper deals with a new structural analysis model implemented to KPM's CAD system for the design of wooden nail plate structures (program called WoDe2000). A new analysis model follows strictly the new Finnish design codes for these structures. The program has been certified by SFS-Sertifiointi Oy.

The new analysis model includes very careful consideration of the geometrical model of joints and their analysis models. In practice this means that the analysis model will include much more finite elements than that used before. The analysis model is a plane frame

model and linear with the respect of geometrical and material entities. The only non-linearities arise when deriving the buckling lengths for composed wooden parts.

The product development of the new program has now been finished and the new program has been used for the design of new products. Some comparisons of the results given by the new structural model and the old one, used over 15 years in the program, will be given in the paper. The old model was quite simplified compared to the new model and it is believed that the new model will give more economical designs than the old system.

2. CODES AND THEORIES

The old version of the structural analysis of the program under consideration was based on an old Finnish code of practice [1] and allowable stress design concept (or to the use of one level safety factor). The ideas for the geometrical modelling have stayed about same about 20 years [2]. The new version is based on two Finnish codes of practice [3[, [4]. The first one is based on the Eurocode system and the second one is based on Finnish local code system. Both are based on so called limit state design or to the use of two level safety factors. The analysis models given in both the codes are same.

The old analysis model for the joint was simple hinge usually without eccentricity. The new model includes eccentricities and moreover the model includes the local translational and rotational flexibilities of joints. The eccentricities and semi-rigidity of joints reflects to the results of the structural analysis. The strength check for bars and joints after structural analysis includes also new theories appearing in the new codes of practice.

Typical local analysis model for the nail plate joint can be seen in Fig. 1. The beam elements are connected at the node J by spring elements, which include both translational and rotational springs. The spring factors are given in the codes. Node J is locating in the centre of effective area of the nail plate covered by separate wooden parts connected in the joint. Node J is connected by orthogonal very stiff beam element (eccentricity element) to the analysis model of a bar between joints. All the nodes J are connected together using very stiff beam elements.

The concept of very stiff element was verified using different cross-sectional properties for very stiff elements and proper wood sizes where found by these means. The minimum lengths of very stiff elements were found also by the method of trial and error and some minimum lengths were put to these elements if needed in order to get proper numerical results.

The program generates similar analysis models for all the joints appearing in the nail plate structure starting from the geometrical model. This means, that the geometrical entities (sizes and locations of nail plates as an example) must be defined before generating the analysis model. In practice this means some iteration when designing these structures.

3. EXAMPLE

Consider a simple nail plate structure shown in Fig. 2. as an example for theories.

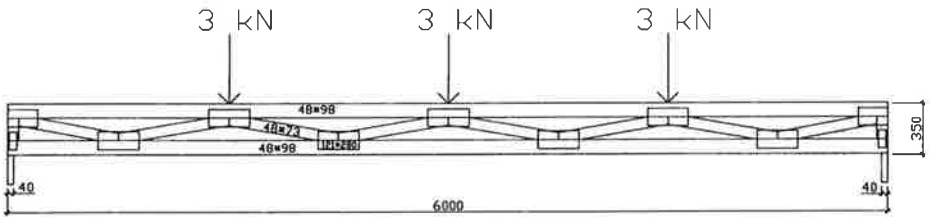


Fig. 2. Example structure

If the structure is analysed using the theory which has been valid about 200 years i.e. it is supposed, that there are hinges at every nodes, then the maximum tensile force of the bottom chord is 35.0 kN in this case. If the old allowable stress design method is used and the bottom chord is made of wood 48*98 mm^2, then the maximum stress is 7.44 MPa. The allowed tensile stress for the Finnish T24 wood is in this case 6.30 MPa, so the used capacity is 7.44/6.30 = 1.18.

If the same structure is analysed using the old version of the program under consideration, then the maximum tensile force is 33.2 kN and the tensile stress is 7.05 MPa. In this analysis it was supposed, that the chords (top and bottom) are continuous, which means, that there are bending moments without distributed loads at the chords, too. The maximum bending moment of the bottom chord is in this case 226 Nm, which causes bending stress 2.94 MPa and the allowed bending stress is for T24 wood quality 9.60 MPa. The used capacity of bottom chord is in this case 7.05/6.30 + 2.94/9.60 = 1.12 + 0.30 = 1.42.

It can be seen, that the very old design was 18% on the unsafe side, but about 20 years used method gave the result of 42% on the unsafe side.

When using the new version of the program for the same problem, then the following results were got. The used capacity of the bottom chord due to the axial force and due to the bending moment are as follows: 0.83 + 0.65 = 1.48.

It can be seen from the results, that the axial force and moreover the stresses caused by the axial forces are smaller when using the new version than those using the old version. On the other hand, it can be seen that the bending stresses of the bottom chord are about twice when using the old version compared to those of the old version. The joint eccentricities are one source for larger bending stresses and the local joint flexibilities are another source.

The similar trend can be seen when considering the displacements. The old version gave the maximum displacement 34 mm and the new version gave the result 41 mm. The increase of the displacement can be explained by the local joint flexibilities.

The most promising results when using the new version seem to be when considering the axially loaded bars. For the structure shown in Fig. 2 the following used capacities are got for the top chord:

- Two hundred years old theory: 1.17
- Twenty years old theory: 1.39 (=1.10 + 0.29)
- The most novel theory: 1.27 (0.76 + 0.51)

It can be seen, that the theories for buckling and moreover the equations appearing in the codes seem to give more economical results than before. The differences in this example are not large (about 10% in used capacities for top chord), but still the direction is proper taking into account the efforts for development and research work and exactness of the model.

The example was only one simple statically determinated structure. The largest development caused by the use of the new version is that the nail plate joint can resist bending moment. This makes it possible to use nail plate structures in applications, where they could not be used before, as shown in Fig. 3.

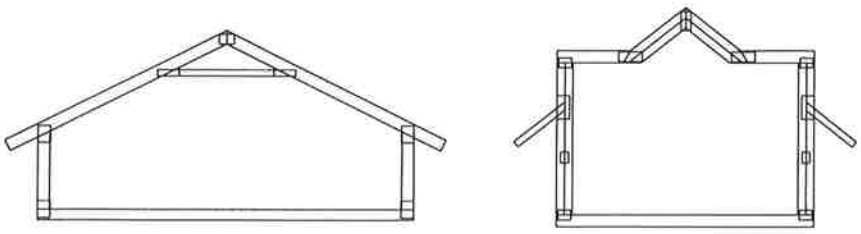


Fig. 3. Wooden frames made as nail plate structures

The new analysis model can be used for the design of many kinds of nail plate structures. It is believed, that this gives more market shares to nail plate structures and at least the position on the market (market leader for roof structures) can be kept as it is today. At least it is believed, that the new analysis model simulates the state of stresses and strains very accurately for the nail plate structures [5].

4. CONCLUSIONS

The new analysis model for nail plate structures can be used to model the structural behaviour rather exactly. The new model takes into account the local flexibilities and eccentricities in joints.

The local flexibilities of joints increase displacements of the structure very much. The force distribution is about the same as before, but the bending stiffness of joints cause bending moments especially to the diagonals and verticals of trusses. The eccentricities enlarge bending stresses of chords but at the same time the axial stresses are smaller than using the old theory. On the other hand, it is possible to design such structures, which were not possible to design before, which is a large step.

The real possible benefits or obstacles of the new model can be seen after wider experiences in practical projects. Now it can be said, that the design of axially loaded bars leads to more economical result than using the old system. Perhaps, the main reasons are the equations in the codes, not the structural model.

The strength check of joints was not considered here. They have been analysed in Ref. [5].

REFERENCES

[1] Naulalevyrakenteiden suunnitteluohjeet, NR-Toimikunta, PUURAKENTEIDEN LAADUNVALVONTAYHDISTYS ry, Espoo, 1983 (in Finnish)

[2] Mikkola M., Computer Aided Design of Wooden Trusses Using Nailplate Connectors, Proceedings of The Second International Conference on Civil and Structural Engineering Computing, CIVIL-COMP 85, Volume 1, Ed. B. H. V. Topping, Civil-Comp Press, 10 Saxe-Coburg Place, Edinburgh, 1985, pp. 149-154

[3] Naulalevyrakenteiden suunnitteluohjeet, Euronormi, SFS-Sertifiointi Oy, Päivitetty 6.10.99 (in Finnish)

[4] Naulalevyrakenteiden suunnitteluohjeet, RakMK:n ohjeisiin B10 perustuva rajatilamitoitus, SFS-Sertifiointi Oy, 6.10.1999 (in Finnish)

[5] Kevarinmäki A. Semi-rigid behaviour of nail plate joints, Helsinki University of Technology, Laboratory of Structural Engineering and Building Physics, Publications 109, Espoo, 2000

ANALYTICAL GENERALIZATION OF WILLIOT-MOHR DISPLACEMENT PLANE FOR LARGE DISPLACEMENT AND INTO THE SPACE.

PERTTI HOLOPAINEN
Laboratory of Applied Mechanics
Tampere University of Technology
P.O. Box 589, FIN-33101 Tampere, FINLAND

ABSTRACT

The basic Williot problem is presented in a plane when the extensions of bars are large. The corresponding basic problem is presented in space also. An exact solving for the vertex point of a bar triangle and a bar tetrahedron is presented and a selection procedure of roots. The presented solving procedures are applied to ordinary trusses when the Mohrs correction is made also.
Numerical iterative solving procedures are presented also using the Newton-Raphson method.

INTRODUCTION

When the truss is statically determinate, it can be given to bars arbitrary extensions for example, those corresponding to allowable stress. The cross sectional areas of the bars can be determined by some loading so, that all the bars get the allowable stresses. The signs of bar forces must be known. In many cases, when the values of displacements are moderate, the signs of the bar forces are the same as in the case of small displacements.
When it is necessary to determine the displacements of many nodes of trusses using equations of work, the bar forces should be calculated many times. M. Williot developed in 1877 a graphical method (Williots displacement plane), which is advantageous in cases such as these. Williots method presupposed however, that the node displacements of trusses are small and that trusses are planar trusses. In the basic case of Williots method the two bar vectors, crossing each other, must start from known points. So it is in the cantilever trusses. But for example simply supported trusses do not have a bar triangle with two known points. In these cases on the fixed support, for example, the direction of one bar must be fixed. Hence the truss is considered to be cantilever. Determining the displacements of nodes like this truss the other end of truss become above the support point. The displacement figure must be corrected so that it corresponds to real supporting. A similar correction was presented by Otto Mohr in 1887,[1] [2].

The analytical generelization presented here is started using the same basic problem as Williot. Here is presented the exact solution, when the extensions of bars are large. The bar forces need not be calculated many times but the extensions of the bars must be known. In space the corresponding basic structure is a bar tetrahedron with all three base points fixed. The exact solving of the displaced vertex point is obtainable and presented here in detail.
Starting from these basic cases and going on node by node the displacement figure can be obtained. For example in the simple supported trusses the Mohrs correction to the displacement figure is needed. As one solving equation is polynomial of the second degree. The other solving equations, one in plane and two in space are linear. The criterion to select the roots is, that the vertex point of the bar triangle or of bar tetrahedron shall not move so, that the bars rotated over each other.

BASIC PROBLEM IN PLANE

Consider the basic problem as shown in Fig. 1. The vertex C_0 of the bar triangle is moved because of the known bar extensions e_A, e_B into position *C*. It is assumed that the positions of the base points A,B are known. The displaced position *(x,y)* of *C* has been solved here both analytically and numerically.

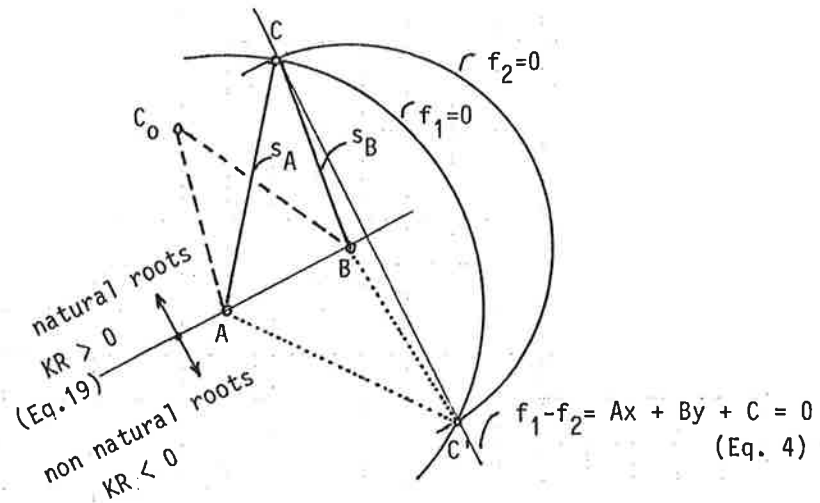


Figure. 1.

The displaced vertex C lies in the other intersection point of circles A with radius s_A and B with radius s_B. The lengths of the radii, or the lengths of the strained bars are

$$\begin{aligned} s_A &= s_{A_0} + e_A \\ s_B &= s_{B_0} + e_B \end{aligned} \tag{1}$$

Circles A and B intersect each other if

$$d_{AB} - s_B \leq s_A \leq d_{AB} + s_B \tag{2}$$

Two equations can be determined (Fig.1)

$$\begin{aligned} f_1 &= (x - x_A)^2 + (y - y_A)^2 - s_A^2 = 0 \\ f_2 &= (x - x_B)^2 + (y - y_B)^2 - s_B^2 = 0 \end{aligned} \tag{3}$$

to solve the displaced position *(x,y)=C*. Writing equations (3) out in full and subtracting f_2 from f_1 the expression for a straight line

$$f_1 - f_2 = Ax + By + C = 0 \tag{4}$$

can be obtained. In Eq. (4)

$$\begin{aligned} A &= 2(x_B - x_A) \\ B &= 2(y_B - y_A) \\ C &= x_A^2 - x_B^2 + y_A^2 - y_B^2 - s_A^2 + s_B^2 \end{aligned} \tag{5}$$

Equation (4) can be rearranged for y as follows

$$y = -\frac{A}{B}x - \frac{C}{B}. \tag{6}$$

Substituting y from Eq. (4) in to Eq.(3)1 the equation

$$ax^2 + bx + c = 0 \tag{7}$$

can be obtained. In Eq. (7)

$$a = 1 + \left(\frac{A}{B}\right)^2$$
$$b = 2\left(\frac{AC}{B^2} + \frac{A}{B}y_A - x_A\right) \qquad (8)$$
$$c = \left(\frac{C}{B}\right)^2 + \frac{C}{B}y_A + x_A^2 + y_A^2 - s_A^2.$$

From Eq. (7) it can be obtained

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \qquad (9)$$

When the right root x has been solved from Eq. (9) y can be calculated from Eq. (6). Both roots $x_{1,2}$ are real if equation (2) holds true.

Two special cases.

A. If the line of basepoints $AB \updownarrow\updownarrow Ox$, the $B=0$ and from Eq. (4)

$$x = -\frac{C}{A}. \qquad (10)$$

Substituting x from Eq. (10) into Eq. (3)1 the equation

$$y^2 + py + q = 0 \qquad (11)$$

can be obtained. In Eq. (11)

$$p = -2y_A$$
$$q = \left(-\frac{C}{A} - x_A\right)^2 + y_A^2 - s_A^2. \qquad (12)$$

Can be solved from Eq. (11)

$$y = -\tfrac{p}{2} \pm \sqrt{\left(\frac{p}{2}\right)^2 - q}. \qquad (13)$$

B. If the line of basepoints $AB \updownarrow\updownarrow Oy$, the A=0 and from Eq. (4)

$$y = -\frac{C}{B}. \tag{14}$$

Substituting y from Eq. (13) into Eq. (3)1 the equation

$$x^2 + px + q = 0 \tag{15}$$

can be obtained. In Eq. (15)

$$\begin{aligned} p &= -2x_A \\ q &= \left(-\frac{C}{B} - y_A\right)^2 + x_A^2 - s_A^2. \end{aligned} \tag{16}$$

Can be solved from Eq. (15)

$$x = -\frac{p}{2} \pm \sqrt{\left(\frac{p}{2}\right)^2 - q}. \tag{17}$$

Numerical solving.

Solving the coordinates of vertex point C from the Eq. (2) using the Newton-Raphson method the exact Jacobian can be derived without difficulty. The solving algoritm is

$$\begin{Bmatrix} x \\ y \end{Bmatrix}^{k+1} = \begin{Bmatrix} x \\ y \end{Bmatrix}^{k} - \alpha^* \cdot 2 \begin{bmatrix} x - x_A & y - y_A \\ x - x_B & y - y_B \end{bmatrix}^{-1} \begin{Bmatrix} f_1(x^k, y^k) \\ f_2(x^k, y^k) \end{Bmatrix} \tag{18}$$

$$\alpha^* \in (0,1]$$

The Jacobian is singular in a line

$$Ax + By + C = 0, \tag{19}$$

where

$$\begin{aligned} A &= y_A - y_B \\ B &= x_B - x_A \\ C &= x_A y_B - y_A x_B. \end{aligned} \tag{20}$$

In many cases can be used the convergence multipler $\alpha^* = 1$.

The line where the Jacobian is singular (Eqs. (19),(20)) passes through the base points A and B.

Selection of + or – in Eqs. (9),(13),(17).

It is assumed here that the displaced vertex point C is on the same side of the base line AB as its initial position C_0. At first the sign in front of the square root in (9), (13) or (17) is selected and x and y shall be calculated. The vector $\bar{k} \times \bar{r}_{AB}$ is perpendicular to the base line AB. If

$$KR = \frac{\bar{k} \times \bar{r}_{AB} \cdot \bar{r}_{AC_0}}{\bar{k} \times \bar{r}_{AB} \cdot \bar{r}_{AC_0}} > 0, \qquad (21)$$

the selection of sign before the square root is correct. Otherwise the opposite sign shall be selected.

BASIC PROBLEM IN SPACE

Consider the basic case in space shown in Fig. 2. The vertex D_0 of the bar tetrahedron is moved into position D because of known bar extensions e_A, e_B, e_C. It is assumed that the positions of basepoints A,B,C are known. The displaced position (x,y,z) of the vertex point D has been solved here both analytically and numerically.

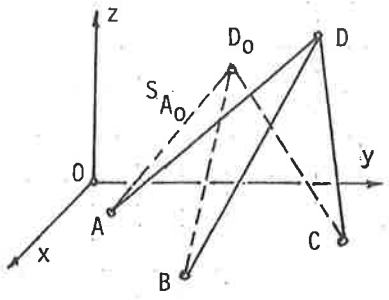


Figure 2.

The lengths of strained bars are

$$\begin{aligned} s_A &= s_{A_0} + e_A \\ s_B &= s_{B_0} + e_B \\ s_C &= s_{C_0} + e_C \end{aligned} \qquad (1)$$

The following three equations can be obtained

$$\begin{aligned} f_1 &= (x - x_A)^2 + (y - y_A)^2 + (z - z_A)^2 - s_A^2 = 0 \\ f_2 &= (x - x_B)^2 + (y - y_B)^2 + (z - z_B)^2 - s_B^2 = 0 \\ f_3 &= (x - x_C)^2 + (y - y_C)^2 + (z - z_C)^2 - s_C^2 = 0 \end{aligned} \qquad (2)$$

to solve x,y,z the displaced position of D. The three spheres (2) have two intersection points with certain boundary conditions of radii s_A, s_B, s_C. Writing out equations (2) in full and subtracting the two, plane equations

$$\begin{aligned} f_1 - f_2 &= A_1x + B_1y + C_1z + D_1 = 0 \\ f_2 - f_3 &= A_2x + B_2y + C_2z + D_2 = 0 \end{aligned} \qquad (3)$$

can be obtained. Equation (3)1 is the equation of the intersection plane of spheres 1 and 2. Equation (3)2 is the equation of the intersection plane of spheres 2 and 3. In Eq. (3)

$$\begin{aligned} A_1 &= 2(x_B - x_A) \\ B_1 &= 2(y_B - y_A) \\ C_1 &= 2(z_B - x_A) \\ D_1 &= x_A^2 - x_B^2 + y_A^2 - y_B^2 + z_A^2 - z_B^2 - s_A^2 + s_B^2 \end{aligned} \tag{4}$$

$$\begin{aligned} A_2 &= 2(x_C - x_B) \\ B_2 &= 2(y_C - y_B) \\ C_2 &= 2(z_C - z_B) \\ D_2 &= x_B^2 - x_C^2 + y_B^2 - y_C^2 + z_B^2 - z_C^2 - s_B^2 + s_C^2. \end{aligned}$$

y and z can be solved as a fuction of x from Eqs. (3). So

$$\begin{Bmatrix} y \\ z \end{Bmatrix} = \begin{bmatrix} B_1 & C_1 \\ B_2 & C_2 \end{bmatrix}^{-1} \begin{bmatrix} -D_1 & -A_1 \\ -D_2 & -A_2 \end{bmatrix} = \frac{1}{D} \begin{bmatrix} C_2 & -C_1 \\ -B_2 & B_1 \end{bmatrix} \begin{bmatrix} -D_1 & -A_1 x \\ -D_2 & -A_2 x \end{bmatrix} \tag{5}$$

$$D = B_1 C_2 - B_2 C_1$$

and further

$$\begin{aligned} y &= (-I_{11}A_1 - I_{12}A_2)x - I_{11}D_1 - I_{12}D_2 = a_y x + b_y \\ z &= (-I_{21}A_1 - I_{22}A_2)x - I_{21}D_1 - I_{22}D_2 = a_z x + b_z \end{aligned} \tag{5}a$$

where

$$\begin{aligned} I_{11} &= C_2 /(B_1 C_2 - B_2 C_1) \\ I_{12} &= -C_1 /(\quad " \quad) \\ I_{21} &= -B_2 /(\quad " \quad) \\ I_{22} &= B_1 /(\quad " \quad) \end{aligned} \tag{6}$$

$$\begin{aligned} a_y &= -I_{11}A_1 - I_{21}A_2 \\ b_y &= -I_{11}D_1 - I_{12}D_2 \\ a_z &= -I_{21}A_1 - I_{22}A_2 \\ b_z &= -I_{21}D_1 - I_{22}D_2 \end{aligned} \tag{7}$$

Substituting the formulae of y and z from Eq. (5)a into Eq. (2)1 the equation

$$ax^2 + bx + c = 0 \tag{8}$$

can be obtained. In Eq. (8)

$$\begin{aligned} a &= 1 + a_y^2 + a_z^2 \\ b &= -2x_A + 2a_y b_y - 2a_y y_A + 2a_z b_z - 2a_z z_A \\ c &= x_A^2 + b_y^2 - 2b_y y_A + y_A^2 + b_z^2 - 2b_z z_A + z_A^2 - s_A^2 \end{aligned} \tag{9}$$

From Eq. (8) can be solved

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \tag{10}$$

Substituting x from Eq. (10) into Eq. (5) y and z can also be obtained numerically.

Three special cases.

A. The base plane $ABC \updownarrow\updownarrow Oxy$.
In this case $C_1 = C_2 = 0$ and $D = 0$ (see Eqs. (4) and (5)). From the Eqs. (3)

$$\begin{aligned} A_1 x + B_1 y + D_1 &= 0 \\ A_2 x + B_2 y + D_2 &= 0 \end{aligned} \tag{11}$$

are obtainable. Solving Eqs. (11)

$$\begin{aligned} x &= \frac{1}{D_{(xy)}}(-B_2 D_1 + B_1 D_2) \\ y &= \frac{1}{D_{(xy)}}(A_2 D_1 - A_1 D_2) \\ D_{(xy)} &= A_1 B_2 - A_2 B_1 \end{aligned} \tag{12}$$

Substituting from Eqs. (12) calculated numerical values of x and y into Eq. (2)1 the equation

$$z^2 + pz + q = 0 \tag{13}$$

can be obtained. In Eq. (13)

$$p = -2z_A$$
$$q = z_A^2 - s_A^2 + x^2 - 2xx_A + x_A^2 + y^2 - 2yy_A + y_A^2. \tag{14}$$

From (13) can be solved

$$z = \frac{p}{2} \pm \sqrt{\left(\frac{p}{2}\right)^2 - q}. \tag{15}$$

B. The basic plane $ABC \updownarrow\updownarrow Oyz$.
In this case $A_1 = A_2 = 0$ (see Eqs. (4) and (5)). From the Eqs. (3)

$$B_1y + C_1z + D_1 = 0$$
$$B_2y + C_2z + D_2 = 0 \tag{16}$$

are obtainable. Solving Eqs. (16) or substituting $x = 0$ into Eqs. (5) the Eqs.

$$y = \frac{1}{D_{(yz)}}(-C_2D_1 + C_1D_2)$$
$$z = \frac{1}{D_{(yz)}}(B_2D_1 - B_1D_2) \tag{17}$$
$$D_{(yz)} = B_1C_2 - B_2C_1) \ (= D)$$

can be obtained. Substituting from Eqs. (17) the calculated numerical values of x and y into (2)1 the equation

$$x^2 + px + q = 0 \tag{18}$$

can be obtained. In Eq. (18)

$$p = -2x_a$$
$$q = x_A^2 + y^2 - 2yy_A + y_A^2 + z^2 - 2zz_A + z_A^2 - s_A^2. \tag{19}$$

From Eq. (18) can be solved

$$x = -\frac{p}{2} \pm \sqrt{\left(\frac{p}{2}\right)^2 - q}. \tag{20}$$

This special case B. can also be solved using the more general equations (1)...(10).

C. The base plane $ABC \updownarrow\updownarrow Ozx$.
In this case $B_1 = B_2 = 0$ and $D = 0$ (see Eqs. (4) and (5)). From Eqs. (3)

$$\begin{aligned} A_1x + C_1y + D_1 &= 0 \\ A_2x + C_2y + D_2 &= 0 \end{aligned} \tag{21}$$

are obtained. Solving Eqs. (21)

$$\begin{aligned} x &= \frac{1}{D_{(zx)}}(-C_2D_1 + C_1D_2) \\ z &= \frac{1}{D_{(zx)}}(A_2D_1 - A_1D_2) \\ D_{(zx)} &= A_1C_2 - A_2C_1 \end{aligned} \tag{22}$$

By substituting from Eqs. (22) calculated numerical values of x and z into Eq. (2)1 the equation

$$y^2 + py + q = 0 \tag{23}$$

can be obtained. In Eq. (23)

$$\begin{aligned} p &= -2y_A \\ q &= y_a^2 + z^2 - 2zz_A + z_A^2 + x^2 - 2xx_A + x_A^2 - s_A^2. \end{aligned} \tag{24}$$

From Eq. (23) it can be solved

$$y = \frac{p}{2} \pm \sqrt{\left(\frac{p}{2}\right)^2 - q}. \tag{25}$$

Selection of + or – in Eqs. (10), (15), (20), (25)

It is assumed that the displaced position D of the vertex point is on the same side of the base plane ABC as its initial position D_0. At first the sign in front of the square root in Eq.(10) is selected and x,y and z calculated. The vector $\bar{r}_{AB} \times \bar{r}_{AC}$ is perpedicular to base plane ABC. If

$$KR_{sp} = \frac{\bar{r}_{AB} \times \bar{r}_{AC} \cdot \bar{r}_{AD_0}}{\bar{r}_{AB} \times \bar{r}_{AC} \cdot \bar{r}_{AD}} > 0 \qquad (26)$$

the selection of sign in front of the square root is correct. Otherwise the opposite sign shall be selected. It is assumed in Eq. (26) that A,B and C follows counterclocwise of each other.

Instead of the analytical consideration in the special cases A. B. C the following procedure is presented also. In these cases the coordinate frame must be rotated, for example by a nearly small angle $\delta\theta$ around the axis $\bar{e} = (\bar{i} + \bar{j} + \bar{k}) / \sqrt{3}$, and x,y,z must be solved using the previously presented procedure in Eqs. (1)...(10). Then the coordinate frame must be rotated by an angle $-\delta\theta$ around the same axis $\bar{e}$ and the coordinates x,y,z of D calculated in the initial global coordinate frame.

Problems A.,B. and C. do not arise when the equations (2) are solved numerically using the Newton-Raphson method.

Numerical solving.

Solving x,y,z from equations (2) using the Newton-Raphson method, the exact Jacobian can be derived without difficulty. The solving algorithm is

$$\begin{Bmatrix} x \\ y \\ z \end{Bmatrix}^{k+1} = \begin{Bmatrix} x \\ y \\ z \end{Bmatrix}^{k} - \alpha^* \cdot 2 \begin{bmatrix} x - x_A & y - y_A & z - z_A \\ x - x_B & y - y_B & z - z_B \\ x - x_C & y - y_C & z - z_C \end{bmatrix}^{-1} \begin{Bmatrix} f_1(x^k, y^k, z^k) \\ f_2(x^k, y^k, z^k) \\ f_3(x^k, y^k, z^k) \end{Bmatrix}.$$

$$\alpha^* \in (0,1] \qquad (27)$$

The Jacobian is singular in plane

$$Ax + By + Cz + D = 0 \qquad (28)$$

where

$$\begin{aligned} A &= z_A y_C - y_A z_C + y_A z_B - z_A y_B + y_B z_C - z_B y_C \\ B &= x_A z_C - z_A x_C + z_A x_B - x_A z_B + z_B x_C - x_B z_C \\ C &= y_A x_C - x_A y_C + x_A y_B - y_A x_B + x_B y_C - y_b x_C \\ D &= x_A z_B y_C + y_A x_B z_C + z_A y_B x_C - x_A y_B z_C - y_A z_B x_C - z_A x_B y_C. \end{aligned} \qquad (29)$$

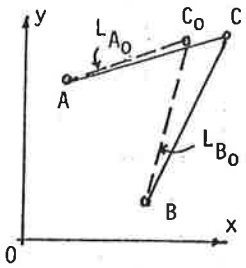
The plane where the Jacobian is singular (Eqs. (28),(29)), passes through the base points A, B and C.

APPLICATIONS

Basic problem in plane.

Example 1. A general case (Fig. 3a).

(a)



$L_{A_0} = 3.1623$ $L_{B_0} = 4.1231$ $e_A = 0.3$ $e_B = 0.1$

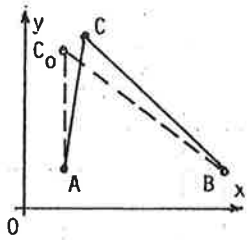
Exact $C_x = 4.310248$, $C_y = 5.014708$

$C'_x = -1.209825$ $C'_y = 1.334659$

N-R.	Start	$C^0(4,5)$	$C'^0(1,1)$
	Final	$C(4.310, 5.015)$	$C'(-1.210, 1.335)$
		$KR = 0.612$	$KR' = -1.846$
		Rounds 4	6
		$\alpha^* = 1$	

Special case A (Fig. 3b).

(b)



$L_{A_0} = 3$ $L_{B_0} = 5$ $e_A = 0.3$ $e_B = 0.1$

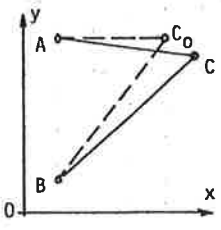
Exact $C_x = 1.11$ $C_y = 4.2981167$

$C'_x = 1.11$ $C'_y = -2.298166$

N-R.	Start	$C^0(1,4)$	$C'^0(1,-4)$
	Final	$C(1.110, 4.298)$	$C'_0(1.110, -2.298)$
		$KR = 0.910$	$KR = -0.910$
		Rounds 4	5
		$\alpha^* = 1$	

Special case B (Fig. 3c).

(c)



$L_{A_0} = 3$ $L_{B_0} = 5$ $e_A = 0.3$ $e_B = 0.1$

Exact $C_x = 4.298167$ $C_y = 4.89$

$C'_x = -2.298166$ $C'_y = 4.89$

N-R.	Start	$C^0(4,5)$	$C'^0(-2,5)$
	Final	$C(4.298, 4.890)$	$C'(-2.298, 4.890)$
		$KR = 0.910$	$KR' = -0.910$
		Rounds 4	4
		$\alpha^* = 1$	

Figure 3.

Basic problem in space.
Example 2.

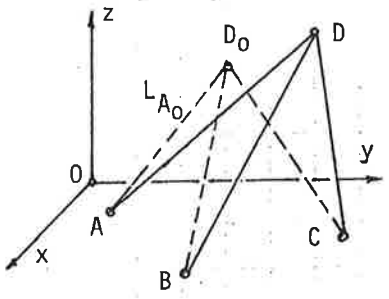


Figure 4.

$A = (1,1,0)$
$B = (3,4,0)$
$C = (2,7, Z_c)$
$D_0 = 2,4,3)$

$L_{A_0} = 4.3589 \quad e_A = 0.3$
$L_{B_0} = 3.1623 \quad e_B = 0.2$
$L_{C_0} = 3.9699 \quad e_C = 0.1$

$$d_k = \max_i \left| x_i^{k+1} - x_i^k \right| \quad \text{(i=1,2,3)}$$

General case: $Z_c = 0.4$:

Exact	N-R Start	Final	Stop: $d_k < 0.0001$
$x_D = 2.184588$	2	2.184588	$\alpha^* = 1$
$y_D = 4.110347$	4	4.110347	Rounds 5
$z_D = 3.260037$	3	3.260037	$KR = 0.9125134$
$x_{D'} = 1.367827$	2	1.267827	$\alpha^* = 1$
$y_{D'} = 4.654855$	4	4.654855	Rounds 5
$z_{D'} = -2.865674$	-3	-2.865674	$KR' = -0.9125132$

Special case A: Plane ABC ↕↕ Oxy.

A=(1,1,0) D_0 ==(2,4,3)

B=(3,4,0) $L_{C_0} = 4.242641$ L_{i_0} as before (i=A,B)

C=(2,7,0) e_i as in general case before (i=A,B,C)

Exact	N-R Start	Final	Stop: $d_k < 0.0001$
$x_D = 2.159008$	2	2.159008	$\alpha^* = 1$
$y_D = 4.127401$	4	4.127401	Rounds 5
$z_D = 3.252908$	3	3.252908	$KR = 0.9222516$
$x_{D'} = 2.159008$	2	2.159008	$\alpha^* = 1$

$y_{D'} = 4.127401$	4	4.1274	Rounds 5
$z_{D'} = -3.252908$	-3	-3.252908	$KR' = -0.9222517$

Special case B: Plane ABC ↕↕ Oyz.

A=(0,1,1) $D_0 = (3,2,4)$

B=(0,3,4) L_{i_0} *and* e_i as before in case A. (i=A,B,C)

C=(0,2,7)

Exact	N-R Start	Final	Stop: $d_k < 0.0001$
$x_D = 3.252908$	3	3.252908	$\alpha^* = 1$
$y_D = 2.159008$	2	2.159008	Rounds 5
$z_D = 4.127401$	4	4.1274	$KR = 0.9222516$
$x_{D'} = -3.252908$	-3	-3.252908	Stop: $d_k < 0.0001$
$y_{D'} = 2.159008$	2	2.159008	Rounds 5
$z_{D'} = 4.127401$	4	4.127401	$KR' = -0.9222516$

Special case C: Plane ABC ↕↕ Ozx.

A=(1,0,1) $D_0 = (4,3,2)$

B=(4,0,3) L_{i_0} *and* e_i as before in case B. (i=A,B,C)

C=(7,0,2)

Exact	N-R Start	Final	Stop: $d_k < 0.0001$
$x_D = 4.127401$	4	4.1274	$\alpha^* = 1$
$y_D = 3.252908$	3	3.252908	Rounds 5
$z_D = 2.159008$	2	2.159007	$KR = 0.9222516$
$x_{D'} = 4.127401$	4	4.127401	$\alpha^* = 1$
$y_{D'} = -3.252908$	-3	-3.252908	Rounds 5
$z_{D'} = 2.159008$	2	2.159007	$KR' = -0.9222517$

The system of equations for unknown node displacemts.

Except the Williot-Mohr method the nonlinear system of equations for all unknown node displacements with known bar extensions can be formulated simultaneously.
For each bar (Fig. 5) of a planar trusses it can be obtained the equations

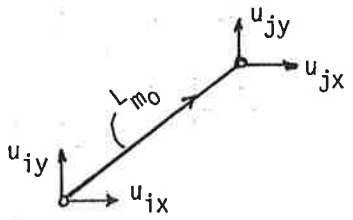


Figure 5.

$$(L_{m_0} + e_m)^2 = (L_{m_{0x}} + u_{jx} - u_{ix})^2 + (L_{m_{0y}} + u_{jy} - u_{iy})^2 \qquad (30)$$

and for each bar of a space truss it can be obtained the equations

$$(L_{m_0} + e_m)^2 = (L_{m_{0x}} + u_{jx} - u_{ix})^2 + (L_{m_{0y}} + u_{jy} - u_{iy})^2 + (L_{m0z} + u_{jz} - u_{iz})^2. \qquad (31)$$

In Eqs. (30) and (31) m=1...M and M is the number of bars. On the supports the corresponding displacementa are equal to zero. Hence the equations

$$u_{sx\ or\ y} = 0 \text{ (s=1...3 with plane truss, s=1...6 with space truss)} \qquad (32)$$

can be obtained in addition. The equations (30,32) or (31,32) can be solved simultaneously by the Newton-Raphson method without fixing none bars direction and without corrections the compatibility on the supports.
An example is presented in the lecture.

REFERENCES

[1] Pentti Loikkanen, Rakenteiden statiikka 1. Otava (1972).

[2] C.H.Norris and J.B.Wilbur, Elementary Structural Analysis. McGraw-Hill (1960).

[3] P.Henrici, Elements of Numerical Analysis. John Wiley & Sons (1964).

SHALLOW FLOOR BEAM BEHAVIOUR – GENERAL THEORY

Matti V. LESKELÄ
Structural Engineering Laboratory
University of Oulu
P.O. Box 191
FIN 90101 OULU, FINLAND
E-mail Matti.Leskela@oulu.fi

ABSTRACT

Shallow floor construction consists of beams supporting the decking structure on their bottom flanges or ledges and thus most of the beam depth is inside the concrete slab, which together form a composite structure. The simple construction does not include a simple theory, however, which would be required to design the structures efficiently. This paper discusses the development of composite behaviour between the beam and the floor slab, its variability as a function of the strain state and the evaluation of the stiffness and resistance to bending. Three systems are classified on account of the slab type: (a) hollow core slabs, (b) composite slabs with deep decking and (c) solid slabs or composite slabs with prestressed planks.

1. INTRODUCTION

Shallow floors or slim floors with steel beams integrated within a concrete slab are a popular form of construction for modern office buildings and typical for the system is that the soffit of the floor allows free access for installations in every direction. The span range for the beams is limited below 10 metres, however, as the floor depth is normally not more that 300 mm. The fast construction using hollow core decking elements has made the system very popular in Scandinavia, and the research on the hollow core slabs supported on beams now covers most of the behavioral aspects the desinger should know for making efficient designs. New developments where deep decking composite slabs are employed as slab structure were started in the UK [1], partly on the basis of the problems associated with the use and design of hollow core slabs. A synthesis of the flexural behaviour of the shallow floor beams is developed, considering mainly the systems classified as (a) and (b) in the abstract. It includes discussion of the interaction between the steel member and the slab structure in the serviceability and ultimate limit states. The conditions for developing the composite behaviour are shown, and partial connection theory for the plastic design of the ultimate bending resistance is introduced as modified from the basic method given in Eurocode 4 [2].

2. GENERAL FLEXURAL THEORY OF BENDING IN COMPOSITE BEAMS

What is normally considered as a composite beam is a compilation of a concrete slab stacked on the top flange of a steel beam and there is only one natural connection interface where the resistance of the connection has to be established. Thus in the 'ordinary' composite beam there is no difficulty in understanding the role and behaviour of the shear connection. While other forms of composite beams were earlier not favoured, a limited understanding of the general behaviour has resulted, and sometimes it is erraneously thought that shear connections can only prevail in horizontal surfaces around the top flange of the steel section. The general theory introduced here shows that the one and only condition required for the development of composite action is that one connection interface, horizontal or vertical, is available, but it is not critical where the interface resides.

From this it follows that by having multiple interfaces or continuous ones, the composite interaction is not intensified or improved, as long as material components in the composite section comply with the flexural theory, and Bernoulli hypothesis is valid.

2.1 Development of composite behaviour

For having full generality, two randomly shaped members in Fig. 1 with indices 'c' and 'a' are considered. Assume that the members are so assembled that they are in contact along the beam axis only on one level marked 'j' or 'jj'. The contact interface can be horizonal or vertical. The centroidal axes of the members are denoted by C_c and C_a for the bending without any composite interaction.

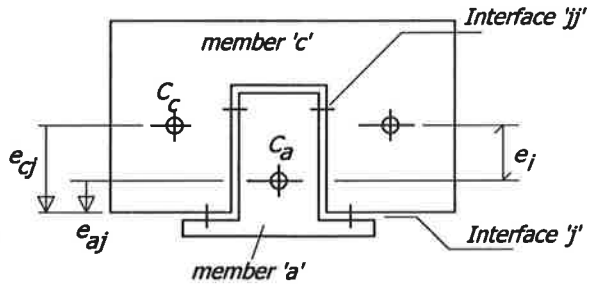


Fig. 1 *Composite section of a flexural member where structural components 'c' and 'a' are interconnected on one interface, 'j' or 'jj' (random selections)*

In any composite cross-section, an external bending moment M is always composed into stress resultants M_a, M_c and Ne_i, with $N = N_a = -N_c$ being the axial stress resultant acting in opposite directions on the neutral axes of the two members (tension considered as positive):

$$M = M_a + M_c + Ne_i \tag{1}$$

The curvature of the members is $\kappa = M/(EI)_{pi}$, in which $(EI)_{pi}$ represents the bending stiffness of the composite structure. Moreover,

$$M_c = \kappa (EI)_c \qquad \text{and} \qquad M_a = \kappa (EI)_a \tag{2}$$

in which $(EI)_c$ and $(EI)_a$ are the flexural stiffnesses of the members, and their sum, denoted by $(EI)_{ni} = (EI)_a + (EI)_c$, indicates the minimum stiffness of the system with no interaction. Thus

$$M = \kappa[(EI)_a + (EI)_c] + Ne_i = \kappa (EI)_{ni} + Ne_i \tag{3}$$

This is valid independent of the interaction rate, as the only condition applied is the common curvature obtained on the basis that the component members are in contact at all points along the beam axis.

Complete interaction is experienced when no strain difference appears at the connection interface, otherwise slipping at the interface is caused by the strain difference and partial interaction follows. The interface strains, ε_{cj} and ε_{aj}, of the connection level are calculated by means of axial stiffnesses and curvature κ:

$$\varepsilon_{aj} = N/(EA)_a + \kappa e_{aj}, \qquad \varepsilon_{cj} = -N/(EA)_c + \kappa e_{cj} \tag{4}$$

Assume now that the strain difference, $\Delta\varepsilon_j = \varepsilon_{cj} - \varepsilon_{aj}$, is known. The axial stress resultant may then be evaluated as

$$N = [\kappa e_i - \Delta\varepsilon_j](EA)_c(EA)_a/(EA)_i \tag{5a}$$

$$(EA)_i = (EA)_a + (EA)_c \tag{5b}$$

To obtain the N – M relationship, N must be inserted in Eq. (3) and then

$$M/\kappa = (EI)_{pi} = (EI)_{ni} + e_i[\kappa e_i - \Delta\varepsilon_j](EA)_c(EA)_a/(EA)_i \tag{6}$$

which is formally shortened as

$$(EI)_{pi} = (EI)_{ni}[(1+\alpha_i) - \Delta\varepsilon_j\alpha_i/(\kappa e_i)] \qquad \text{with} \qquad \alpha_i = \frac{e_i^2}{(EI)_{ni}}\frac{(EA)_c(EA)_a}{(EA)_i} \tag{7}$$

As the strain difference is variable along the span, Eq. (7) represents a non-constant stiffness, which has its minimum at the points of maximum moment, and the effective value is always smaller than $(EI)_i = (EI)_{pi}(\Delta\varepsilon_j = 0)$, but for small $\Delta\varepsilon_j$ (serviceability limit states) the effective stiffness practically equals to $(EI)_i$.

2.2 Shear connection and shear flow

Shear connection is required for maintaining the longitudinal force equilibrium between the material components, i.e. longitudinal shear forces $v_l dx = dN$ are activated, and

$$v_l = \frac{dN}{dx} \approx \frac{V(x)}{(EI)_{pi}} e_c (EA)_c - \frac{d(EI)_{pi}}{dx} \frac{M(x)}{(EI)_{pi}^2} e_c (EA)_c - \frac{d\Delta\varepsilon_j}{dx} \frac{(EA)_c (EA)_a}{(EA)_i} \tag{8}$$

which yields its simplest form in the case of complete interaction ($(EI)_{pi} = (EI)_i$ = constant, $\Delta\varepsilon_j = 0$):

$$v_{l.max} = \frac{V(x)}{(EI)_i} e_c (EA)_c \tag{9}$$

Lever arm e_c is measured to the centroid of area A_c from the neutral axis of the composite section.

3. INELASTICITY ISSUES

All the structures shall be designed to have a given resistance for bending and shear. While it is normally no problem to find enough resistance for shear in shallow floor beams, there is a bigger problem in defining the proper resistance to bending.

For applying plastic theory, stress blocks are used both for steel and concrete materials. The size of the stress blocks depends on the resistance of the shear connection and partial connections are defined as those affecting the value of the bending resistance, M_{Rd}. After achieving full connection the improvements in the connection do not enhance the bending resistance, being now $M_{pl.Rd}$. The degree of shear connection, η, is used for designing the beams for partial resistance $M_{Rd} < M_{pl.Rd}$, and in ordinary beams with ductile mechanical connection η is directly proportional to the number of connectors. It is then easy to interpolate $M_{apl.Rd/\eta=0} < M_{Rd} \leq M_{pl.Rd/\eta=1}$ with respect to η.

3.1 Partial connection theory for shallow floor beams [3]

Theoretical and experimental evidence shows that during bending longitudinally extending vertical (or mainly vertical) cracks emerge on the sides of the beam section starting from the points of zero moment in the span and extending towards the section of maximum moment. While the adhesive bond strength of the steel-concrete interfaces is lower than shear strength of the concrete, the cracks follow the surfaces of the steel section. This behaviour separates the composite section into two parts and after slipping has effectively started, both parts will have their own neutral axes for bending. As the concrete flange section does not have reinforcement in the direction of the beam axis, its compressive stress resultant must be balanced with the resistance of the connection, v_{Rd}, and

considering the shear span, L_s, and design widths b_{d1} the value for the compressed depth of the concrete flange must be

$$x \le \frac{v_{Rd} L_s}{2b_{d1}(0.85 f_{ck} / \gamma_c)} \le h_c \tag{10}$$

and then the plastic balance of internal forces is satisfied by having the depth of the web in compression as

$$h_{wc} = \frac{A_2 + A_w - A_1 - (2b_{d1}x + k_{top} b_1 h_{ct}) \dfrac{0.85 f_{ck} \gamma_a}{f_y \gamma_c}}{2t_w} \tag{11}$$

M_{Rd} is then calculated for the effective parts in steel and concrete.

Formally it may seem that the model presented will not work logically with respect to x, but it is easily shown that the result is correct and conservative, as h_{wc} is a function of x. Moreover, section $b_1 \times h_{ct}$ has only a minimal effect on the resistance and should be omitted (set $k_{top} = 0$), if it is not properly connected to the flange.

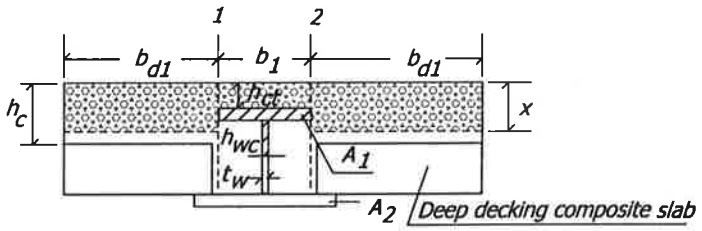


Fig. 2 *Effective cross-section for considering the partial connection theory. Note: it is not essential what is the real form of the steel section, as long as it can be divided into sections A_1, A_2 and A_w.*

3.2 Influence of the shear connection on the beam behaviour [3]

For obtaining best efficiency from the shear connection its behaviour should be ductile, i.e. after reaching the maximum load during slipping, the load should remain as high as possible when the slipping continues. This would justify the use of the maximum load as the resistance of the connection. Unfortunately bond connections between concrete and steel do not behave in this way but rather in the manner shown in Fig. 4, where a general load-slip relationship of the connection is divided into two parts, (s) and (u). To obtain an understanding on the general behaviour, the effects of the shear connection properties were parametrically analyzed by the non-linear finite element method after the system was initially calibrated with the beam tests and special connection tests [3].

Part (s) is responsible for the serviceability properties of the system, e.g. bending stiffness, and in this part slips δ_s are small and only cause negligible $\Delta\varepsilon_j$. In part (u) the slips will be such that only a fraction of the maximum resistance might be available, and the average resistance of the connection within the shear span depends on the end slip developed before the ultimate bending moment is reached.

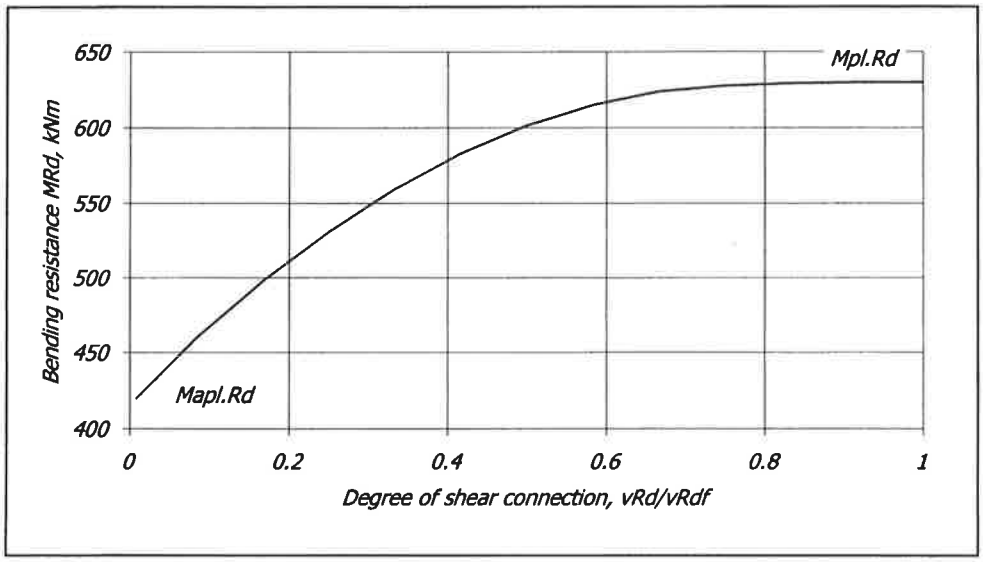


Fig. 3 *Typical form of the bending resistance - shear connection degree diagram for a shallow floor beam. The degree of the shear connection is defined as relative to the resistance of the full connection, v_{Rdf}*

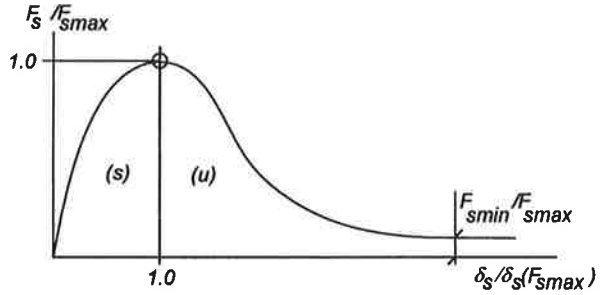


Fig. 4 *General form of the load-slip relationship of the connection*

If no special measures are taken to provide ductile connection and only adhesive bond between materials is considered, the typical overall behaviour of the shallow floor beam would be as follows:

- For serviceability loads the bending stiffness is obtained by full interaction theory, and typically the stiffness of the steel section is doubled. This is justified by the slips remaining in part (s).
- After that a continuous transition from the composite stiffness into a smaller one may happen due to higher strains and cracking in the concrete part which induce larger slips falling in part (u). This is indicated by a progressive increase of deflection with only a small increase of load.
- Finally not much more than $M_{apl.Rd}$ may be expected, if η remains small (small residual strength in the connection). Referring to Fig. 3, in such cases η is typically not more than 0.2.

The behaviour described is typically seen in one-dimensional beam tests if only bond connections are provided and width b_{d1} is large enough to activate bond failure. Finally not more than 5 to 10 % residue of the maximum strength will remain, unless pressure of concrete against the steel web and the due frictional effects are considered (hogging bending of the slab transverse to the beam, see Fig. 5). These are proportional to the amount of the hogging reinforcement and they conservatively evaluate to $\mu A_e f_{sk}/\gamma_s$ with μ = 0.5 and A_e being the reinforcement area per unit length of the beam. Theoretical and experimental evidence for this effect exists, and it is justified to allow for it in real floor construction, and it will considerably improve the bending resistance, depending on the amount of A_e. With reference to Fig. 3, η will typically be in the range of 0.4 to 0.5.

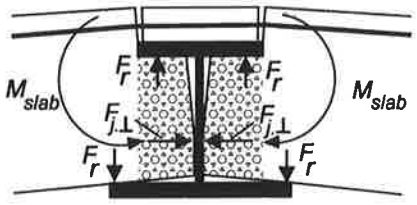


Fig. 5 *Transverse hogging bending will improve the behaviour of the shear connection in the beam due to compressive forces $F_{j.\perp}$.*

3.3 Influence of the slab type on the floor behaviour

In Scandinavia the most important decking elements for shallow floors are hollow core slabs, and by integrating beams with these elements a similar effective serviceability stiffness to that with deep decking composite slabs can be obtained. The main difference in behavior is that the resistance of the hollow core units often becomes critical well before the plastic behavior of the beams is entered [5, 6, 7 and 8].

Comparison of the system with deep decking composite slabs to a similar system with hollow core decking revealed that the effectiveness of the bond shear connection in

relation with hollow core slabs is better, and this means that the hollow core units suffer more from the composite behavior than do the composite slabs. This may seem irrational but is attributable to the more flexible web section of hollow core slabs for transverse shear forces (so called *'Vierendeel effect'* between the hulls of the slab). Another issue is that the compression flange due to the top hull of the slab is smaller in size so as not to activate as high stress resultants as will emerge in deep decking composite slabs and composite slabs with prestressed planks.

No part of the hollow core units can reasonably be exploited in the plastic section analysis when designing the bending resistance of the beam, but it has proved affordable to calculate the plastic moment for the composite section, where the concrete is only allowed for in the width of the steel section [4]. The other option would be the more cumbersome way to evaluate elastic resistance based on load histories and stress analysis on account of the full composite section with compression flanges on sides of the beam section. Numerically the results from the two options do not differ much, but slightly more efficiency may be obtained with plastic analysis [3].

Resistance in the shear connection interfaces [9]

The resistance of the connection on one side of the steel member should be evaluated as the minimum of the two values, $v_{Rd.1}$ and $v_{Rd.2}$:

- $v_{Rd.1}$ is the resistance in the steel-concrete interface which follows the surface of the steel section,

$$v_{Rd.1} = v_{Rd.mc} + \mu A_e f_{sk} / \gamma_s \tag{12}$$

- $v_{Rd.2}$ is the resistance in the vertical interface 1 or 2 through concrete slab,

$$v_{Rd.2} = 2.5 A_{cv} \eta \tau_{Rd} + A_e f_{sk} / \gamma_s \tag{13}$$

with $v_{Rd.mc} \geq 0$ being the total resistance of the bond and mechanical connection. It is clear from Eqs. (12) and (13) that $v_{Rd.1}$ will be critical when $v_{Rd.mc}$ is small. For bond connection $v_{Rd.mc} = h_w \tau_{Rd}$ and $\tau_{Rd} \leq 0.2$ MPa.

4. CONCLUSIONS

Shallow floor integrated beams are a form of composite beams in which both intentional and unintentional composite behaviour may be classified. It has been customary to refer to unintentional composite behaviour in connection with hollow core slabs supported on beams, but in case of beams integrated with deep decking composite slabs the composite system is exploited more generally and plastic behaviour in the composite system can be accounted for.

Bond connections are only applicable in case of hollow core slabs, and typically in these systems the interaction effectiveness is initially high (refer to Eq. (9)) but deteriorates

when the loads are increased (especially last component in Eq. (8)). Prior to failure of the slabs the composite interaction has reduced to a level which provides satisfactory total safety with respect to characteristic loads normally appearing e.g. in commercial buildings.

For design purposes of the shallow floor beams it is most desirable to be able to employ plastic methods of analysis, and partial connection theory is introduced for the purpose. The properties of the connection should then be considered carefully so as not to over-estimate the bending resistance. In beams supporting hollow core slabs the plastic section analysis may only be done for the beam body with no allowance of the composite interaction with the slab units. However, real composite behaviour should be considered in the serviceability analyses so as not to under-estimate the bending stiffness which critically affects the precambering design.

NOTATION

A_1, A_2, A_w Areas for the steel section parts, top flange, bottom flange and web,
A_e Area of the slab reinforcement across the beam,
A_{cv} Area of the slab cross-section,
b_{d1} Design width of the concrete compression flange outside the beam section,
x Compressed depth of the concrete flange,
$M_{apl.Rd}$ Plastic bending resistance of the steel section,
$M_{pl.Rd}$ Plastic bending resistance of the composite section, full connection,
M_{Rd} Plastic bending resistance of the composite section, partial connection,
v_{Rd} Resistance of the shear connection per unit length,
L, L_s Span of the beam and the shear span,
h_w Depth of the steel web,
h_{wc} Depth of the steel web in compression,
h_c Depth of the composite slab over the sheeting,
$v_{Rd.mc}$ Resistance of the mechanical shear connection per unit length,
f_{ck} Cylinder strength of the concrete,
f_y Yield strength of the steel,
f_{sk} Strength of the reinforcement,
η Degree of the shear connection = v_{Rd}/v_{Rdf},
τ_{Rd} Design bond strength of the steel-concrete interface.

REFERENCES

1. Lawson, R.M., Mullet, D.L. and Rackham, J.W., "Design of Asymmetric *Slimflor*® Beams using Deep Composite Decking". The Steel Construction Institute, Publication SCI-P-175, 1997
2. ENV 1994-1-1, Design of steel and concrete composite structures. Part 1-1: General rules and rules for buildings, CEN 1992

3. Leskelä, M.V. and Hopia, J., "Steel Sections for Composite Shallow Floors". Report RTL 0053E, University of Oulu, Structural Engineering Laboratory, March 2000, Oulu, Finland
4. Bode, H., Dorka, U.E., Stengel, J., Sedlacek, G. and Feldmann, M., "Composite Action in Slim Floor Systems". In: Composite Construction in Steel and Concrete III, Eds. C. Dale Buckner and Bahram M. Shahrooz. Proceedings of an Engineering Foundation Conference, Irsee, Germany 1996, American Society of Civil Engineers, 1997, 472-485.
5. Leskelä, M.V. and Pajari, M., "Reduction of the Vertical Shear Resistance in Hollow Core Slabs when Supported on Beams". Concrete 95, Conference Papers, Volume One, CIA and FIP, Brisbane, Australia 1995, 559-568
6. Pajari, M., "Shear resistance of prestressed hollow core slabs on flexible supports". VTT Publications 228. Espoo 1995
7. Pajari, M. and Koukkari, H., "Shear Resistance of PHC Slabs Supported on Beams. I: Tests". ASCE Journal of Structural Engineering, Vol. 124, No 9, September 1998, 1050-1061.
8. Pajari, M., "Shear Resistance of PHC Slabs Supported on Beams. II: Analysis". ASCE Journal of Structural Engineering, Vol 124, No. 9, September 1998, 1062-1073.
9. Leskelä, M.V., "Shallow Floor Integrated Beams and Their Components: Comparison of Behaviour". A paper for the Unified Engineering Foundation Conference 'Composite Construction in Steel and Concrete IV' to be held in Banff, Canada 28.5-2.6.2000, 12p. To be published.

ALUMIINISAUVOJEN PURISTUSLUJUUDESTA

P.Hassinen & J.Jensen
Teknillinen korkeakoulu
Rakenteiden mekaniikka
PL 2100, 02015 TKK

TIIVISTELMÄ

Alumiinin pieni kimmokerroin korostaa erilaisten epästabiiliusilmiöiden tärkeyttä alumiinirakenteiden mitoituksessa. Alumiiniseosten jännitysvenymäriippuvuus on jatkuva kuvaaja ilman selkeää myötörajaa. Tästä syystä alumiinirakenteiden kestävyyden laskemisessa alumiinin käyttäytymistä kuvataan kimmo-plastisten ainemallien ohella jatkuvilla Ramber-Osgood malleilla. Artikkelissa tarkastellaan ainemallien merkitystä alumiinisauvojen puristuskestävyyteen. Laskennallisesti määritettyjä tuloksia verrataan pienillä koekappaleilla tehtyjen kokeiden tuloksiin.

1. ALUMIINIRAKENTEIDEN OMINAISPIIRTEITÄ

Alumiiniseoksista pursotettuja profiileita käytetään keveyden ja suuren lujuuden vuoksi vaativissa rakenteissa. Pursotustekniikka mahdollistaa alumiiniprofiileiden vapaan muotoilun rakenteellisten vaatimusten ja ulkonäköseikkojen asetttamien vaatimusten mukaan. Teräkseen verrattuna pieni kimmokerroin korostaa epästabiiliusilmiöiden merkitystä puristettujen alumiinisauvojen käyttäytymisessä ja suunnittelussa. Alumiiniprofiilit ovatkin usein poikkileikkaukseltaan monimutkaisia erilaisilla reuna- ja välijäykisteillä varustettuja sauvoja. Lisäksi alumiiniseosten jatkuva, vailla myötörajaa oleva jännitysvenymäriippuvuus tuo erityispiirteitä alumiinirakenteiden mitoittamiseen. Alumiinirakenteiden mitoitusohjeissa ja käsikirjoissa annetaan monimutkaisia ja osin epähavainnollisia lausekkeita kriittisen ominaisjännityksen laskemiseksi. Kriittisen ominaisjännityksen avulla lasketaan tyypillisesti suhteellinen hoikkuusluku ja sen jälkeen puristuslujuus Perry-Robertson lausekkeita käyttäen.

TKK:ssa toteutettiin kolme vuotta sitten diplomityötehtävänä pieni tutkimushanke, jossa luotiin kokeellista tietoa alumiinisten puristettujen sauvojen käyttäytymisestä ja verrattiin kokeellisia tuloksia laskennallisesti eri standardien laskentamallien avulla saataviin tuloksiin /Jensen 1997/. Kokeet tehtiin poikkileikkaukseltaan pienillä alumiinisauvoilla, jotta pienehkössä tutkimushankkeessa voitaisiin luoda käytännössä kiinnostavalle hoikkuusalueelle kaikki epästabiiliusilmiöt kattava kokeellinen aineisto. Laskennallisia vertailuja on sittemmin jatkettu mallintamalla käytetyn alumiiniseoksen jännitys-

venymäriippuvuus erilaisilla ainemalleilla ja hakemalla puristettujen profiileiden kantokyvyn arvoja elementtimenetelmää käyttäen.

Puristetun alumiinisauvan kantokyvyn, tai kestävyyden suomalaisissa laskentaohjeissa käytettyä käsitettä käyttääksemme, voi määrätä poikkileikkauksen myötääminen, poikkileikkauksen osan paikallinen lommahtaminen tai sauvan vääntö- tai tasonurjahdus. Nämä ilmiöt voivat esiintyä myös yhdessä, jolloin epästabiiliusilmiöiden yhteisvaikutus tekee kestävyyden arvioimisesta entistä vaikeamman tehtävän.

Puristetun rakenteen kestävyyden arvioimiseen on epästabiiliusilmiöiden ja aineen plastisoitumisen vuoksi vaikea kehittää yksinkertaisia lausekkeita. Laskentaohjeissa ja standardeissa esiintyvät lausekkeet ovat yksinkertaistettuja ja ne peittävät helposti laskentamallin perustana olevan fysikaalisen ilmiön /ENV 1999-1-1 1998, RIL87-1998/. Nykyaikana puristetun sauvan kestävyys voitaisiin laskea numeerisia menetelmiä käyttäen. Epästabiiliusilmiöihin liittyvien ominaisarvojen ja -muotojen laskeminen onnistuukin jollakin tavalla monimutkaisillekin poikkileikkauksille ja rakenteille, mutta rakenteen kestävyyden arvioimiseen ovat numeeriset menetelmät käytännön suunnittelutyötä ajatellen kuitenkin vielä liian raskaita ja niitä käytetään vain erikoistilanteissa.

2. AINEMALLIT

Alumiiniseosten jännitysvenymäriippuvuus on jatkuva kuvaaja vailla selkeää myötörajaa. Laskennallisissa analyyseissa alumiiniseoksille käytetään kimmoplastisten ainemallien lisäksi jatkuvaa Ramberg-Osgood mallia. Yksiaksiaalisessa tapauksessa kimmoplastinen ja Ramberg-Osgood malli voidaan esittää yhtälöillä (1) ja (2) /Mazzolani 1995, Hibbitt, Karlsson & Sorensen 1998/.

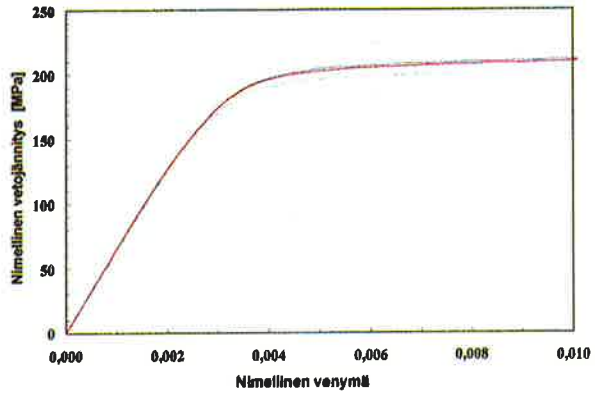
$$\varepsilon = \frac{\sigma}{E}, \text{ kun } |\varepsilon| \le f_y/E \text{ ja} \tag{1a}$$

$$\varepsilon = \frac{\sigma}{E} + \varepsilon^{pl}, \text{ kun } |\varepsilon| > f_y/E \tag{1b}$$

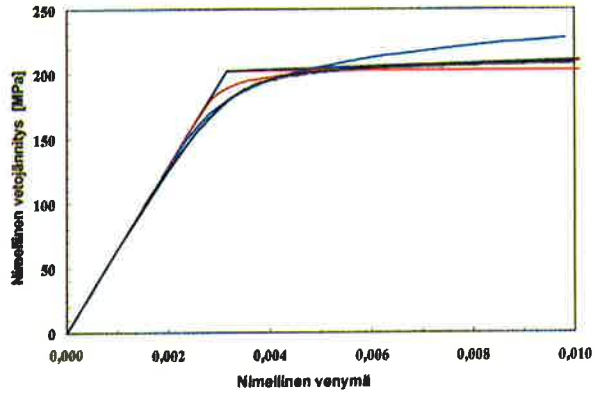
$$\varepsilon = \frac{\sigma}{E} + \frac{\alpha}{E}\left(\frac{|\sigma|}{\sigma^0}\right)^{n-1}\sigma \tag{2}$$

Numeerisia laskelmia varten on kokeellisesti määritettyä jännitysvenymäriippuvuutta kuvattu ideaalisella kimmo-plastisella ja kinemaattisesti lujittuvalla kimmo-plastisella ainemallilla. Ramberg-Osgood malli on sovitettu yhteensopivaksi keskisuurten ($\varepsilon < 0.015$) ja toisaalta pienten ($\varepsilon < 0.005$) kokeellisten venymien kanssa. Lisäksi kokeellista jännityksen ja venymän välistä yhteyttä on mallinnettu paloittain jatkuvien suoran yhtälöiden avulla, joka malli kuvaa suhteellisen tarkasti todellista keskimääräistä jännitysvenymäriippuvuutta (kuva 1).

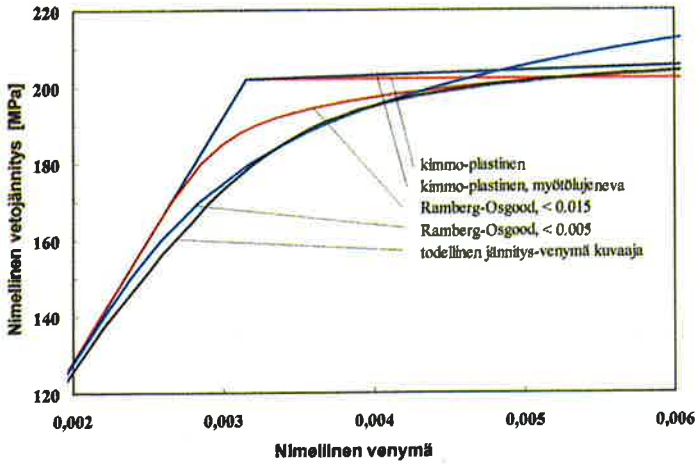
a



b



c

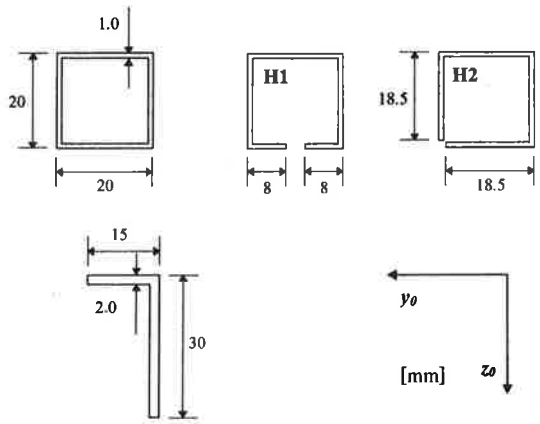


Kuva 1. Poikkileikkaukseltaan epäsymmetrisen kulmaprofiilin muotoisten koekappaleiden alumiiniseoksen AW6063 T6 a) kokeellisia ja b, c) laskentamallien mukaisia jännitysvenymäriippuvuuksia.

Vetokokeisiin perustuvat neliöputkiprofiilin materiaaliparametrit laskelmissa ovat olleet $E = 65200$ N/mm^2, $\nu = 0.33$, $f_{0.2} = 178$ N/mm^2 ja $E_t = 480$ N/mm^2. Epäsymmetrisen kulmaprofiilin parametrit ovat vastaavalla tavalla olleet $E = 64100$ N/mm^2, $\nu = 0.33$, $f_{0.2} = 202$ N/mm^2 ja $E_t = 1100$ N/mm^2. Kulmaprofiilin alumiiniseokselle pienten venymien perusteella määritetyt Ramberg-Osgood mallin parametrit ovat $\alpha = 0.4286$, $\sigma^0 = 200.5$ N/mm^2 ja $n = 12$. Vastaavasti keskisuurten venymien mukaan kiinnitetyiksi ainevakioiksi saadaan $\alpha = 0.4286$, $\sigma^0 = 199.4$ N/mm^2 ja $n = 33.7$.

4. KOETULOKSIA

Puristettujen alumiinisauvojen käyttäytymistä tutkittiin kokeilla, joihin valittiin neliöputkiprofiili ja epäsymmetrinen kulmaprofiili /Jensen 1997/. Neliöputkiprofiilin poikkileikkauksen sivumitta oli 20 mm ja nimellinen seinämävahvuus 1 mm. Kulmaprofiilin laippojen nimelliset leveydet olivat 30 ja 15 mm ja seinämävahvuus 2 mm. Putkiprofiilista valmistettiin lisäksi kahdenlaisia, poikkileikkaukseltaan avoimia profiileita avaamalla putket sivun keskiviivaa tai nurkkaa pitkin (kuva 2). Laskennallisia tarkasteluja varten koekappaleiden geometriset mitat ja materiaaliominaisuudet määritettiin huolellisesti.



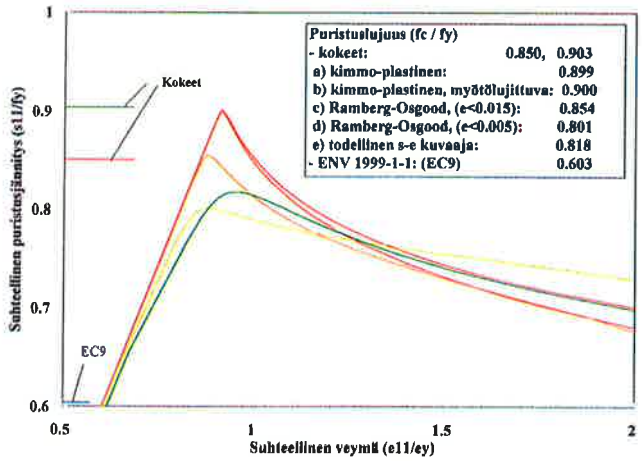
Kuva 2. Alumiiniprofiileista valmistettujen koekappaleiden nimelliset poikkileikkausmitat

Kokeissa sauvat kiinnitettiin molemmista päistään jäykästi kiinni puristuskoneeseen ja kuormitettiin keskisellä kuormalla kasvattamalla sauvan puristumaa vakionopeudella. Kokeissa mitattiin voiman ja aksiaalisen puristuman lisäksi sauvan sivuttaiset siirtymät sauvan puolivälissä. Sivuttaisten siirtymien avulla laskettiin poikkileikkauksen siirtymiä ja kiertymiä alkuperäisessä pääjäyhyyskoordinaatistossa (kuva 3). Koesauvojen pituuksia muutettiin sellaisella alueella, joka puristettujen sauvarakenteiden hoikkuuden puolesta on käytännössä mielenkiintoinen. Puristetun sauvan nurjahdusherkkyyttä kuvaava suhteellinen hoikkuus $\bar{\lambda} = \sqrt{f_y / \sigma_{cr}}$ vaihteli poikkileikkaukseltaan neliöputken muotoisilla koekappaleilla välillä $\bar{\lambda} = 0.11$ ja 1.34 ja vastaavasti kulmaprofiilin muotoisilla

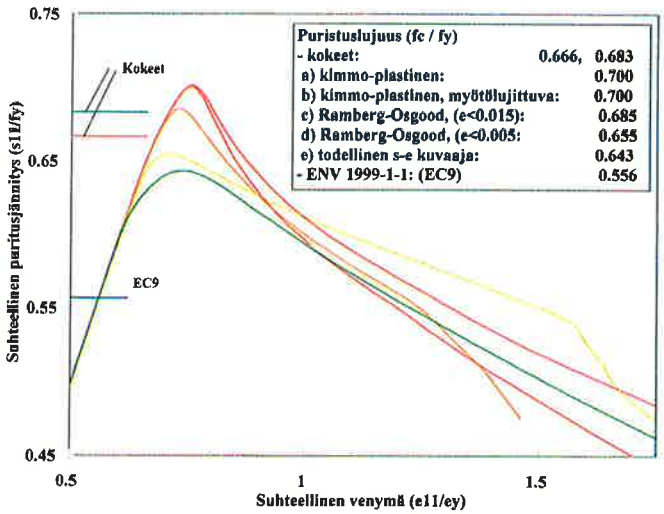
jotka taivuttavat ja kiertävät sauvaa pituuden puolivälissä. Pistekuormat häviävät, kun suurimmat venymät ylittävät myötövenymän arvon.

Laskentatulosten mukaan kimmoplastisiin ainemalleihin perustuvat analyysit tuottavat korkeimmat arvot sauvan puristuslujuudelle. Alhaisin lujuus saadaan eurooppalaisen alumiinirakenteiden esistandardin /ENV 1999-1-1 1998/ mukaisilla lausekkeilla. Numeeristen analyysien tuottamat puristuslujuudet yhtyvät hyvin kokeellisiin tuloksiin. Sen sijaan laskentastandardin tulokset alittavat kokeellisesti määritetyt lujuuden arvot voimakkaasti silloin, kun vääntönurjahdusmuoto hallitsee profiilin käyttäytymistä (kuva 4).

a

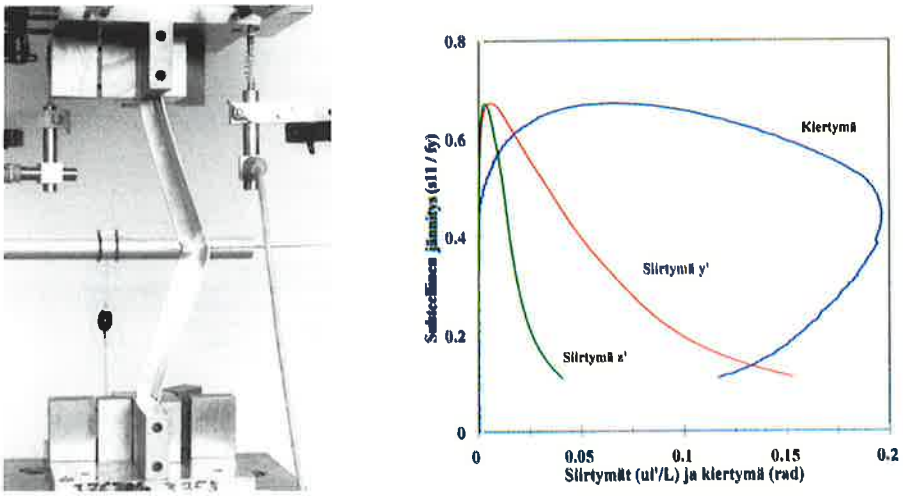


b



Kuva 4. Laskettuja ja kokeellisia puristuslujuuden arvoja epäsymmetriselle kulmaprofiilille, kun profiilin vapaa pituus on a) 130 mm ja b) 250 mm.

koekappaleilla välillä $\bar{\lambda} = 0.20$ ja 1.74. Kokeissa pyrittiin hakemaan esille kaikki tärkeät epästabiiliusmuodot. Neliöputken seinämäpaksuus osoittautui liian suureksi niin, että seinämän paikallisen lommahduksen vaikutusta puristuskestävyyteen ei saatu esille.



Kuva 3. a) Alumiinisauvojen puristuskokeiden koejärjestely ja b) kokeelliset siirtymät ja kiertymä sauvan puolivälissä pääjäyhyyskoordinaattien suunnissa. Kuvassa olevan epäsymmetrisen kulmaprofiilin vapaa pituus on 250 mm.

5. LASKENNALLISIA TULOKSIA

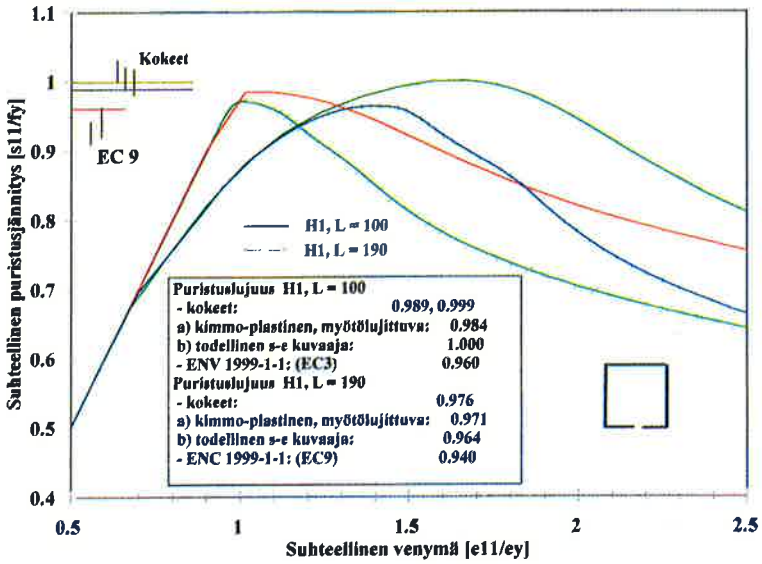
Vertailulaskelmien tarkoituksena oli tutkia mahdollisuuksia kuvata numeerisilla menetelmillä puristetun sauvan käyttäytymistä ja ennustaa rakenteen kestävyyttä sekä edelleen selvittää erilaisten ainemallien vaikutusta laskettuihin kestävyysarvoihin. Laskentamalleissa profiilit kuvattiin kahdeksan solmuisilla kuorielementeillä, joita putkiprofiilien yhdessä seinämässä seinämän leveyssuunnassa oli yhteensä neljä. Epäsymmetrisen kulmaprofiilin leveämmässä laipassa oli leveyssuunnassa viisi ja kapeammassa laipassa kolme kuorielementtiä. Elementtitiheys kasvoi sauvan päitä ja sauvan puoliväliä kohti. Numeeristen analyysien lisäksi koesauvojen puristuskestävyyttä on arvioitu eurooppalaisessa alumiinirakenteiden esistandardissa annettujen laskentamallien avulla /ENV 1999-1-1 1998/.

5.1 Avoin kulmaprofiili

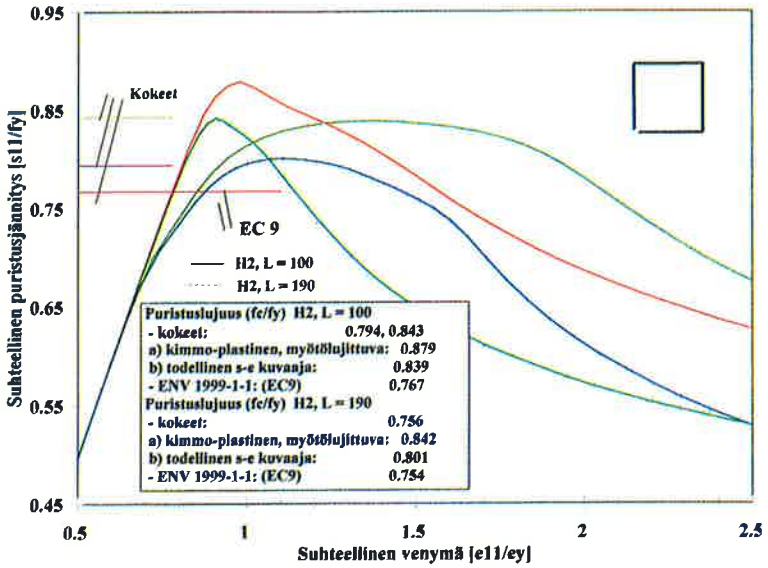
Avoimen kulmaprofiilin ensimmäinen ominaismuoto on lyhyillä sauvoilla laippojen paikallinen lommahdusmuoto, joka sauvan pituuden kasvaessa liittyy vääntönurjahdusmuotoon (kuva 3). Alkuhäiriönä laskentamallissa on käytetty kahta pientä pistekuormaa,

5.2 Avattu putkiprofiili

a



b



Kuva 5. Laskettuja ja kokeellisia puristuslujuuden arvoja avatulle neliöputkiprofiilille, kun profiilin on avattu a) sivun keskiviivaa ja b) nurkkaa pitkin. Profiilien vapaat pituudet ovat L = 100 mm ja 190 mm.

Avatun neliöputkiprofiilin ensimmäiseen ominaismuotoon vaikuttaa reunaltaan vapaiden seinämien aiheuttama poikkileikkauksen muodonvääristyminen. Numeerisissa laskelmissa

on avatuille neliöprofiileille käytetty vain lujittuvaa kimmoplastista ja toisaalta todellista jännityksen ja venymän välistä yhteyttä jäljittelevää ainemallia. Alkuhäiriönä laskelmissa on käytetty L = 100 mm pituisille sauvoille ensimmäistä ominaismuotoa ja L = 190 mm pituisille sauvoille ensimmäisen ja toisen ominaismuodon summaa. Kunkin ominaismuodon mukaisen alkuhäiriön amplitudin arvoksi on valittu 10% seinämän paksuudesta, joka kuvaa geometriselta muodoltaan lähes virheettömien koekappaleiden alkumuotoa. Molempiin ainemalleihin perustuvat numeeriset analyysit ja myös europpalaisen esistandardin antama laskentamalli tuottavat lähellä koetuloksia olevia puristuslujuuden arvoja (kuva 5).

6. YHTEENVETO JA JOHTOPÄÄTÖKSET

Epästabiiliusilmiöt vaikuttavat merkittävästi puristettujen alumiinisauvojen käyttäytymiseen ja kestävyyteen. Numeeriset analyysit osoittavat, että alumiinisauvojen kestävyyttä voidaan luotettavasti arvioida laskennallisilla menetelmillä. Alumiiniseosten jännitysvenymäriippuvuus on jatkuva kuvaaja vailla selkeää myötörajaa. Yksinkertainen kimmoplastinen ainemalli ei silloin kuvaa tarkasti materiaalin ominaisuuksia. Numeerisissa analyyseissa käytettävällä ainemallilla on merkitystä laskentatulokseen. Eurooppalaisessa alumiinirakenteiden suunnittelun esistandardissa annetut laskentamallit tuottavat selkeästi alaraja-arvoja alumiinisauvojen puristuskestävyydelle silloin, kun vääntönurjahdus hallitsee rakenteen käyttäytymistä.

LÄHTEET

1. Jensen, J. 1997, Jämförelse mellan beräknad och genom provning fastställd bärförmåga hos tryckta aluminiumprofiler. Diplomityö. Teknillinen korkeakoulu.

2. ENV 1999-1-1 1998, Eurocode 9: Design of aluminium structures - Part 1 - 1: General rules.

3. Suomen Rakennusinsinöörien Liitto RIL r.y. 1998, Alumiinirakenteiden suunnitteluohjeet RIL 87-1998.

4. Mazzolani, F. 1995, Aluminium alloy structures. E & FN Spon, London.

5. Hibbitt, Karlsson & Sorensen, Inc. 1998, Abaqus / Standard User's Manual. Version 5.8.

ANALYSIS OF DISTORTION OF A THIN-WALLED RECTANGLE BEAM USING THE THEORY OF GUIDED VLASOV BEAMS

Risto Koivula

Department of Mechanical Engineering
Lappeenranta university of Technology
PL 20, 53851 Lappeenranta

ABSTRACT

The coupled deformations of distortion, torsion and warping of a beam with a thin-walled closed rectangular cross-section can be analysed by dividing the beam into four guided Vlasov beams on elastic foundation according to Figure 1.

Due to the double symmetry, the stresses caused by the normal force and the bending moments are orthogonal to those caused by the warping and the distortion.

The four identical part profiles are angle cross-sections with a non-compressible axis along both edges, figure 2. These are due to the symmetry properties on the total beam. The direction of the deflection of the torsional axis of the part beams is guided by a lateral plane support. The in-plane deflection is supported by a spring, figure 3. The spring force, due to the plate bending of the walls, does not pass through the guided torsional axis on the part beam. So the torsion always gives rise to the distortion and on the contrary, except in the case when the cross-section is such that torsion does not give rise to normal stresses, for instance the square cross-section with a constant wall thickness. The last case under distorting load is solved analytically by Asko Kähönen and Erkki Niemi. /3/

Guided Vlasov beams with more than one non-compressible axes constitute an intermediate link between the theories of open and closed thin-walled cross-sections.

KEY WORDS

Thin-walled beams; Vlasov beams; guided beams; distortion; warping torsion; bimoment; torsional axis, extensional axis

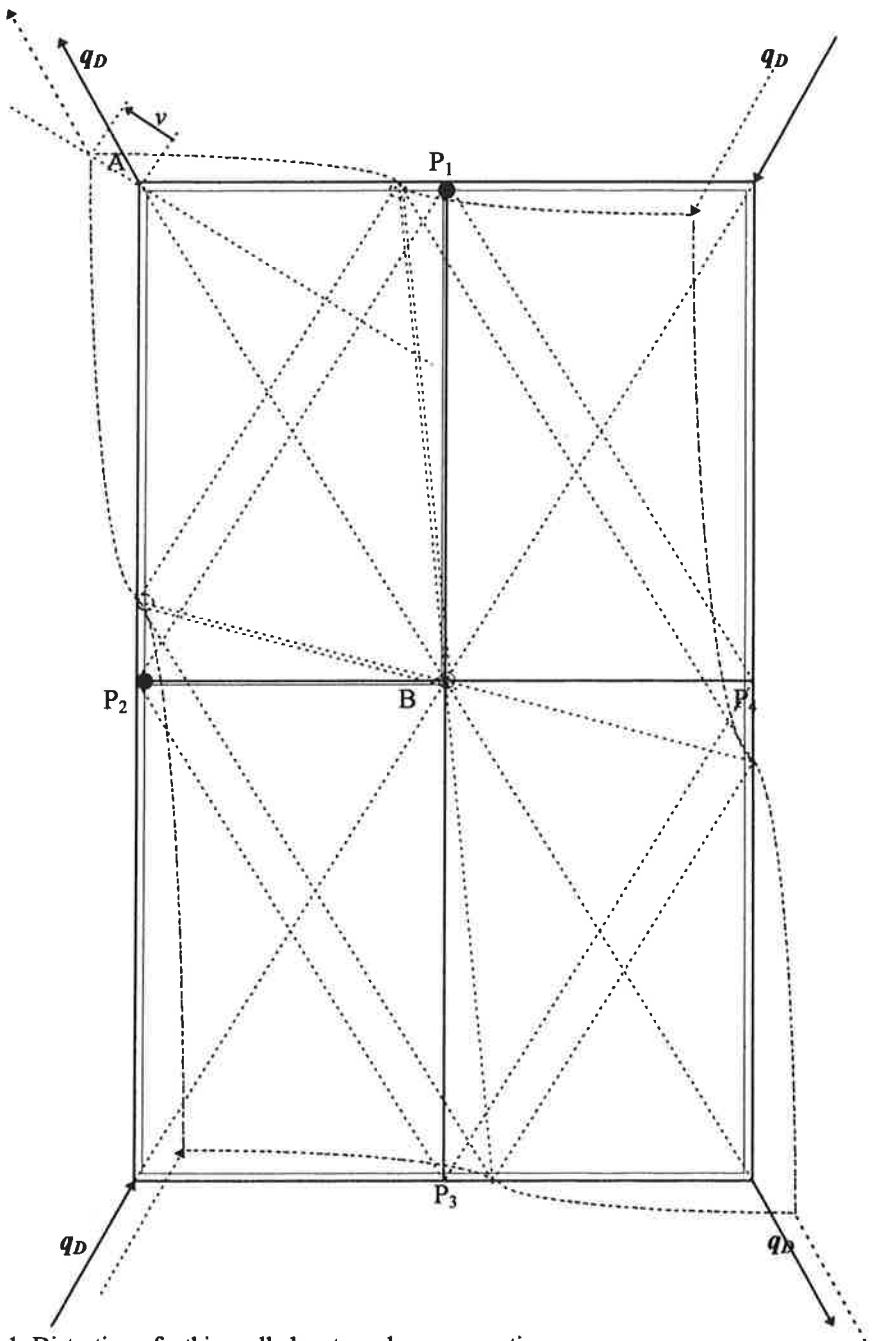
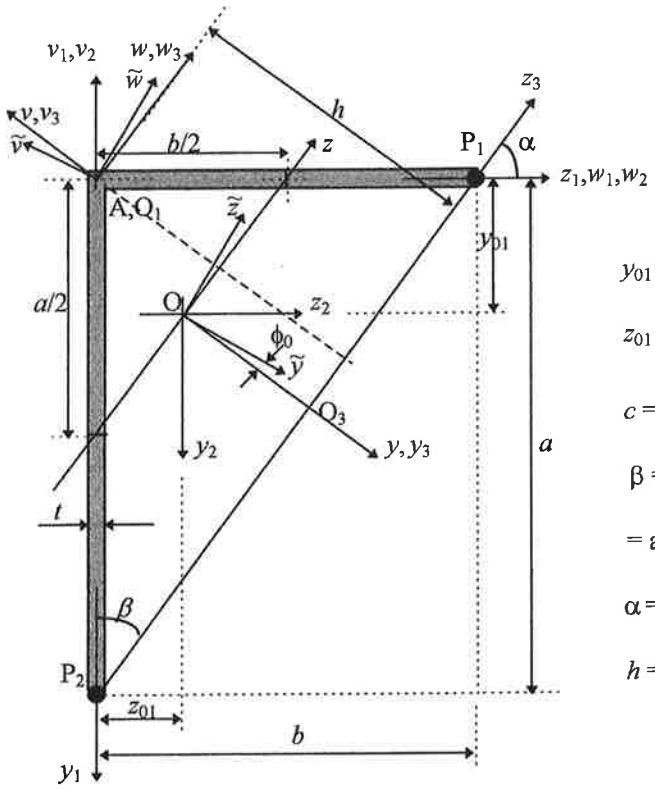


Figure 1. Distortion of a thin-walled rectangular cross-section

The mid-axes P_i of the webs are non-compressible hinges. The hinged "plane supports" BP_i are rigid. The "deflection" v of the part beam is normal to the line P_1P_2.

Angle cross-section, guided by non-compressible axes at the edges

Our basic element is an angle cross-section, the deformation of which is restricted by a non-compressible axis, a continuous axial support at both edges. Its degrees of freedom, until further notice neglecting the shear deformation, are bending around the line P_1P_2, the deflection being normal to this line, and the angle of twist around the torsional axis, the corner axis A. The indices refer to corresponding systems of axes and special points. As to the transformations between these systems I refer to the reference /1/.



$$y_{Q1} = a^2/(a+b)/2 = S_{w1}/A \quad (1)$$

$$z_{Q1} = b^2/(a+b)/2 = S_{v1}/A \quad (2)$$

$$c = \sqrt{a^2 + b^2} \quad (3)$$

$$\beta = \arctan(b/a) = \arcsin(b/c)$$

$$= \arccos(a/c) = \pi/2 - \alpha \quad (4)$$

$$\alpha = \arcsin(a/c) = \arccos(b/c) \quad (5)$$

$$h = b\sin\alpha = a\cos\alpha = ab/c \quad (6)$$

Figure 2. An angle cross-section with non-compressible axes at the edges

The displacement and force quantities like deflections $v(x)$ and $w(x)$, rotations $\theta_v(x) = -w'(x) = -\,dw(x)/dx$ and $\theta_v(x) = v'(x)$, bending moments and shear forces are given in a right-handed (anti-clockwise) system of axes, but the measure quantities like x, y, z and the sectorial area $\omega(s)$ in a left-handed (clockwise) system. Then some important matrices are symmetric, and the force and displacement quantity "vectors" $\{F\}$ and $\{u\}$ are in the same time without minus signs.

In figure 2 five systems of axes and special points are shown. Each of them has the sectorial pole at point A.

- $AO_1xy_1z_1$, the axes of which coincide with the mid-lines of the webs, is a generic system, in which the cross-section integrals are easy to calculate.
- $AOxy_2z_2$ is a parallel to the web mid-lines centroidal system
- $AOx\tilde{y}\tilde{z}$ is the fundamental system with the origin at the centroid and axes parallel to the

principal directions. In this case it has not much use.
- AOxyz is a centric system with the axes parallel to those of the solving system. The guided (eccentric) special points are mostly determined using this system.
- $AO_3xy_3z_3$ is the solving system, in which the effective degrees of freedom are mutually orthogonal and the force quantities can be computed independently from the active loads.

For this cross-section the sectorial area at the wall mid-lines is zero, without lateral supports, for the arm $h_s(s)$ of a parallel to s unit vector $\bar{i}_s(s)$ around the pole A is zero for all s:

$$h_s(s) \equiv 0\ ,\ \tilde{\omega}(s) \equiv 0 \tag{7}$$

This fulfils at the non-compressible axes the condition $\tilde{\omega}(s_1) = \tilde{\omega}(s_2) = 0$. For the solving system $AO_3xy_3z_3$ the pole is still at the corner A and sectorial area is zero:

$$a_{y3} - a_{y0} = 0\ ,\ a_{z3} - a_{z0} = 0\ ,\ \omega_3(s) \equiv 0 \tag{8}$$

The guided origin of the solving system is placed at the intersectional point O_3 of the straight line between the non-compressible axes with the normal drawn to this line through the centroid. In this system the cross-sectional integrals are expressed below. The subscript zero ($_0$) refers to the centroidal system AOxyz, parallel to the solving system.

$$I_{v3} = I_{v0} = \frac{t(a^6 + 4a^5b + 6a^3b^3 + 4ab^5 + b^6)}{12(a^2 + b^2)(a+b)} \tag{9}$$

$$I_{w0} = \frac{ta^2b^2(a+b)}{12(a^2+b^2)} \tag{10}$$

$$I_{w3} = I_{w0} + Ah^2/4 = \frac{ta^2b^2(a+b)}{12(a^2+b^2)} + \frac{ta^2b^2(a+b)}{4(a^2+b^2)} = \frac{ta^2b^2(a+b)}{3(a^2+b^2)} = \frac{ta^2b^2(a+b)}{3c^2} \tag{11}$$

$$I_{vw3} = I_{vw0} = \frac{tab(a^3 - b^3)}{12c^2}\ ,\quad S_{w3} = -Ah/2 = -\frac{tab(a+b)}{2c}\ ,\quad S_{v3} = 0 \quad (12),(13),(14)$$

The model of the quarter of the box beam as a guided Vlasov beam on elastic foundation

The functioning of the upper left corner part of the box beam of figure 1 with its effective degrees of freedom can be modelled as a guided Vlasov beam in figure 3. The active degrees of freedom of the cross-section are the deflection of some pole A_5 of the axis w_5 and the angle of twist . In the reality the centre B of the box beam, in figure 3 the point A_4, has no deflections due torsion or distortion, but in the analogy with the Vlasov BEF (beam on elastic foundation) it has a motion $v_4(x)$ normal to the guided neutral line w_4 (P_1P_2). In direction w it is fixed both in the model and the reality. The corner A, the concentration axis of the loads q_D and q_T, has deflections as if the part beam with rigid-in-plane cross-section were supported according to figure 3.

If the shear deformation $\gamma^*(x)$ were neglected at the profile line like usually is supposed for the usual sectorial area

$$\gamma_{xs}(x,s) = 0\ , \tag{15}$$

it would not be possible for a beam of figure 3 to twist at all, except if the parallel to P_1P_2 plane support passed through corner A of the cross-section. That is to say it is not possible to draw any usual sectorial area $\omega_5(s) - \omega_5(s_{01})$ around any pole axis A_5 in the support plane w_5 so that it were zero in the same time at both non-compressible axes:

$$u(x,s_{01}) = u(x,s_{02}) = 0 \tag{16}$$

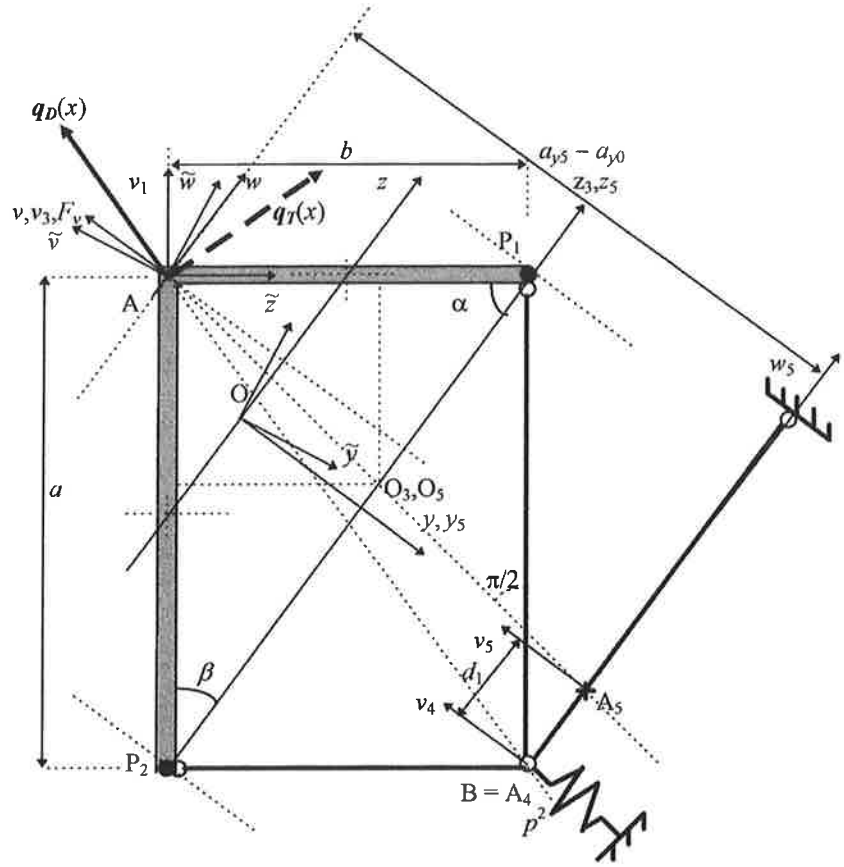


Figure 3 Analogous model for a quarter segment of a box beam under torsion and distortion
The spring acting in direction v_4 is due to the plate bending of the walls. $q_D(x)$ is a distorting and $q_T(x)$ a twisting load. There influences are connected by the spring.

In this case equation (15) must be replaced by an equation taking into account the shear deformation from constant shear stress flow

$$\gamma_{xs}(x,s) = \frac{\partial u(x,s)}{\partial s} + \frac{\partial v_s(x,s)}{\partial x} = \gamma^*_{zs}(x,s) = \frac{\tau^*_{zs}(x,s)}{G} = \frac{T^*_s(x)}{Gt(s)} \tag{17}$$

$T^*_s(x)$ is the constant shear stress flow

The warping gives now rise to axial displacement distribution

$$u(x,s) = -\int_{s_{01}}^{s} [\theta'_u(x) h_{s5}(s) - \frac{T^*_s(x)}{Gt(s)}] ds = \theta_u{}'(x)[\omega_5(s) - \omega_5(s_{01})] + \frac{T^*_s(x)}{tG}(s - s_{01}) \tag{18}$$

Equation (16) is now valid, if

$$T^*_s(x) = -Gt \frac{\omega_5(s_{02}) - \omega_5(s_{01})}{s_{02} - s_{01}} \theta'_u(x), \tag{19}$$

if the wall thickness $t(s) \equiv t$ is constant. This shear stress flow results now in a torque corresponding the StVenant torque of closed beams:

$$M_{u4}^{*}(x)=T_s^{*}(x)\int_S h_5(s)\mathrm{d}s=-T_s^{*}(x)\left[\omega_5(s_{02})-\omega_5(s_{01})\right] \tag{20}$$

$$=G\frac{[\omega_5(s_{02})-\omega_5(s_{01})]^2}{s_{02}-s_{01}}\theta_u'(x)=GI_t^{*}\theta_u'(x) \tag{21}$$

For a thin-walled cross-section with constant t the StVenant torsional stiffness moment

$$I_t^{*}=\frac{t[\omega_5(s_{02})-\omega_5(s_{01})]^2}{|s_{02}-s_{01}|}=\frac{4tA_\omega^2}{S}=\frac{4ta^2b^2}{a+b}=I_t \tag{22}$$

in which S is the arc length between the non-compressible axes and A_ω is the area enclosed by the profile line and the lateral restraint planes. This formula coincides with the formula for torsion of closed beams.

The usual sectorial area from P_1 to P_2 is equal around any sectorial pole A_5 of the axis w_5 in figure 3:

$$\omega_5(s_{02})-\omega_5(s_{01})=2ab=2ch\,, \tag{23}$$

because this sectorial area equals with double the area swept by the pole radius which for points on w_5 consist of two triangles with the same height h and same base c. The sectorial area taking in attention the shear deformation effect is then

$$\omega_5^{*}(s)=\omega_5(s)-\omega_5(s_{01})-\frac{s-s_{01}}{s_{02}-s_{01}}[\omega_5(s_{02})-\omega_5(s_{01})] \tag{24}$$

Placing the pole A_5 to the intersectional point of the restraint plane and the bisector of the right angle A, the sectorial area vanishes at every s:

$$\omega_5^{*}(s)\equiv 0 \tag{25}$$

This is a well defined sectorial area with a unique pole and unique value for all s. It means that torsion around this pole is StVenant torsion and the bimoment $B_5(x)\equiv 0$ is orthogonal with the bending moments. So the axis A_5 is the torsional axis of the construction.
Now we can form the essential in the theory of guided beams definition matrix of normal stress force quantities

$$\begin{Bmatrix} F_u \\ M_{v5} \\ M_{w5} \\ B_5 \end{Bmatrix}=E\begin{bmatrix} A & 0 & S_{w3} & 0 \\ 0 & I_{v0} & -I_{vw0} & 0 \\ S_{w3} & -I_{vw0} & I_{w3} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\begin{Bmatrix} 0 \\ 0 \\ \theta_{w5}' \\ \theta_u''' \end{Bmatrix} \tag{26}$$

$$M_{w5}(x)=EI_{w3}\theta_{w5}'(x)=EI_{w3}v_5''(x) \tag{27}$$

According to the definition, the shear force normal to the neutral axis w_3 is obtained from

$$F_v(x)=\iint_A \tau_{xs}(x,s)\cos(\bar{i}_x,\bar{i}_s)\mathrm{d}A=-\int_S \tau_{xs}(x,s)\frac{\mathrm{d}y(s)}{\mathrm{d}s}t(s)\mathrm{d}s$$

$$=-\int_{s_1}^{s_2} t(s)\tau_{xs}(x,s)[y(s)-y_{03}]+\int_S \frac{\partial[t(s)\tau_{xs}(x,s)]}{\partial s}[y(s)-y_{03}]\mathrm{d}s$$

$$=-\iint_A \frac{\partial\sigma_x(x,s)}{\partial x}t(s)[y(s)-y_{03}]\mathrm{d}s=-M_{w1}'(x) \tag{28}$$

In equation (28) is used the equilibrium condition of a wall element d*x*d*s* in absence of surface or volume loadings:

$$t(s)\frac{\partial \sigma_x(x,s)}{\partial x}+\frac{\partial[t(s)\tau_{xs}(x,s)]}{\partial s}=0 \tag{29}$$

The axial equilibrium condition of a beam element d*x* with equation (27) gives

$$EI_{w3}v_5^{(4)}(x)=M_{w5}''(x)=F_v(x)=-q_v(x) \tag{30}$$

The presupposition of eq. (30) to be valid is, that at the edges either the shear stresses $\tau_{xs}(s)$ or the unit warping $y(s)-y_{03}$ (in terms of the generalized beam theory of R. Schardt) is zero at the edges. Then this force quantity of the solving system is determined by external loadings only. For the StVenant torsion

$$M_u^*(x)=GI_t\theta_u'(x) \tag{31}$$

$$GI_t\theta_u''(x)=M_u^{*\prime}(x)=-m_5(x) \tag{32}$$

When to this system at point A_4 is applied a spring resisting the deflection $v_4(x)=v_5(x)+d_1\theta_u(x)$ with a spring constant p^2, for displacement quantities $v_5(x)$ ja $\theta_u(x)$ is obtained a coupled group of differential equations

$$\begin{cases} EI_{w3}v_5^{(4)}(x)=q_v(x)-p^2[v_5(x)+d_1\theta_u(x)] & (33)\\ -GI_t\theta_u''(x)=m_5(x)-p^2d_1[v_5(x)+d_1\theta_u(x)]\text{ , eli} & (34)\end{cases}$$

$$\begin{bmatrix} EI_{w3}D^4+p^2 & p^2d_1 \\ p^2d_1 & -GI_tD^2+p^2d_1^2\end{bmatrix}\begin{Bmatrix} v_5(x)\\ \theta_u(x)\end{Bmatrix}=\begin{Bmatrix} q_v(x)\\ m_5(x)\end{Bmatrix} \tag{35}$$

$$d_1=\frac{a-b}{a+b}c \tag{36}$$

The determinant of the differentiation matrix is

$$(Ei_{w3}D^4+p^2)(-GI_tD^2+p^2d_1^2)-p^4d_1^2=-D^2(EI_{w3}GI_tD^4-EI_{w3}p^2d_1^2D^2+GI_tp^2) \tag{37}$$

$$D^2(D^4-q^2d_1^2D^2+q^2k^2)=0 \tag{38}$$

$$k^2=\frac{GI_t}{EI_{w3}}=\frac{1}{2(1+\nu)}\frac{4ta^2b^2}{a+b}\frac{3(a^2+b^2)}{a^2b^2(a+b)}=\frac{6}{1+\nu}\frac{c^2}{(a+b)^2} \tag{39}$$

$$q^2=\frac{p^2}{GI_t} \tag{40}$$

The roots of equation (38) are

$$D_1=D_2=0\,,\quad D_{3,4,5,6}=\pm\sqrt{q^2d_1^2\pm\sqrt{q^4d_1^4-4k^2q^2}}\,/\sqrt{2} \tag{41}$$

The spring constant can be calculated for instance using the Mohr integrals. At the corner A of the cross-section the plate bending moment in obtained from

$$m_{sA}(x)=q_D(x)ab/c \tag{42}$$

The plate stiffness constant of the wall is

$$K = \frac{Et^3}{12(1-\nu^2)} \tag{43}$$

The displacement in direction F_D is then

$$\delta_D = \frac{4(1-\nu^2)q_D a^2 b^2 (a+b)}{Et^3(a^2+b^2)} = q_D I_{w3}/(tK)\,, \tag{44}$$

from which the spring constant

$$p^2 = q_D/\delta_D = tK/I_{w3} \tag{45}$$

The twisting component q_T of the corner load does not bend the walls, but it has a component bending the beam. So the distorting component has the direction q_D. The spring constant calculated in direction q_D is valid in direction v, the w-component of $q_D - p^2\delta_D$ constitutes the reaction force in the lateral support.

The bending load $q_v(x)$ and the twisting moment load $m_5(x)$ are obtained form the corner forces:

$$q_v(x) = q_D(x)\cos(\alpha-\beta) - q_T(x)\sin(\alpha-\beta) \tag{46}$$

$$q_w(x) = q_D(x)\sin(\alpha-\beta) + q_T(x)\cos(\alpha-\beta) \tag{47}$$

$$m_5(x) = q_v(x)d_1 - q_w(x)2h = -q_D(x)d_1\cos(\alpha-\beta) - q_T(x)2h \tag{48}$$

Normal stresses are obtained from the bending moment of the solving system.

$$\sigma_x(x,s) = \frac{M_{w5}(x)}{I_{w3}}y_3 = \frac{M_{w5}(x)}{I_{w3}}(y-y_{03}) \tag{49}$$

Discussion

The idea of the theory of guided beams is the fact that for thin-walled beams, the deformations of which are restricted by external continuous lateral or longitudinal restraints, the mutually independent effective force quantities, which are determined by only external loads, are defined in a **solving system of axes and special points**, which does not coincide with the **fundamental system** with the origin at the centroid, the pole at the shear centre, and the axes parallel to the (central) principal directions. Besides they can be solved, they even ought to be solved in the solving system, if they are solved right. The choice of the correct system is analogous to the separating correctly a group of mutually dependent differential equations to mutually independent equations.

The group of equations (33) is a model of a Vlasov BEF beam. It has solutions in elementary functions, if the load deviations are simple enough functions of x. Maybe this makes it possible to develop for a thin-walled rectangle beam an exact stiffness matrix including the torsional and distortional degrees of freedom, which are mutually coupled, because the line of influence of the spring reaction does not pass through the guided torsional axis A_5 of the segment cross-section. If $d_1 = 0$, as for a square cross-section the wall thickness t being constant, the group is separated into two separate differential equations.

The theory of guided beams is now extended to the field of closed thin-walled cross-sections. The cases with two or several non-compressible axes constitute an intermediate link between the theories of open and closed thin-walled cross-sections.

References:

/1/ Koivula Risto: Avoprofiilisen sauvan ohjattu vääntö (Guided torsion of Vlasov beams, in Finnish), Rakenteiden mekaniikka, Vol. 18 No 4 1985, pp 57 - 94

/2/ Koivula Risto: Transformations of force and displacement quantities between different systems of axes and special points, Proceeding of 2nd International Conference of Thin-Walled Structures, Singapore 1998

/3/ Kähönen, A., Niemi, E.: Distortion of a double symmetric box section subjected to eccentric loading using the beam on elastic foundation approach, Lappeenranta University of Technology 1986

DYNAMIC BUCKLING OF ELASTIC-PLASTIC BEAMS

ÜLO LEPIK
Institute of Applied Mathematics
Tartu University, 51014 Tartu, ESTONIA

ABSTRACT

Buckling of axially compressed elastic-plastic beams is discussed. The load is applied instantaneously and remains unaltered during the motion. The effect of stress waves travelling along the beam is taken into account. It is assumed that the material of the beam has linear-strain hardening. A method of solution, based on the Galerkin technique, is proposed; this method is applicable to an arbitrary number of degrees of freedom. Numerical examples are presented.

1. INTRODUCTION

Elastic-plastic buckling of beams under dynamic loads is an important problem in dynamics. Usually it is assumed that the beam is instantaneously brought to the state of uniform compressive stress (the so-called uniform thrust theory, see e.g. Lindberg and Florence [1]). This problem has been a subject of extensive research; the results are presented in many text-books about structural dynamics.

If we want to examine the initial phase of buckling the effect of travelling axial waves must be taken into account. This brings about an interaction of longitudinal and transverse waves. For finite length beams a rather complicated state of stress is created because of reflections of various waves. The effect of a travelling axial wave on the buckling of elastic beams was discussed by Sevin [2] and Lindberg [3]. Lindberg carried out also some camera observations of dynamic elastic buckling. An experimental study of dynamic plastic buckling was reported by Abrahamson and Goodier [4]; they also discussed some aspects of interaction between axial and flexural waves. A theoretical study on elastic buckling forms of elastic beams under impact loading was considered by Kornev [5], who showed that solutions appear growing exponentially with time; his solution holds within the time required for the elastic wave to transverse the bar. Making use of the quasi-bifurcation method Lee [6] studied the effect of travelling elastic-plastic waves on the buckling phenomenon.

In nonlinear dynamics for integrating the equations of motion usually different variants of the finite-element method have been applied. Modern FEM codes treat also axial-wave lateral-bending interactions in the range beyond strain-rate reversal. Nevertheless it seems to us that to develop alternative codes could be reasonable. In the present paper a Galerkin-type solution is proposed. The advantage of this technique is that the modal structure of the solution is maintained. This allows us to follow how the amplitudes of different modes grow in time; useful information can also be obtained from modal interactions. So the Galerkin type solutions enable us to provide deeper understanding of the problems which are usually difficult to see in stricktly numerical results. For the solution of elastic problems the Galerkin method has turned out to be a valuable tool, but not much has been done in the case of plastic deformations.

In this paper early buckling process for a beam under an instantaneously applied load is examined; it is assumed that the load is constant in time and of infinite duration. The end-sections of the beam are simply supported. The material of the beam is linear strain-hardening. The axial displacement is calculated according to the method of characteristics. For determining the flexural motion we shall proceed from the Hamilton principle and making use of the Galerkin procedure separate the variables. The method is applicable to an arbitrary number of degrees-of-freedom. The applicability and efficiency of the recommended technique is demonstrated by means of numerical examples.

2. PROBLEM STATEMENT

Let us consider a beam with rectangular cross-section; B, h and L are the width, thickness and height of the beam, respectively. The x^*-axis is directed along the beam and its origin is located at one end-section of the beam. At $x^* = 0$ a sudden constant axial load of infinite duration P^* is applied. It is assumed that P^* exceeds the critical load at which bifurcation takes place. The beam has an initial imperfection $w_0^* = w_0^*(x)$, which is regarded to be small (Fig. 1).

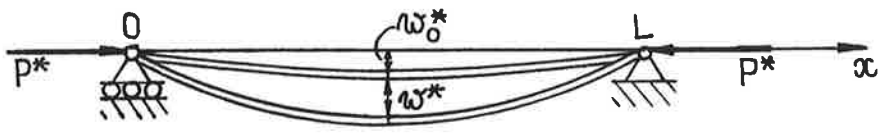


Fig. 1. Beam dimensions.

The following notations are introduced: t^*-time, ϱ-density, u^*-axial displacement, w^*-deflection, σ^*-stress, e-strain, E-Young's modulus, σ_s-yield stress, T^*-axial force, M^*-bending moment, $e_s = \sigma_s/E$, $\gamma = h/L$. Instead of these quantities it is convenient to use the following nondimensional parameters:

$$x = \frac{x^*}{L}, \quad z = \frac{z^*}{h}, \quad t = \frac{t^*}{L}\sqrt{\frac{E}{\varrho}}, \quad u = \frac{u^*}{L}, \quad w = \frac{w^*}{h}, \quad w_0 = \frac{w_0^*}{h},$$
$$\sigma = \frac{\sigma^*}{\sigma_s}, \quad P = \frac{P^*}{\sigma_s B h}, \quad T = \frac{T^*}{\sigma_s B h}, \quad M = \frac{4M^*}{\sigma_s B h^2}. \tag{1}$$

We shall start from Hamilton principle

$$\int_{t_0^*}^{t_1^*} (\delta K^* - \delta U^* + \delta W')\, dt^* = 0, \tag{2}$$

where (δ-is the symbol of isochronic variation)

$$\delta K^* = B \int_0^L \int_{-h/2}^{h/2} \varrho \left[\frac{\partial u^*}{\partial t^*} \delta\left(\frac{\partial u^*}{\partial t^*}\right) + \frac{\partial w^*}{\partial t^*} \delta\left(\frac{\partial w^*}{\partial t^*}\right) \right] dx^* dz^*,$$
$$\delta U^* = B \int_0^L \int_{-h/2}^{h/2} \sigma^* \delta e \, dx^* dz^*,$$
$$\delta W' = P^*(t^*) \delta u^*(0, t^*).$$

Let us confine our analysis to the Bernoulli-Euler beam for which

$$e = \varepsilon - z^* \frac{\partial^2 w^*}{\partial x^{*2}} \tag{3}$$

and

$$\varepsilon = \frac{\partial u^*}{\partial x^*} + \frac{1}{2}\left[\left(\frac{\partial w^*}{\partial x^*} + \frac{\partial w_0^*}{\partial x^*} \right)^2 - \left(\frac{\partial w_0^*}{\partial x^*} \right)^2 \right].$$

After some transformations the variational equation (2) can be rewritten in the form

$$\int_{t_0}^{t_1} \left\{ \int_0^1 \left[\frac{\ddot{u}}{e_s} \delta u + T \delta u' + \gamma^2 \left(\frac{\ddot{w}}{e_s} \delta w + T(w' + w_0') \delta w' - \right.\right.\right.$$
$$\left.\left.\left. - \frac{1}{4} M \delta w'' \right) \right] dx - P \delta u(0, t) \right\} dt = 0. \tag{4}$$

where

$$T = \int_{-0.5}^{0.5} \sigma \, dz, \quad M = 4 \int_{-0.5}^{0.5} \sigma z \, dz. \tag{5}$$

For the sake of concreteness we shall consider the following boundary conditions:
(i) the beam is simply supported, so that $w = M = 0$ for $x = 0$ and $x = 1$;
(ii) the end-section $x = 1$ is axially restrained: $u(1) = 0$.

3. PROPAGATION OF AXIAL WAVES

Let us assume that the beam material has linear strain-hardening (Fig. 2).

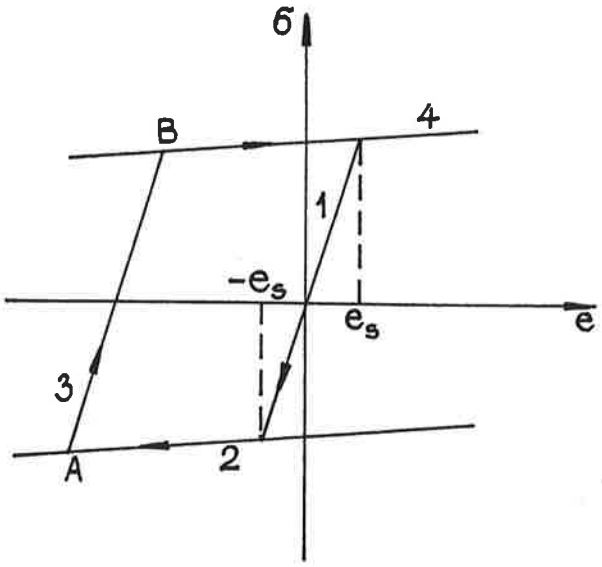


Fig. 2. Stress-strain diagram with linear strain-hardening.

In this case the strain-hardening parameter $\lambda = 1 - (1/E)(d\sigma^*/de)$ is constant; limit cases $\lambda = 0$ and $\lambda = 1$ correspond to elastic and perfectly plastic materials, respectively. The speed of the elastic and plastic wave fronts is

$$c_0^* = \sqrt{\frac{E}{\varrho}}, \quad c_1^* = \sqrt{\frac{E(1-\lambda)}{\varrho}}. \tag{6}$$

In the following we shall use nondimensional speeds, which are defined as

$$c_0 = 1, \quad c_1 = \frac{c_1^*}{c_0^*} = \sqrt{1-\lambda}. \tag{7}$$

For calculating the strains in the beam we use the method of characteristics. Since this method is generally known, here only final results are given. The characteristic relations are

$$\begin{aligned} v + c_j e &= const \quad \text{for } x = c_j t, \\ v - c_j e &= const \quad \text{for } x = -c_j t, \ (j = 0, 1) \end{aligned} \tag{8}$$

Here $v = v^*/c_0^*$ denotes nondimensional particle velocity. Thus it is suitable to note that $t = 1$ is the nondimensional time required for the elastic wave to transverse the

beam. Since we have assumed that $P(t) = const$, also $e_0 = e(0,t) = const$. The xt-plane is now divided into regions in which v and e have constant values calculated from the characteristic relations of Eqs. (8). The three following cases can be distinguished (numerical subscript in the symbols e and v denotes the number of the region in Fig. 3).

Case 1. Here $|e_0| < 0.5e_s$. This is the elastic case; when the wave front OA is reflected from the boundary $x = 1$ the amplitude of the reflected wave is doubled (Fig 3a). It follows from (8) that

$$e_1 = v_1 = 0, \quad e_2 = e_0, \quad v_2 = -e_0, \quad e_3 = 2e_0, \quad v_3 = 0,$$
$$e_4 = e_0, \quad v_4 = e_0, \quad v_5 = e_5 = 0, \quad e_6 = -e_0, \quad v_6 = -e_0.$$

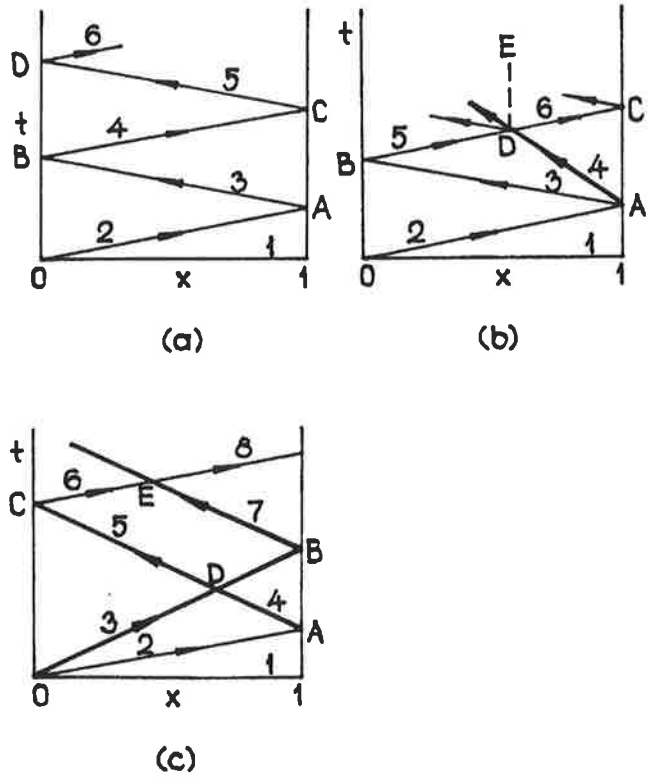


Fig. 3. Position-time diagram of wave fronts in the beam. Thin lines denote elastic wave fronts and thick lines plastic wave fronts; (a) $|e_0| < 0.5e_s$, (b) $0.5e_s < |e_0| < e_s$, (c) $|e_0| > e_s$.

Case 2. If $0.5e_s < |e_0| < e_s$ then the elastic wave front OA is reflected at $x = 1$ as an

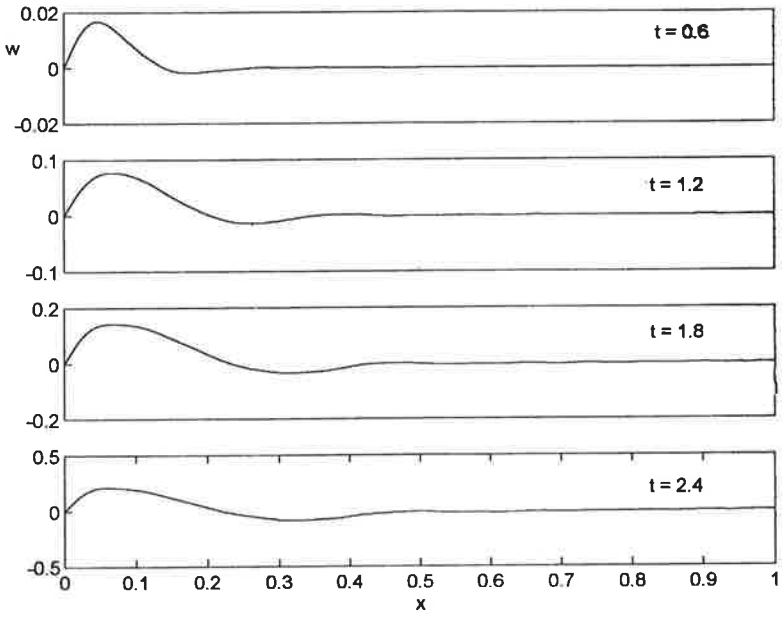


Fig. 4. Deflection profiles.

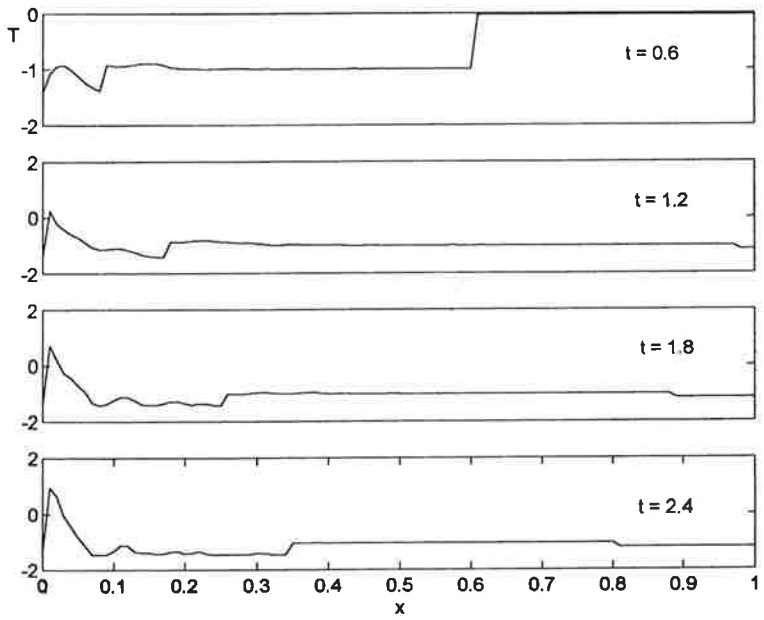


Fig. 5. Axial stress profiles.

It follows from Fig. 4 that the biggest amplitudes of the buckled form appear near the end-section $x = 0$ (this is the effect of local buckling which is well known from observations and experiments).

In many papers it has been assumed that the axial force $T(x)$ is constant along the beam. As it follows from Fig. 5 this assumption does not hold in the present case. It should be mentioned that this case corresponds to the "fast loading". For "slow loading" the axial inertia forces can be neglected, consequently $\ddot{u} = 0$ and it follows from the equation of motion that $T' = \ddot{u} = 0$; consequently $T(x)$ would be a piecewisely constant function. This case of "slow loading" is discussed in the paper by Lepik [7].

More details about our method of solution and further examples can be found in the paper [8].

REFERENCES

1. H.E. Lindberg and A.L. Florence. *Dynamic pulse buckling-theory and experiment.* Defence Agency, Washington, Contract No DNA 001-78-C-0287 and Martinus Nijhof Publishers Norwell. MA (1987).
2. E. Sevin. On the bending of columns due to dynamic axial forces including effects of axial inertia. *ASME Trans.* **82**, *Ser. E, J. Appl. Mech.* **27**, 125-131 (1960).
3. H.E. Lindberg. Impact buckling of a thin bar. *ASME Trans.* **87**, *Ser. E, J. Appl. Mech.* **32**, 315-322 (1965).
4. G.R. Abrahamson and J.N. Goodier. Dynamic flexural buckling of rods within an axial plastic compressive wave. *ASME Trans.* **88**, *Ser. E, J. Appl. Mech.* **33**, 241-247 (1966).
5. V.M. Kornev. On buckling forms of elastic-plastic beams under impact (in Russian). *Prikl. Mek. Tek. Fiz.* No 3, 63-66 (1968).
6. L.H.N. Lee. Flexural waves in rods within an axial compressive wave. *Wave Motion* **3**, 243-255 (1981).
7. Ü. Lepik. On dynamic buckling of elastic-plastic beams. *Int. J. Non-linear Mech.* **35**, 721-734 (2000).
8. Ü. Lepik. Dynamic buckling of elastic-plastic beams including effects of axial stress waves. *Int. J. Impact Engng.* (in print) (2000).

KIMMOISELLA ALUSTALLA OLEVAN PALKIN TEORIAN SOVELTAMISESTA ÄÄRELLISEN PITKIEN JÄYKISTETTYJEN LAATTAKAISTOJEN ANALYSOINTIIN

Teuvo Partanen

Lappeenrannan Teknillinen Korkeakoulu
Konetekniikan osasto
PL 20, 53851 Lappeenranta

TIIVISTELMÄ

Tutkimuksessa esitetään, kuinka kimmoisella alustalla olevan palkin, Beam on Elastic Foundation (BEF), teoriaa voidaan hyödyntää analysoitaessa kahdelta sivulta tuetun laattakaistan siirtymiä ja jännityksiä keskilinjatasossa. BEF parametrit määritetään asettamalla yhtä suuriksi laatta- ja puoliäärettömän BEF ratkaisuissa suurin taipuma ja suurin taivutusmomentti laatan keskilinjalla, kun peruskuormituksena on vakiosuuruinen poikittainen viivakuorma. Pitkittäissivuiltaan vapaasti ja jäykästi tuetun äärellisen pitkän laattakaistan laattaratkaisuun perustuvat keskilinjan voima- ja siirtymätulokset esitetään formuloituina BEF jäykkyysmatriisimuotoon. Ratkaisut formuloidaan niin, että tuloksia voidaan verrata kahden parametrin puoliäärettömään BEF ratkaisuun. Tulosten mukaan BEF malli kuvaa hyvin puoliäärettömän ja äärellisen pitkän laatan keskilinjan käyttäytymistä.

AVAINSANAT

Laatta, laattateoria, palkki, palkkiteoria, kimmoinen alusta, BEF, jäykkyysmatriisi.

JOHDANTO

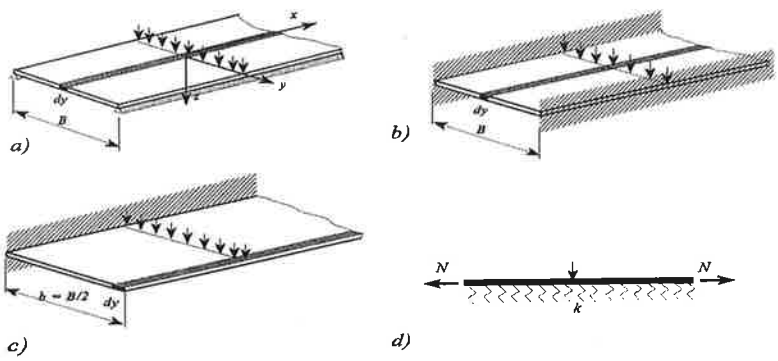
Klassinen laattateoria on matemaattisesti vaativaa. Tästä johtuen teoriaan perehtyvät syvällisemmin korkeakouluissa rakenteiden mekaniikkaan erikoistuneet ja vasta opintojen loppuvaiheessa. Sekä kone- että rakenuustekniikan suunnittelijat tarvitsevat laattarakenteita toteuttaessaan työtehtäviään. Tänä päivänä suunnittelijalla on käytettävissään yleiset FE-ohjelmistopaketit, joiden avulla voidaan mallittaa hyvinkin komplisoituja rakenteita, joiden rakenteellisesta käyttäytymisestä ei välttämättä ole kovinkaan selkeää käsitystä.

Rakenteiden mekaniikassa on palkkirakenteiden taivutusteoria lujuusopillisen ymmarryksen pohjana. Täten olisi toivottavaa, että laatta- ja kuorirakenteiden taivutuskäyttäytymistä voitaisiin ainakin likimääräisesti analysoida palkkiteoriaa hyödyntäen. Peruspalkkiteoria lisättynä aksiaalivoiman vaikutuksella, ns. toisen kertaluvun efektillä, tunnetaan nurjahdustaivutuksena. Ilmiölle matemaattisesti analoginen on Vlasovin kehittämän estetyn väännön teoriaan perustuva differentiaaliyhtälö ja sille johdetut ratkaisut. Winklerin kehittämässä kimmoisalla alustalla olevan palkin teoriassa tavallisen palkin differentiaaliyhtälöön on liitetty nollannen asteen termi, joka tuottaa alustareaktion. Viime vuosina on esitetty taivutuskäyttäytymiselle ratkaisuja ongelmissa, joissa mukana on sekä toisen että nollannen asteen termit. Teoria tunnetaan kehittäjänsä prof. Schardtin mukaan yleistettynä palkkiteoriana.

Esittävä tutkimus pohjautuu kirjoittajan väitöskirjaan [1]. Laattakaistojen käyttäytyminen poikittaissuuntaisen viivakuorman alaisena osoitettiin toimivan kimmoisalla alustalla olevan palkin, Beam on Elastic Foundation (BEF), teorian mukaisesti. Tämä esitys keskittyy kuvaamaan, kuinka ja kuinka tarkasti BEF-teoria kuvaa äärellisen pitkän laattakaistan käyttäytymistä. Tavoitteena on osoittaa, että puoliäärettömälle laattakaistalle määritettyjen BEF-parametrien avulla äärellisen laattakaistan siirtymätila ja pituussuuntaiset jännitykset voidaan määrittää FE-palkkielementillä, jossa on mukana sekä toisen että nollannen asteen termit eli yleistetyn palkkiteorian mukaiset parametrit.

BEF TEORIASTA VÄLINE LAATTOJEN ANALYSOINTIIN

Kimmoisalla alustalla olevan palkin, BEF teoria antaa vastauksen moniin äärettömyydessä vaimeneviin taivutustehtäviin, joista tämän esityksen lähteenä käytetyssä kirjoittajan väitöskirjassa [1] tarkasteltiin kuvassa 1 esitettyjen lattakaistojen käyttäytymistä. Tuloksia on esitetty lähteessä [2] ja perusteita aiemmin myös lähteessä [3].



Kuva 1. *Pituussuunnassa tuettuja laattakaistoja kuormitettuna poikittaissuuntaisella viivakuormalla: a) vapaasti tuettu, b) jäykästi tuettu, c) ulokelaatta, d) dy levyisen laattasiivun BEF malli.*

Toisena BEF teorian sovellutuksena kuoriteoriassa johdetaan taivutuskäyttäytymisen määrittäviä siirtymäratkaisuja pyörähdyssymmetrisesti kuormitetuille lieriökuorille. Vertaamalla BEF ja kuoriteorian ratkaisuja lähtökohdiltaan ja tuloksiltaan voidaan havaita niiden olevan lähes identtisiä. Pieni ero jäykkyyksissä muodostuu kuorirakenteen biaksiaalisesta jännitystilasta. Täten suunnittelija opetellessaan esim. BEF teoriaan sisältyviä formulointeja samalla löytää konkreettisia sovellutuskohteita aiemmin opettelemilleen matemaattisille valmiuksille ratkaista trigonomi- ja eksponenttifunktioiden käyttöön liittyviä tehtäviä.

Pituus-leveyssuhteeltaan kapeissa suorakaidelaatoissa puoliääretön taivutuskäyttäytyminen esiintyy hyvin monissa käytännön rakenteissa. Tästä esimerkki on mm. taivutuspalkin käyttäytyminen, kun laipassa on pieni taite, josta aiheutuu palkin pituussuunnassa suhteellisen nopeasti vaimeneva taivutusjännityshäiriö laippaan. Kun palkki on väsyttävästi kuormitettu, on jännitys- ja kestoikäanalyysi määritettävä rakenteelliseen jännitysvaihteluun pohjautuen. Olisi toivottavaa, että tarvittava jännitysanalyysi olisi suhteellisen nopea ja vaivaton. Sekä teoreettisesti että opittavaksi vaadittavien valmiuksien yhdenmukaistamisen vuoksi ideaalista olisi, jos BEF teoria voisi kyetä ainakin rajoitetusti toimimaan palkkiteoriaa, laatta- ja kuoriteoriaa yhdistävänä tukijalkana. Esimerkkitapauksena olevan alipan taitteen yksinkertaistettua BEF analysointia hyödyntäen pitäisi siis pystyä määrittämään liitoksen suurin rakenteellinen jännitysvaihtelu niin, että kestoikämääritys vastaa todellista käyttäytymistä. Liitosalueen muiden kohteiden osalta jännitysanalyysi voi tehtäväasettelun johdosta olla epätarkempi

Laatan teoreettinen käyttäytyminen

Laatan käyttäytymistä *xy*-tasossa hallitsee *z*-suuntaisen taipumafunktion differentiaaliyhtälö

$$\frac{\partial^4 w}{\partial x^4} + 2\frac{\partial^4 w}{\partial x^2 \partial y^2} + \frac{\partial^4 w}{\partial y^4} = \frac{p_z}{K}\,, \tag{1}$$

jossa p_z on laatan painekuormitus ja K laatan taivutusjäykkyys. Pitkittäisreunoiltaan vapaasti tuettujen laattakaistojen klassinen analysointi voidaan toteuttaa Levy'n menetelmällä, kun ulkoinen kuormitus lausutaan superpositiona harmonisista komponenteista. Vapaasti tuettuja laattakaistojen ratkaisuja superponoimalla voidaan ratkaista myös rotaatiojäykkyydeltään jäykästi tuetun (tai jatkuvan) laatan siirtymätila.

$$\begin{aligned} w(x,y) &= X(x)\,Y(y) = \sum_m X_m(x)\cos(\eta_m y)\,, \\ \eta_m &= \frac{m\pi}{B}\,, \quad m = 1,\ 3,\ 5\ldots \end{aligned} \tag{2}$$

Yhtälö (2) sijoittamalla yhtälöön (1) tuottaa homogeeniselle osuudelle ($p_z = 0$) kullekin arvolle m

$$X_m''''(x) - 2\eta_m^2 X_m''(x) + \eta_m^4 X_m(x) = 0 . \tag{3}$$

Puoliäärettömän kaksiparametrisen BEF elementin määritys

Kaksiparametrisen BEF palkin differentiaaliyhtälö on muotoa

$$EI\frac{d^4}{dx^4}w(x) - N\frac{d^2}{dx^2}w(x) + kw(x) = q(x) . \tag{4}$$

Puoliäärettömässä BEF palkissa siirtymät häviävät etäännyttäessä nollakohdasta. Tällöin siirtymätila on muotoa

$$w(x) = e^{-\alpha x}[A_1\cos(\beta x) + A_2\sin(\beta x)] , \tag{5}$$

jossa A_1 ja A_2 ovat tuntemattomat amplitudit. Käytetyt symbolit ovat

$$\lambda = \sqrt[4]{\frac{k}{4EI}} , \quad n = \frac{N}{N_W} = \frac{N}{4EI\lambda^2}, \quad N_W = 4EI\lambda^2 ,$$

$$\alpha = \sqrt{\sqrt{\frac{k}{4EI}} + \frac{N}{4EI}} = \lambda\sqrt{1+n} , \quad \beta = \lambda\sqrt{1-n} . \tag{6}$$

Vapausasteita U_1 ja U_2 vastaavaksi jäykkyysmatriisiksi [S] saadaan

$$[S] = EI\begin{bmatrix} 4\lambda^2\alpha & -2\lambda^2 \\ -2\lambda^2 & 2\alpha \end{bmatrix} . \tag{7}$$

Ratkaisu (5) on voimassa, kun $-1 < n < 1$, jäykkyysmatriisi (7) on aina voimassa, kun $-1 < n$.

PUOLIÄÄRETTÖMÄN LAATTAKAISTAN BEF PARAMETRIT

Harmoninen kuormitus vapaasti tuetussa laatassa

Asettamalla homogeeniset osuudet differentiaaliyhtälöissä (3) ja (4) yhtä suuriksi saadaan

$$w_m'''' - \frac{N_m}{K} w_m'' + \frac{k_m}{K} w_m = X_m'''' - 2\eta_m^2 X_m'' + \eta_m^4 X_m , \quad \eta_m = \frac{m\pi}{B} . \tag{8}$$

Asettamalla nollannen asteen ja toisen asteen termit termit laattateoreettisessa ja BEF ratkaisussa yhtä suuriksi saadaan alustaparametriksi k_m, (Winkler-parametri) ja vastaavasti toisen asteen termiksi (Pasternak-parametri) N_m

$$k_m = K\eta_m^4 = K\frac{m^4\pi^4}{B^4} , \quad N_m = 2K\eta_m^2 = 2K\frac{m^2\pi^2}{B^2} . \tag{9}$$

Kun nämä termit sijoitetaan yhtälöön (6), saadaan

$$\frac{1}{\lambda_m} = \sqrt[4]{\frac{4K}{k_m}} = \frac{\sqrt{2}}{m\pi} B , \quad n_m = \frac{N_m}{4K\lambda_m^2} = 1 ,$$
$$\frac{1}{\alpha_m} = \frac{1}{\sqrt{2}\,\lambda_m} = \frac{1}{\eta_m} , \quad \beta_m = 0 . \tag{10}$$

Kahdelta sivulta vapaasti ja jäykästi tuetut laatat [1]

BEF ratkaisussa tarvittavat kaksi parametria, $1/\lambda$ ja n, voidaan määrittää kahdesta yhtälöstä: i) suurin taipuma; ii) suurin taivutusmomentti esim. tasaisella viivakuormituksella asetetaan laatta- ja BEF ratkaisuissa yhtä suuriksi. Vapaasti tuetulle laatalle (kuva 1a) saadaan

$$\frac{1}{\lambda} = \sqrt{2}\frac{B}{\pi}\sqrt{\frac{0.989}{0.916}} = 1.47\frac{B}{\pi} = 0.468B , \quad n = 0.588. \tag{11}$$

Vastaavalla periaatteella saatiin pitkien laskutoimitusten jälkeen päistä jäykästi tuetulle laatalle (kuva 1b)

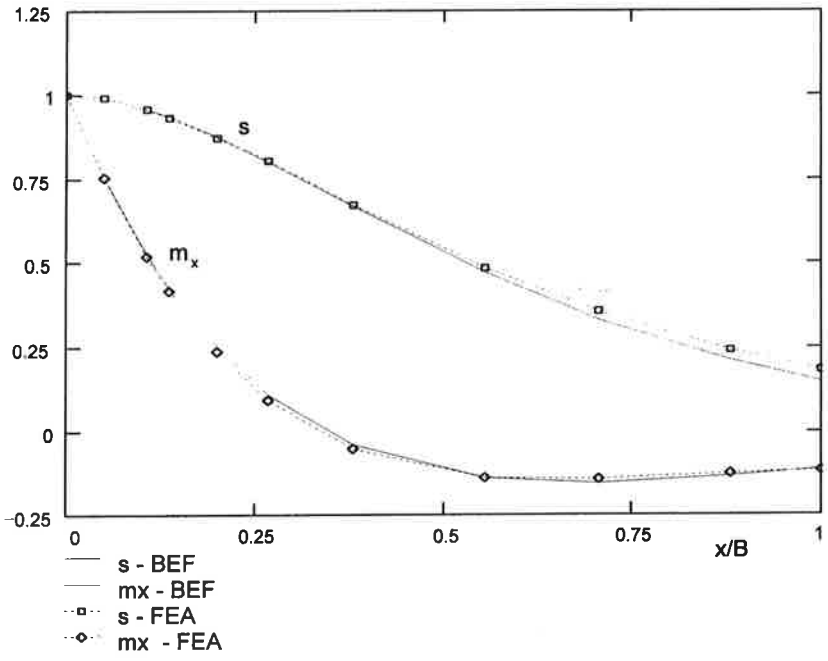
$$\frac{1}{\lambda} = 0.985\frac{B}{\pi} = 0.314B , \quad n = 0.245. \tag{12}$$

BEF mallilla taipuma ja taivutusmomenttifunktiot vastaavat suhteellisen hyvin laattateorian mukaista todellista käyttäytymistä, jota kuorielementtimallin FE analyysitulos edustaa, kuva 2. Kun pitkittäisreunojen rotaatiojäykkyys vaihtelee, osoitettiin [1], että $1/\lambda$ määritetään yhtälön (6) avulla jännevälin keskikohdan taipuman käänteisarvon perustuen, joka vastaa alustajäykkyyden arvoa k.

$$\frac{1}{\lambda} = \frac{B}{\pi}\sqrt[4]{\frac{5+r}{1+r}} \quad , \quad r = k_\varphi \frac{L}{4K} = k_\varphi \frac{B}{8K} \; . \tag{13}$$

Laatan vääntöjäykkyyteen liittyvän toisen asteen termin N osoitettiin olevan kahdelta sivulta tuetussa laatassa likimain vakio eli riippumaton päiden rotaatiojäykkyydestä k_φ . Normalisoiduksi suhteeksi n saadaan

$$n = \frac{1}{\pi}\sqrt{\frac{5+r}{1+r}} \; . \tag{14}$$



Kuva 2 *Vapaasti tuetun laatan normalisoitu taipuma, s, ja taivutusmomentti, m_x laatan keskellä $y = 0$ normalisoidun pituuskoordinaatin x/B funktiona.*

Vapaareunainen laatta [1]

Teoreettinen ratkaisu on kuvan 1c mukaiselle laattakaistalle parametrista FE laskentaa käyttäen saatiin

$$\frac{1}{\lambda} = 1.288\frac{B}{\pi} = 0.82b\,, \quad n = 0.65\,, \quad b = \frac{B}{2}\,, \tag{15}$$

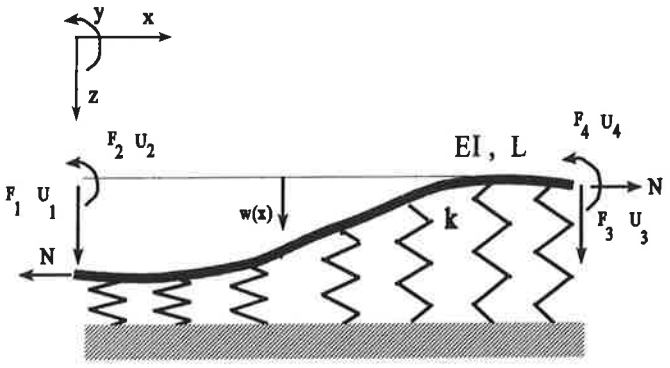
missä ulokkeen leveys b on puolet koko leveydestä B.

ÄÄRELLISEN PITKIEN BEF ELEMENTTIEN FORMULOINTI

Puhtaille Winkler-alustoille, kun $\alpha = \beta = \lambda$ kaavassa (6), FE-formulaatioita perustuen tarkalle taipuman muotofunktiolle kehittivät itsenäisesti Ting ja Mockry vuonna 1984, sekä Eisenberger ja Yankelevsky 1985. Ottaen huomioon myös toisen asteen termin N vaikutuksen Yankelevsky and Eisenberger kehittivät vuonna 1986 [4] tarkan siirtymäratkaisun yhtälölle (4) ja siitä edelleen jäykkyysmatriisin, kun normalisoitu aksiaalivoima n yhtälössä (6) oli välillä $-1 < n < 1$. Tarkempi kirjallisuuskatsaus löytyy lähteestä [1]. Seuraavassa esitetään keskeisimmät tulokset, joita voidaan käyttää neljän vapausasteen kaksiparametrista BEF elementtiä formuloitaessa. Tässä tutkimuksessa tuloksia sovelletetaan arvioitaessa BEF teorian käytettävyyttä äärellisen pitkien laattakaistojen analysointiin.

Siirtymäfunktio w(x), -1 < n < 1

Kuvassa 3 on esitetty BEF FE-elementti ja käytetty koordinaatisto.



Kuva 3 *Äärellisen pitkän BEF elementin koordinaattijärjestelmä, vapausasteet ja voimasuureet.*

Kaavan (6) symbolit on käytössä. Differentiaaliyhtälö (4) on perustana määritettäessä elementin siirtymätila ja kaksisolmuisen, neljän vapausasteen BEF elementin sisäiset voimasuureet. Vapausasteet ja voimasuureet näkyvät kuvassa 3. Muuttujan n suhteen voidaan määrittää kaksi reaaliarvoista ratkaisulohkoa:

i) $-1 < n < 1$;
ii) $n > 1$.

Ensimmäisessä tapauksessa, $-1 < n < 1$, ($-N_{cr} < N < N_{cr}$) sekä α että β yhtälössä (6) ovat reaaliset. Siirtymäfunktio $w(x)$, joka toteuttaa differentiaaliyhtälön (4), voidaan esittää hyperboli ja trigonometrisista funktioista koostuvana:

$$w(x) = A_1\cosh(\alpha x)\cos(\beta x) + A_2\cosh(\alpha x)\sin(\beta x) + \\ + A_3\sinh(\alpha x)\cos(\beta x) + A_4\sinh(\alpha x)\sin(\beta x) \ . \tag{16}$$

Vetovoiman dominoidessa, $n > 1$ ($N > N_{cr}$), β yhtälössä (6) muuttuu nollan kautta kompleksiseksi. Reaaliarvoinen ratkaisu saadaan vaihtamalla trigonometriset termit kosini and sini yhtälössä (16) hyperbolisiksi kosiniksi ja siniksi, kun muuttujaksi samalla valitaan β_1, ks. kaava (28). Kun $n > 1$, saadaan kaikki siirtymätilasta johdettavat eksplisiittiset ratkaisut tunnetuista perustapauksen $-1 < n < 1$ ratkaisuista, joista ratkaisuna esitetään kaavassa (28) jäykkyysmatriisi. Lisää ratkaisuja löytyy myös lähteestä [1]. Seuraavana keskitytään perustapauksen $-1 < n < 1$ esittämiseen.

Käytettävät erikoissymbolit, $-1 < n < 1$:

$$CH = \cosh(\alpha L), \ SH = \sinh(\alpha L) \ , \\ c = \cos(\beta L), \quad s = \sin(\beta L) \ , \tag{17}$$

Tuntemattomat siirtymäamplitudit, vektori {A}, saadaan määritettyä yksikkösiirtymämenetelmällä. Tuloksena saadaan

$$\{A\} = [E]^{-1}\{U\} = [G]\{U\} \ . \tag{18}$$

Lopulta, (tapaukselle $-1 < n < 1$), kerroinmatriisi [G] on eksplisiittisesti

$$[G] = \frac{1}{\Delta}\begin{bmatrix} \Delta & 0 & 0 & 0 \\ \frac{\alpha}{\beta}(SHCH+\frac{\alpha}{\beta}sc) & -\frac{SH^2}{\beta} & -\frac{\alpha}{\beta}(SHc+\frac{\alpha}{\beta}CHs) & -\frac{\alpha}{\beta^2}SHs \\ -(SHCH+\frac{\alpha}{\beta}cs) & \frac{\alpha}{\beta^2}s^2 & SHc+\frac{\alpha}{\beta}CHs & \frac{1}{\beta}SHs \\ -\frac{\alpha}{\beta}(SH^2+s^2) & \frac{SHCH}{\beta}-\frac{\alpha}{\beta^2}sc & (1+\frac{\alpha^2}{\beta^2})SHs & -\frac{SHc}{\beta}+\frac{\alpha}{\beta^2}CHs \end{bmatrix}$$

$$\Delta = SH^2 - \frac{\alpha^2}{\beta^2}s^2 \,. \tag{19}$$

Elementin siirtymätila on nyt tunnettu

$$w(x) = [N(x)]\{U\} = [H(x)][G]\{U\} \,, \tag{20}$$

missä [H(x)] on kaavan (16) x:stä riippuvista termeistä koostuva osuus kaavan (20) muotofunktiosta [N(x)]. Palkin kallistumiskulma, $\varphi(x)$, saadaan derivoimalla joko $w(x)$, yhtälö (16) tai [H(x)]. Kuvaamalla derivointia matriisilla [D]

$$\frac{d}{dx}[H(x)] = [H(x)]\begin{bmatrix} 0 & \beta & \alpha & 0 \\ -\beta & 0 & 0 & \alpha \\ \alpha & 0 & 0 & \beta \\ 0 & \alpha & -\beta & 0 \end{bmatrix} = [H(x)][D] \,, \tag{21}$$

$\varphi(x)$ voidaan kirjoittaa seuraavasti

$$\varphi(x) = \frac{d}{dx}[H(x)][G]\{U\} = [H(x)][D][G]\{U\} \,. \tag{22}$$

Voimasuureet saadaan edelleen derivoimalla $w(x)$:tä tai matriisitulojen [D] [D]... avulla, esim $[D^2]$ = [D] [D]. Taivutusmomentiksi saadaan siten

$$M(x) = -EI\frac{d^2}{dx^2}w(x) = -EI\frac{d^2}{dx^2}[H(x)][G]\{U\} = -EI[H(x)][D^2][G]\{U\} \tag{23}$$

Yleistetty leikkausvoima $V(x)$ on

$$V(x) = Q(x) + N\varphi(x) = -EI\frac{d^3w}{dx^3} + N\frac{dw}{dx}\,, \tag{24}$$

jossa leikkausvoima on

$$Q(x) = -EI\frac{d^3}{dx^3}w(x) = -EI\frac{d^3}{dx^3}[H(x)][G]\{U\} = -EI[H(x)][D^3][G]\{U\}. \tag{26}$$

i) Jäykkyysmatriisi, $-1 < n < 1$, $-N_{cr} < N < N_{cr}$

$$[S] = \frac{EI}{\Delta}\begin{bmatrix} 4\lambda^2\alpha(SHCH + \frac{\alpha}{\beta}sc) & S_{21} & S_{31} & S_{41} \\ -2\lambda^2(SH^2 + \frac{\alpha^2}{\beta^2}s^2) & 2\alpha(SHCH - \frac{\alpha}{\beta}sc) & S_{32} & S_{42} \\ -4\lambda^2\alpha(SHc + \frac{\alpha}{\beta}CHs) & -S_{41} & S_{11} & S_{43} \\ -4\lambda^2\frac{\alpha}{\beta}SHs & 2\alpha(SHc - \frac{\alpha}{\beta}CHs) & -S_{21} & S_{22} \end{bmatrix},$$

$$\Delta = SH^2 - \frac{\alpha^2}{\beta^2}s^2,$$

$$\alpha = \lambda\sqrt{1+n}, \quad \beta = \lambda\sqrt{1-n}, \quad n = \frac{N}{4EI\lambda^2},$$

$$SH = \sinh(\alpha L),\ CH = \cosh(\alpha L),$$
$$s = \sin(\beta L),\ c = \cos(\beta L)\ . \tag{27}$$

ii) $n > 1$, $N > N_{cr}$

$$[S_t] = \frac{EI}{\Delta}\begin{bmatrix} 4\lambda^2\alpha\left(SHCH + \frac{\alpha}{\beta_t}shch\right) & S_{21} & S_{31} & S_{41} \\ -2\lambda^2\left(SH^2 + \frac{\alpha^2}{\beta_t^2}sh^2\right) & 2\alpha\left(SHCH - \frac{\alpha}{\beta_t}shch\right) & S_{32} & S_{42} \\ -4\lambda^2\alpha\left(SHch + \frac{\alpha}{\beta_t}CHsh\right) & -S_{41} & S_{11} & S_{43} \\ -4\lambda^2\frac{\alpha}{\beta_t}SHsh & 2\alpha\left(SHch - \frac{\alpha}{\beta_t}CHsh\right) & -S_{21} & S_{22} \end{bmatrix},$$

$$\Delta = SH^2 - \frac{\alpha^2}{\beta_t^2}sh^2, \quad \alpha = \lambda\sqrt{1+n}, \quad \beta_t = \lambda\sqrt{n-1}, \quad n > 1,$$

$$SH = \sinh(\alpha L), \quad CH = \cosh(\alpha L),$$

$$sh = \sinh(\beta_t L), \quad ch = \cosh(\beta_t L). \qquad (28)$$

ÄÄRELLISEN PITKIEN JÄYKISTETTYJEN LAATTAKAISTOJEN ANALYSOINTI BEF TEORIALLA

BEF laskennan soveltuvuutta äärellisen pitkien laattakaistojen, pituus L, analysointiin testattiin tutkimalla jäykkyysmatriisin (27) toimivuutta. Vertailussa analysoitiin yhdellä äärellisen pitkällä BEF elementillä sekä vapaasti tuettuja että reunoiltaan jäykästi tuettuja laattoja, ks. kuva 1. Verifioinnissa käytettiin FE mallia, joka koostui kuorielementeistä. Palkin kohdasta $x = 0$, laatan linjalta $x = 0$ vapautettiin joko siirtymä tai rotaatio kuormituksen ollessa vastaavasti joko viivavoima tai viivamomentti. Varioiden pituutta L kohdassa $x = L$ kaikki vapausasteet kiinnitettiin. Kalvovoiman vaikutusta ei systemaattisesti tutkittu, eli tutkimus rajattiin välille $-1 < n < 1$.

BEF tulokset kuormituksen ollessa poikittainen viivakuorma

BEF ratkaisussa siirtymä U_1 voiman vaikutuskohdassa, $x = 0$, taivutusmomentit M_1 ja M_2 kohdissa $x = 0$ ja $x = L$ saadaan matriisin [S] ensimmäisen sarakkeen avulla, kaava (27), koska $U_2 = U_3 = U_4 = 0$. Taipuma U_1 pistevoimalla V_0 kohdassa $x = 0$ on kahden tekijän tulo: puoliäärettömän BEF elementin taipuma $U_{1,\infty}$, yhtälöstä (7), ja palkin äärellisestä pituudesta johtuva kerroin $u_1(\lambda L)$, joka saadaan kaavasta (27) määritettävän U_1:n avulla, eli

$$U_1 = U_{1,\infty}\, u_1(\lambda L)\ ,\quad U_{1,\infty} = \frac{V_0}{4EI}\frac{1}{\lambda^2\alpha}\ ,\quad u_1(\lambda L) = \frac{SH^2 - \frac{\alpha^2}{\beta^2}s^2}{SHCH + \frac{\alpha}{\beta}sc}\ . \tag{29}$$

Taivutusmomentiksi M_1 kohdassa $x = 0$ saadaan

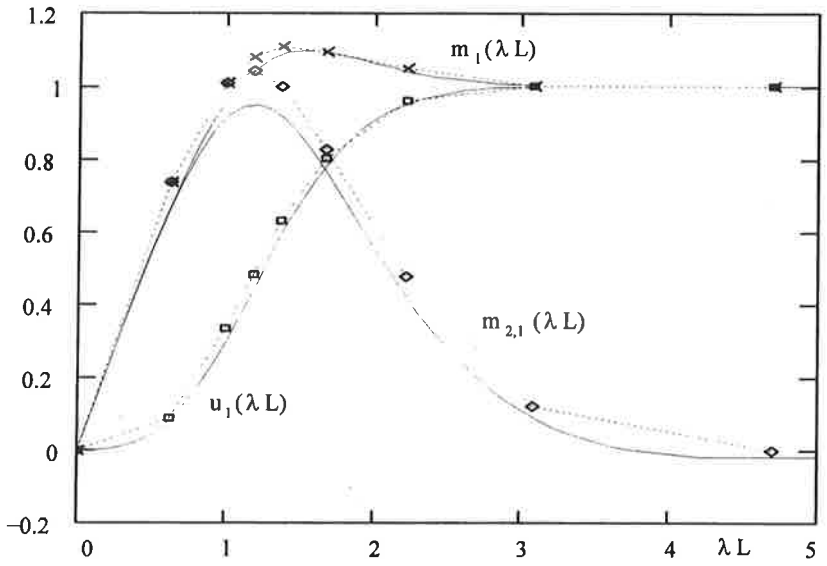
$$M_1 = M_{1,\infty}\, m_1(\lambda L)\ ,\quad M_{1,\infty} = -V_0\frac{1}{2\alpha}\ ,\quad m_1(\lambda L) = \frac{SH^2 + \frac{\alpha^2}{\beta^2}s^2}{SHCH + \frac{\alpha}{\beta}sc}\ . \tag{30}$$

Taivutusmomentti M_2 kohdassa $x = L$ on vastaavasti

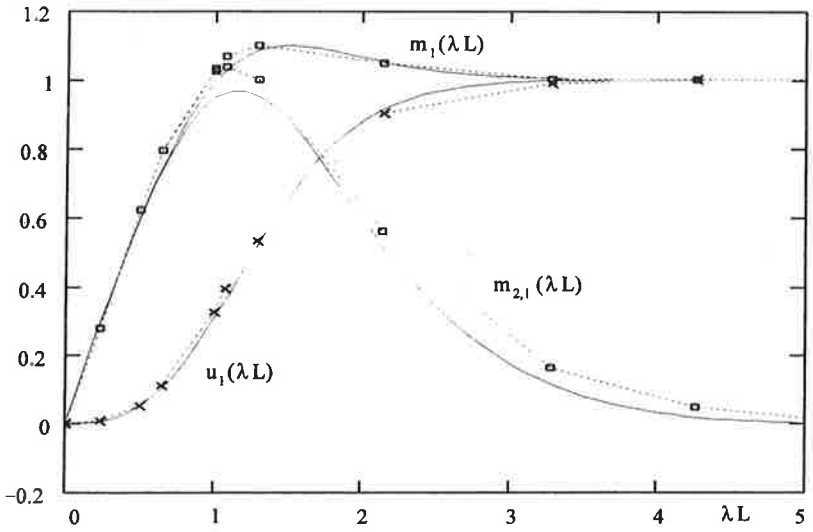
$$M_2 = M_{1,\infty}\, m_{2,1}(\lambda L)\ ,\quad m_{2,1}(\lambda L) = 2\frac{\alpha}{\beta}\frac{SH\,s}{SH\,CH + \frac{\alpha}{\beta}s\,c} \tag{31}$$

BEF mallin ja kuorielementti FE laskennan tulosten vastaavuus

Vertailun tuloksia on esitetty kuvissa 4 ja 5: BEF mallin tulokset on esitetty yhteneväisillä viivoilla ja tehtyjen FE mallinnusten tulokset symbolein ja interpoloitikatkoviivoin. Funktiot $u_1(\lambda L)$, $m_1(\lambda L)$ and $m_{2,1}(\lambda L)$ on kuvattu funktiona laatan normalisoidusta pituudesta λL. BEF malleissa karakteristisina pituuksina käytettiin $1/\lambda = 1.5B/\pi$ vapaasti tuetulle laatalle ja B/π jäykästi tuetulle. Vastaavasti toisen asteen parametrit olivat $n = 0.71$ and $1/\pi$. Kuva 4 esittää vertailun tulokset vapaasti tuetulle laatalle, kuva 5 jäykästi tuetulle.



Kuva 4 *Vapaasti tuetun laatan äärellisen pituuden kertoimet funktiona suhteesta λL: kohdassa x = 0 taipuma $u_1(\lambda L)$ ja taivutusmomentti $m_1(\lambda L)$; ja $m_{2,1}(\lambda L)$ kohdassa x = L, kun kuormituksena on poikittainen viivakuorma.*



Kuva 5 *Jäykästi tuetun laatan äärellisen pituuden kertoimet funktiona suhteesta λL: kohdassa x = 0 taipuma $u_1(\lambda L)$ ja taivutusmomentti, $m_1(\lambda L)$; ja $m_{2,1}(\lambda L)$ kohdassa x = L, kun kuormituksena on poikittainen viivakuorma.*

YHTEENVETO

Puoliäärettömän reunoiltaan vapaasti tai jäykästi laattakaistan käyttäytymistä keskilinjatasossaan voidaan rajoitetusti kuvata kimmoisalla alustalla olevan palkin BEF-teorialla. Malli toimii myös vapaareunaiselle laattakaistalle, kuten esim. I-palkin laipan ulkoreunalle. Kimmoisana alustana toimii laatan poikittaisjäykkyys. Jousivakio k määräytyy pitkittäissivujen reunaehdoista, kun kuormituksena on painekuorma p: $k = p/w$, jossa w on tarkastelukohdan taipuma. Laatan vääntöjäykkyys toimii aksiaalisen vetoesijännitysvoiman N tapaan taipuman differentiaaliyhtälön toisen asteen termin osana.

Kun tarkastellaan pituudeltaan äärellistä laattaa, puoliäärettömään BEF-malliin perustuva laskentamalli kuvaa myös hyvin viivaelementin käyttäytymistä. Laskennassa voidaan käyttää neljän vapausasteen muotofunktiota ja siihen perustuvaa palkkielementtiä.

Kehitettyä laskentamallia voidaan käyttää arvioitaessa hitsausliitoksissa asennustarkkuuksista ja hitsausliitoksen jäähtymisestä aiheutuvia rakenteellisia laatan taivutusjännitysten suuruuksia. Näiden huomioonottaminen on tärkeää määritettäessä väsyttävästi kuormitettujen hitsausliitosten väsymiskestoikää.

LÄHTEET

1. Partanen T. (1999). *On the application of beam on elastic foundation theory to the analysis of stiffened plate strips.* Doctor of Sciencis (Technology) Thesis. Acta Universitatis Lappeenrantaensis 86, 102 p.

2. Partanen T.(1997). Puoliäärettömät sauva- ja palkkielementit rakenteiden laskennassa. *Teräsrakenteiden tutkimus- ja kehityspäivät 16 - 17.1.1997.* 5 s.

3. Partanen T., (1999). Kimmoisella alustalla olevan palkin teorian soveltamisesta jäykistettyjen laattakaistojen analysointiin. *Teräsrakenteiden tutkimus- ja kehityspäivät 18-19.1.2000.* s. 223 - 230.

4. Yankelevsky, D. and Eisenberger, M. 1986. Analysis of beam column on elastic foundation. *Computers & Structures Vol. 23, No. 3*, pp. 351-356.

SUUNNAT SEPAROIVAAN SIIRTYMÄYRITTEESEEN PERUSTUVA OHUTSEINÄMÄISTEN KANNATTAJIEN LASKENTAMENETELMÄ

H. Perttola ja J. Paavola
Rakenteiden mekaniikan laboratorio
Teknillinen korkeakoulu
PL 2100
02015 TKK

TIIVISTELMÄ

Tarkastellaan ohutseinämäisten kannattajien analysointiin tarkoitettua laskentamenetelmää, joka perustuu semidiskreettiin formulaatioon. Menetelmässä siirtymäyritteen riippuvuus kannattajan aksiaali- ja poikittaissuunnassa juoksevista koordinaateista separoidaan. Kuvataan ensin dimensioreduktio kolmiulotteisesta yksiulotteiseksi, semidiskreettiä muotoa olevaksi tehtäväksi, jossa tuntemattomat funktiot riippuvat yksin aksiaalisesta koordinaatista. Redusoidun tehtävän diskretointi elementtimentelmän tapaan ohutseinämäisen kannattajan aksiaalisuunnassa synnyttää niin sanotut VI-elementit. Näiden muodostaminen esitetään systemaattisena menettelynä, jossa kaksiulottisessa alueessa (seinämänosassa) toteutettu siirtymien interpolaatio konstruoidaan aksiaalisuuntaan kulkevilla solmuviivoilla määritellyn parametrisen interpolaation ja poikittaissuunnassa luonteeltaan ei-parametrisen interpolaation avulla. Jälkimmäisen "parametrit" määräytyvät tuntemattomien funktioiden arvoista eivätkä ne siis ole todellisia parametreja. Lopuksi esitetään kaksi kuvatulla menetelmällä laskettua esimerkkiä.

JOHDANTO

Kolmiulotteisen probleeman korvaamista yksinkertaistetulla tehtävällä, jossa dimensioiden lukumäärä on supistunut kahteen tai yhteen, kutsutaan dimensioreduktioksi. Siirtymämenetelmän mukaisissa kimmo-opin tehtävissä se yleensä perustuu jossakin suunnassa tehtyyn kinemaattiseen olettamukseen, eli olettamukseen siitä, millainen tarkasteltavien siirtymäsuureiden jakaumien muoto on tietyllä koordinaattiviivalla, joka vastaa yhtä rakenteen ulottuvuuksista. Reduktioon johtava kinemaattinen olettamus tehdään useinmiten siinä suunnassa, jossa kappaleen geometrinen ulottuvuus on suhteellisesti ottaen pieni sen muihin mittoihin verrattuna. Vaaditaan siis, että vähintään yksi dimensio on paljon muita pienempi. Massiivisilla kappaleilla, joilla kaikki geometriset ulottuvuudet ovat yhtä merkittäviä, dimensioreduktioon perustuvan yleispätevän laskentamallin muodostamisen edellytykset ovat siten heikommat.

Palkin ulottuvuuksista kaksi (poikkileikkausmitat) ovat pieniä verrattuna sen pituuteen. Kolmiulotteinen malli voidaan korvata yksiulotteisella, kun alkuaan palkin akselia vastaan kohtisuoran poikkileikkaustason liike deformaatiossa kuvataan yksin aksiaalisesta koordinaatista riippuvien siirtymäsuureiden avulla perustuen esimerkiksi Eulerin-Bernoullin tai Timoshenkon palkkimalliin. Pintarakenteen (levy, laatta tai kuori) ulottuvuus pintaa vastaan kohtisuorassa suunnassa on pieni verrattuna rakenteen pinnan suuntaisiin dimensioihin, joita vastaavasti pintakoordinaatit määritellään. Pintarakenteen paksuuden suunnassa tehty kinemaattinen otaksuma (Kirchhoffin tai Mindlinin teoriat) johtaa kaksiulotteiseen malliin, jossa rakenteen deformaatio kuvataan yksin pintakoordinaattien funktiona esitettyjen siirtymäsuureiden avulla. Ohutseinämäisen palkkimaisen kannattajan seinämän paksuus on pieni verrattuna sen poikkileikkauksen karakteristiseen mittaan, joka on puolestaan pieni verrattuna kannattajan pituuteen. Ohutseinämäisen kannattajan kaikki kolme dimensiota ovat siis eri suuruusluokkaa. Tälle rakennetyypille voidaan periaatteessa muodostaa laskentamalli, jossa yhdistyvät reduktio kaksiulotteiseksi pintarakenteeksi ja edelleen reduktio yksiulotteiseksi palkkirakenteeksi.

Haettaessa siirtymämenetelmän mukaista ratkaisua kaksiulotteisessa alueessa määritellylle kimmo-opin tehtävälle voidaan ulottuvuuksista toinen periaatteessa redusoida pois valitsemalla etukäteen toiseen ulottuvuuteen (koordinaattiviivan α_1 suuntaan) liittyvät (kanta)funktiot $\Phi_i(\alpha_1)$ $(i = 1..n)$, joiden yhdistelmänä siirtymien jakaantuminen tässä suunnassa esitetään. Suunnista toisessa (α_2-koordinaattiviivan suunnassa) siirtymäjakauman muoto riippuu tuntemattomista funktioista $f_i(\alpha_2)$, jotka toimivat siirtymäyritteessä kantafunktioiden $\Phi_i(\alpha_1)$ kertoimina. Kaksiulotteinen tehtävä voidaan nyt korvata laskentateknisessä mielessä yksiulotteisella, jonka ratkaisuna ovat yksin α_2:sta riippuvat funktiot $f_i(\alpha_2)$. Tiettyä kantafunktioiden valintaa vastaavan redusoidun tehtävän ratkaisu on likiratkaisu alkuperäiselle tehtävälle. Menettelyä voidaan kuvata sanalla *semidiskreetti*, mikä viittaa interpolaation äärellistämiseen vain toisessa suunnassa. Tulkitaan, että kysymys on joko tiettyyn ongelmaan sopivan kinemaattisen olettamuksen asettamisesta (tapauskohtainen näkökulma) tai systemaattisesta tavasta interpoloida siirtymiä (yleinen näkökulma). Variaatiolaskennassa suuntien separointiin perustuva menettely tunnetaan *Kantorovichin menetelmänä (Kantorovich and Krylov: Approximate Methods of Higher Analysis, 1958)*, jossa funktiot $f_i(\alpha_2)$ ratkaistaan redusoidun probleeman Eulerin yhtälöstä. Tämä edellyttää, että likiratkaisu antaa semidiskreetin mallin kokonaispotentiaalienergialle minimiarvon. Viime vuosina semidiskreettiä formulaatiota on tutkinut kreikkalainen Xantis. Hän on esittänyt yleisen menettelyn, joka johtaa dimensioiltaan supistettuun tehtävään (*Chleboun and Xanthis: The Method of Arbitrary Lines in Optimal Shape Design: Problems with an Elliptic State Equation, 1996).* Menetelmän avulla etsitään elliptisille reunaarvotehtäville likimääräistä ratkaisua.

Ohutseinämäisen kannattajan siirtymätilaa voidaan kuvata aksiaalisen ja poikittaisen suunnan suhteen separoituvan siirtymäyritteen avulla. Valitsemalla siirtymäjakaumaa kannattajan poikkileikkauksessa kuvaavat kantafunktiot etukäteen, tehtävä voidaan redusoida yhteen, akselin suuntaiseen, ulottuvuuteen. Tavallisesti kantafunktiot mää-

ritellään joko koko poikkileikkaukselle tai sitten seinämäosittain (poikkileikkausprofiilin nurkasta nurkkaan). Tämän lähestymistavan voidaan katsoa edustavan tapauskohtaista näkökulmaa, jossa poikkileikkaukselle annetaan suoraan "järkevää kinemaattista oletusta" edustavat kantafunktiot. Yleisemmästä näkökulmasta kannattajan seinämät tulkitaan kaksiulotteisiksi rakenneosiksi, joille haetaan semidiskreettiä likiratkaisua. *Vlasov* esitteli kirjassaan *Thin-Walled Elastic Beams* [9] suuntien separointiin perustuvan ohutseinämäisten kannattajien analysointimenetelmän. Siinä kannattajan tasapainoehdot muodostetaan virtuaalisen työn periaatteen mukaan lähtemällä liikkeelle irtileikatusta, paksuudeltaan differentiaalisesta poikkileikkausviipaleesta *(elementary transverse strip)*, jota tarkastellaan siirtymäyritteen muodosta riippuen joko ristikkona tai kehänä. Tuloksena saadaan tavallisten differentiaaliyhtälöiden muodostama (redusoitu) yhtälöryhmä. Sen ratkaisemiseksi joudutaan yleensä turvautumaan numeerisiin menetelmiin. Elementtimenetelmää käytettäessä ratkaisuyhtälöiden äärellistettyä muotoa vastaa tietty elementtijako kannattajan akselin suunnassa. Teknillisessä korkeakoulussa on kehitetty [5] Vlasovin ideaan perustuva ohutseinämäisten kannattajien analysointiin tarkoitettu elementtimenetelmäsovellus. Sen elementtejä kutsutaan VIT-elementeiksi (**V**lasov **I**ntegrated **T**hin-Walled Element). Niiden avulla voidaan laskea semidiskreetin mallin tapaukselle likimääräinen ratkaisu.

TEHTÄVÄN MÄÄRITTELY, DIMENSIOREDUKTIO

Tarkastellaan levymäisistä osista koottua ohutseinämäistä kannattajaa. Olkoon seinämänosan i ($i = 1..N$) keskipintaa (tasoa) vastaava kaksiulotteinen alue ${}^i\Omega_2$ ja sen paksuus $2a_i$, joka oletetaan seinämänosassa vakioksi. Deformoitumattomassa alkutilanteessa ohutseinämäinen kannattaja vie tilan

$$\Omega_3 = \bigcup_i^N {}^i\Omega_3, \tag{1}$$

jossa ${}^i\Omega_3 = {}^i\Omega_2 \times [-a_i, a_i]$ vastaa seinämänosan i ottamaa tilaa. Olkoon $(x, y, z) \in {}^i\Omega_3$ materiaalipartikkelin koordinaatit suorakulmaisessa koordinaatistossa ja $\mathbf{r}^{(0)}$ vastaava paikkavektori. Deformaatiossa partikkeli siirtyy niin, että sen uusi asema on

$$\mathbf{r} = \mathbf{r}^{(0)} + \mathbf{u}, \tag{2}$$

missä $\mathbf{u} = u(x, y, z)\mathbf{i} + v(x, y, z)\mathbf{j} + w(x, y, z)\mathbf{k}$ on siirtymävektori (annettuna tässä suorakulmaisessa kannassa $\{\mathbf{i},\ \mathbf{j},\ \mathbf{k}\}$). Kokonaispotentiaalienergian minimin periaatteen mukaan kinemaattisesti hyväksyttävien siirtymätilojen joukosta $\mathbf{v} \in \mathrm{V}$ todellinen siirtymätila $\mathbf{u}(x, y, z)$ antaa kimmoisen kappaleen potentiaalienergialle I minimin. Kolmiulotteista ohutseinämäisen kannattajan mallia vastaava kimmo-opin tehtävä voidaan siis antaa yleisessä muodossa:

$$\mathcal{M}: \quad \text{Etsi } \mathbf{u}(x, y, z) \in \mathrm{V} \text{ se. } I(\mathbf{u}) \leq I(\mathbf{v}) \ \ \forall \mathbf{v}(x, y, z) \in \mathrm{V}. \tag{3}$$

Tarkastellaan sitten seinämänosaa ${}^i\Omega_3$ ja määritellään z-akseli kohtisuoraksi keskitasoa vastaan sekä x- ja y-akselit yhtymään keskitasoon ($z = 0$). Seinämänosan siirtymätila

voidaan tällöin esittää Taylorin sarjana

$$\mathbf{u}(x,y,z) = \hat{\mathbf{u}}_0 z^0 + \hat{\mathbf{u}}_1 z^1 + \hat{\mathbf{u}}_2 z^2 + \cdots \tag{4}$$

jossa potenssien z^n $(n = 0, 1, ..)$ kertoimien komponentit määritetään z:n suhteen otettujen siirtymävektorin osittaisderivaattojen arvoista seinämänosan keskipinalla $z = 0$ (ks. esim. [10]). Muodostamalla siirtymäyrite $\mathbf{u}$:lle siten, että interpoloidaan kertoimien $\hat{\mathbf{u}}_n(x,y,0)$ komponentteja keskipinnalla, syntyy kaksiulotteinen tehtävä, joka edustaa alkuperäisen kolmiulotteisen tehtävän reduktiota. Kun sarja (4) katkaistaan jättämällä siihen vain kaksi ensimmäistä termiä, saadaan siirtymävektorille approksimaatio

$$\mathbf{u} \approx \hat{\mathbf{u}} = \hat{\mathbf{u}}_0(x,y,0) + \hat{\mathbf{u}}_1(x,y,0)z, \tag{5}$$

joka määräytyy kuudesta keskipinnan suureesta eli kolmesta keskipinnan ${}^{i}\Omega_2$ siirtymävektorin $\hat{\mathbf{u}}_0 = u_0(x,y)\mathbf{i} + v_0(x,y)\mathbf{j} + w_0(x,y)\mathbf{k}$ komponentista ja kolmesta vektorin $\hat{\mathbf{u}}_1 = u_z(x,y,0)\mathbf{i} + v_z(x,y,0)\mathbf{j} + w_z(x,y,0)\mathbf{k}$ komponentista. Probleema (3) voidaan siis korvata tehtyä likimääräistystä vastaavalla, kahteen ulottuvuuteen redusoidulla tehtävällä

$$\hat{\mathcal{M}}: \quad \text{Etsi } \hat{\mathbf{u}} \in \hat{\mathrm{V}} \text{ se. } I(\hat{\mathbf{u}}) \leq I(\hat{\mathbf{v}}) \ \ \forall \hat{\mathbf{v}} \in \hat{\mathrm{V}}, \tag{6}$$

jossa kokonaispotentiaalienergian minimoivaa funktiota $\hat{\mathbf{u}}$ etsitään kinemaattisesti hyväksyttävistä siirtymätilojen joukosta $\hat{\mathrm{V}}$. Tämän alkiot konstruoidaan esimerkiksi approksimaatioon (5) perustuvina. Luonteeltaan tämä approksimaatio on kinemaattinen otaksuma, jonka mukaan alunperin keskipintaa vastaan kohtisuorat viiva-alkiot pysyvät seinämänosan deformaatiossa suorina. Jos näiden viiva-alkioiden suhteellinen pituudenmuutos on merkityksettömän pieni, niin saadaan lisäehto, joka Greenin-Lagrangen muodonmuutostensorin avulla kirjoitettuna kuuluu: $e_{zz} = w_z + 0.5(u_z^2 + v_z^2 + w_z^2) = 0$. Linearisoituna (otaksutaan pienet siirtymägradientin komponentit) se saa muodon: $w_z = 0$. Jäljellä on enää viisi keskipinnalla määriteltyä siirtymäkomponenttia, joista kaksi (u_0 ja v_0) liittyy seinämänosan toimintaan membraanina ja kolme (w_0, u_z ja v_z) sen deformaatioon taivutuksessa niin sanotun Mindlinin laattamallin mukaan. Jos vielä vaaditaan, että leikkausmuodonmuutos seinämänosan taivutuksessa häviää ($u_z = -w_{o,x}$ ja $v_z = -w_{o,y}$), päädytään Kirchhoffin laattamalliin, jossa taivutus kuvataan yksin keskipinnan taipuman w_o avulla. Olkoon minimointitehtävä $\hat{\mathcal{M}}_m$, kun se liittyy seinämänosan toimintaan membraanina sekä olkoon Kirchhoffin laattamalliin perustuva taivutustehtävä $\hat{\mathcal{M}}_{b(K)}$ ja Mindlinin mallia vastaavaa taivutustehtävä $\hat{\mathcal{M}}_{b(M)}$.

Ohutseinämäisen kannattajan analyysi on seinämänosittain redusoitavissa yllä kuvatulla tavalla kolmiulotteisesta tehtävästä $\mathcal{M}$ kaksiulotteiseksi tehtäväksi $\hat{\mathcal{M}}$. Tällöin se yleensä vastaa yhdistettyjä tehtäviä ($\hat{\mathcal{M}}_m$ ja $\hat{\mathcal{M}}_{b(K)}$) tai ($\hat{\mathcal{M}}_m$ ja $\hat{\mathcal{M}}_{b(B)}$). Ennen kuin ehdotetaan menettelytapaa $\hat{\mathcal{M}}$:n redusoimiseksi edelleen yhteen dimensioon, niin tarkastellaan yleistä kaksiulotteista tehtävää, jossa etsitään alueessa $(x,y) \in \Omega_2$ määriteltyä funktiota

$$\hat{f} = \hat{f}(x,y), \qquad (i = 1..n) \tag{7}$$

siten, että se minimoi funktionaalin

$$I(f) = \iint_{\Omega_2} F(x, y, \hat{f}, \partial_x \hat{f}, \partial_y \hat{f}, \partial_{xy} \hat{f}, ...) dx dy. \tag{8}$$

ja toteuttaa vaaditut ehdot tarkastelualueen reunalla. Likimääräinen ratkaisu tälle voidaan saavuttaa esimerkiksi *Ritzin menetelmällä*, jolloin yritefunktio on muotoa

$$\hat{f}(x, y) \approx \hat{f}_R(x, y) = a_1 \phi_1(x, y) + a_2 \phi_2(x, y) + \ ... \ + a_N \phi_N(x, y), \tag{9}$$

jossa etukäteen valitut (kanta)funktiot $\phi_j(x, y)$ määritellään alueessa Ω_2 ja kertoimet a_i ovat määräämättömiä vakioita. Esitellään sitten yleistetty yritefunktio

$$\hat{f}(x, y) \approx \hat{f}^\star(x, y) = a_1(x) \phi_1(x, y) + a_2(x) \phi_2(x, y) + \ ... \ + a_N(x) \phi_N(x, y), \tag{10}$$

jossa kertoimet a_j eivät olekaan enää määräämättömiä vakioita, vaan tuntemattomia koordinaatin x funktioita $a_j = a_j(x)$. Kun yrite (9) sijoitetaan funktionaaliin (8), niin syntyy funktio $I = I(a_1, a_2, ..., a_N)$, jolle minimin antava kertoimien $a_1, a_2, ..., a_N$ yhdistelmä määräytyy algebrallisesta yhtälöryhmästä $(I_{,a_i} = 0, i = 1..N)$. Jos sijoitetaankin yleistettyä muotoa olevat yritefunktiot (10) funktionaaliin (8) ja integroidaan toisen koordinaatin y suhteen, niin on periaatteessa saavutettavissa uusi funktionaali muotoa

$$I(a_1, ..., a_N) = \int_{I_x} \mathcal{F}(x, a_1, ..., a_n, \partial_x a_1, ..., \partial_x a_n, ...) dx, \tag{11}$$

jossa integroidaan enää yhdessä suunnassa. Kantorovich ehdotti tuntemattomien, yksin koordinaatista x riippuvien kerroinfunktioiden $a_j(x)$ yhdistelmän määrittämistä niin, että ne antavat minimin funktionaalille (11). Syntyy siis uusi muodollisesti yksiulotteinen tehtävä:

$$\hat{\mathcal{M}}^\star: \quad \text{Etsi } \hat{f}^\star \in \hat{\mathrm{V}}^\star \text{ se. } I(\hat{f}^\star) \leq I(\hat{g}^\star) \ \ \forall \hat{g}^\star \in \hat{\mathrm{V}}^\star, \tag{12}$$

jossa funktionaalille (11) minimiarvon antavaa funktiota $\hat{f}^\star$ haetaan approksimaatioon (10) perustuvien kinemaattisesti hyväksyttävien siirtymätilojen joukosta $\hat{\mathrm{V}}^\star$.

Likimääräinen ratkaisumenetelmä, joka perustuu modifioituun tehtävään $\hat{\mathcal{M}}^\star$ tunnetaan *Kantorovichin menetelmänä* (erityisesti [4], myös [2], [6] ja [7].

Semidiskreetti tehtävä $\hat{\mathcal{M}}^\star_K$

Erikoistapauksena yritefunktiosta (10) voidaan esitellä riippuvuuden koordinaateista x ja y separoiva muoto

$$\hat{f}(x, y) \approx \hat{f}^\star(x, y) = h(x) \Psi(y) \tag{13}$$

jossa (kanta)funktio $\Psi(y)$ riippuu vain y:stä ja yksin x:stä riippuva $h(x)$ on tuntematon *kerroinfunktio*. Funktio $\Psi(y)$ voidaan valita rationaalisin perustein vain, jos on riittävä tieto siitä, miten interpoloitava funktio $\hat{f}(x, y)$ muuttuu y-koordinaattiviivan

suunnassa. Jos tämän kuvaaminen vain yhden kantafunktion avulla on riittämätöntä, niin yritteeseen (13) lisätään termejä. Tällöin se saa muodon

$$\hat{f}^{\star}(x,y) = h_1(x)\Psi_1(y) + h_2(x)\Psi_2(y) + \ ... \ + h_N(x)\Psi_N(y) \tag{14}$$

jossa jokainen termi on x:n ja y:n suhteen separoituva. Lisäksi vaaditaan, että funktiot $\Psi_j(y)$ $(j = 1..N)$ muodostavat lineaarisesti riippumattoman kannan

$$K = \{\Psi_1, \Psi_2, ..., \Psi_N\}. \tag{15}$$

Kiinnittämällä x $(x = x_k = vakio)$ ja merkitsemällä $h_j(x_k) = c_j$, saadaan (14):sta luonteeltaan *ei-parametrinen interpolaatio* y-koordinaattiviivan suunnassa eli

$$\hat{f}(\cdot, y) = c_1\Psi_1(y) + c_2\Psi_2(y) + \ ... \ + c_3\Psi_N(y), \tag{16}$$

jossa $\hat{f}_i(x,y)$:n jakaumaa interpoloidaan kerroinfunktioiden $h_j(x)$ arvoista. Nämä eivät suoraan toimi ratkaistavan tehtävän tuntemattomina. Funktio $\hat{f}_i(x,y)$ siis esitetään y-koordinaattiviivan suunnassa diskretoituna, mutta x-koordinaattiviivan suunnassa jatkuvana, tuntemattomien funktioiden $h_j(x)$ $(j = 1..N)$ yhdistelmänä. Tehtävän $\hat{\mathcal{M}}^{\star}$ modifikaatiota, joka perustuu termeittän separoituvaan yritteeseen (14), nimitetään *semidiskreetiksi tehtäväksi* $\hat{\mathcal{M}}^{\star}_K$. $\hat{\mathcal{M}}^{\star}_K$:n oikea alaindeksi K viittaa kantaan (15), joka valitaan etukäteen.

Palataan takaisin ohutseinämäisen kannattajan dimensioreduktioon. Ensimmäisessä vaiheessa jokainen seinämänosa supistuu kaksiulotteiseksi. Tehty reduktio

$$\mathcal{M} \longrightarrow \hat{\mathcal{M}}$$

perustuu kinemaattiseen olettamukseen siirtymäjakauman muodosta seinämän keskipinnan kohtisuoran suunnassa. Toinen vaihe kaksiulotteisesta tehtävästä yksiulotteiseen perustuu yllä kuvattuun reduktioon

$$\hat{\mathcal{M}} \longrightarrow \hat{\mathcal{M}}^{\star}_K.$$

Tällöin interpoloitavan suureen riippuvuus separoidaan seinämänosan pintakoordinaateista, jotka nyt vastaavat kannattajan aksiaaliseen ja poikkileikkauksen tasossa määriteltyihin suuntiin kulkevia koordinaattiviivoja. Kannan K funktiot puolestaan määrittelevät ohutseinämäisen kannattajan poikkileikkaukselle mahdolliset deformaatiomuodot, jotka voidaan periaatteessa esittää niin tarkasti kuin halutaan kantaa K parantamalla.

VI-ELEMENTIT

Otetaan lähtökohdaksi tehtävä $\hat{\mathcal{M}}^{\star}_K$, johon liittyy semidiskreettiä muotoa (14) oleva siirtymäyrite. Tehtävä $\hat{\mathcal{M}}^{\star}_K$ ei useinkaan ole ratkaistavissa analyyttisesti, jolloin myös

tuntematonta, yksin koordinaatista x riippuvaa funktiota $h_i(x)$ on approksimoitava. Muodostaan siis interpolaatio

$$h_i(x) \approx \sum_{j=1}^{J} \Phi_j(x)\, {}^{i}p_j, \tag{17}$$

jossa jokaiseen funktioon $h_i(x)$ $(i = 1..I)$ liittyvät parametrit ${}^{i}p_j$ $(j = 1..J)$ toimivat ratkaistavan tehtävän tuntemattomina. Tämän viimeiseksi tehdyn likimääräistyksen synnyttämä tehtävä voidaan antaa muodossa

$$\hat{\mathcal{M}}^{\star}_{KL}: \quad \text{Etsi } \hat{f}^{\star}_{KL} \in \hat{\mathrm{V}}^{\star}_{KL} \text{ se. } I(\hat{f}^{\star}_{KL}) \leq I(\hat{g}^{\star}_{KL}) \;\; \forall \hat{g}^{\star}_{KL} \in \hat{\mathrm{V}}^{\star}_{KL}. \tag{18}$$

Joukkoon $\hat{\mathrm{V}}^{\star}_{KL}$ kuuluvat nyt ne kinemaattisesti hyväksyttävät funktiot, jotka on mahdollista esittää approksimaationa

$$f(x,y) \approx \sum_{i=1}^{I} \sum_{j=1}^{J} \Psi_i(y)\Phi_j(x)\, {}^{i}p_j. \tag{19}$$

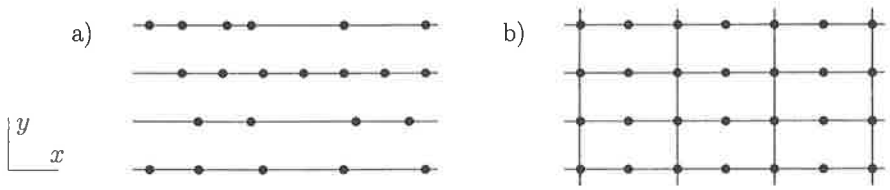
Tällöin funktioiden Φ_j oletetaan muodostavan lineaarisesti riippumattoman kannan

$$L = \{\Phi_1, \Phi_2, ..., \Phi_N\} \tag{20}$$

ja antavan ratkaisulle muodon x-akselin suunnassa. Approksimaatioon (17) perustuvan muunnoksen

$$\hat{\mathcal{M}}^{\star}_{K} \longrightarrow \hat{\mathcal{M}}^{\star}_{KL}$$

tarkoituksena on korvata semidiskreetti tehtävä tehtävällä, joka voidaan ratkaista algebrallisesta yhtälösysteemistä. Kantafunktioiden $\Phi(x) \in L$ valinta tekee siis mahdolliseksi saavuttaa likimääräinen ratkaisu semidiskreetille tehtävälle $\hat{\mathcal{M}}^{\star}_{K}$.



Kuva 1: Solmujen sijoittaminen solmuviivoille: a) Ilman ryhmittelyä ja b) ryhmiteltynä elementtijakoa varten (kolme kvadraattista elementtiä).

Oletetaan nyt, että kerroinfunktiot $h_i(x)$ määritellään tietyllä x-koordinaattiviivalla $y = y_l$ eli *solmuviivalla*. Tällä sijaitsevat *solmupisteet*, joihin kerroinfunktioiden interpolaation (17) parametrit liittyvät, voidaan periaatteessa valita täysin vapaasti. Tätä havainnollistaa kuva 1 a. Kuva 1 b esittää puolestaan tilannetta, jossa kaikkien funktioiden $h_i(x)$ interpolaatio solmuviivoilla on samanlainen ja analoginen eräälle elementtimenetelmän muotofunktioapproksimaatiolle, jolloin solmupisteet ryhmitellään

tietyille koordinaattiviivoille $x = x_k = vakio$. Näin syntyy x-akselin suunnassa tietty elementtijako, jota b-kohdan esimerkkitapauksessa vastaa paloittain toisen asteen Lagrangen interpolaatio. Kun ohutseinämäinen kannattaja, jolle muodostetaan suunnat separoiva siirtymäyrite "Vlasovin tapaan", diskretoidaan yllä kuvatun menettelyn mukaisesti akselinsa suunnassa, niin muodostuu niin sanottuja *Vlasov-integroituja-elementtejä (=VI-elementtejä).*

Toisaalta kantafunktiot $\Psi_i(y) \in K$ voidaan valita siten, että nekin vastaavat tiettyä yksiulotteista elementtimenetelmän muotofunktioapproksimaatiota. Tällöin syntyy elementtijako myös y-akselin suunnassa. VI-elementit, jotka diskretoivat ohutseinämäisen kannattajan sen akselin suunnassa, jakautuvat tällöin itsekin useaan elementtiin, joita tässä kutsutaan *seinämänosiksi.* Kun siirtymien interpolaatiot sekä x- että y-akselin suuntaan toteutetaan muotofunktioapproksimaatioille analogisesti, niin muodostuu paloittain (seinämänosittain) määriteltyjä interpolaatioita, jotka voivat olla x- ja y-suuntien suhteen joko symmetrisiä tai epäsymmetrisiä. Seuraavassa esitetään, kuinka tällä tavoin voidaan muodostaa esimerkiksi bikuubinen elementti, joka laatan tapauksessa tunnetaan BSF-elementtinä.

Tehtävää $\hat{\mathcal{M}}_{KL}$ määriteltäessä ei kiinnitetty mitään huomiota reunaehtoihin. Tämä siksi, että kuvattuun ratkaisumenettelyyn liittyen reunaehdot annetaan samaan tapaan kuin elementtimenetelmässä yleensä. Huomautetaan vielä, että jos tehtävä $\hat{\mathcal{M}}^*_{KL}$ muotoillaan siten, että funktiot $\Psi \in K$ vastaavat Fourier-sarjan kantafunktioita ja funktiot $\Phi \in L$ valitaan muotofunktioapproksimaatioille analogisina kantafunktioina, niin syntyy laskentamalli, joka tunnetaan *äärellisten kaistojen menetelmänä* (*Cheung:* [1]).

Bikuubinen elementti

Tarkastellaan kaksiulotteista aluetta, jossa interpoloidaan suorakulmaisista koordinaateista x ja y riippuvaa funktiota $w(x, y)$. Se voidaan tulkita esimerkiksi laatan taipumaksi. Olkoon $w(x, y)$:n interpolaatio semidiskreettiä muotoa

$$\hat{f}(x, y) = \sum_{i=1}^{n} g_i(x) H_i(y), \tag{21}$$

jossa etukäteen valitut funktiot $H_i(y)$ muodostavat lineaarisesti riippumattoman kannan $K = \{H_i(y)\}$. Kun taipuman interpolaatio mielivaltaisella y-koordinaattiviivalla $(x = x_k)$ toteutetaan C_1-jatkuvana, niin osavälillä $y \in [y_i, y_{i+1}]$ voidaan kirjoittaa

$$\hat{w}(\cdot, y) = {}^iH_1 \; w_i + {}^iH_2 \; \phi_i + {}^iH_3 \; w_{i+1} + {}^iH_4 \; \phi_{i+1}, \tag{22}$$

jossa funktiot ${}^iH_j = {}^iH_j(y)$ $(j = 1..4)$ ovat kyseiselle osavälille i määritellyt Hermiten kuubiset interpolaatiopolynomit sekä

$$w_l = w(x_k, y_l) \text{ ja } \phi_l = w_{,y}(x_k, y_l).$$

Määrittelemällä koordinaattiviivalla $y = y_l$ yksin x:stä riippuvat funktiot

$$ {}^{l}\mathcal{W}(x) = w(x, y_l) \quad \text{ja} \quad {}^{l}\mathcal{R}(x) = w_{,y}(x, y_l) $$

ja sijoittamalla ne yritteessä (21) $g_i(x)$-funktioiden tilalle, saadaan osa-alueella $(x, y) \in {}^{l}\Omega_2 = [x_0, x_1] \times [y_l, y_{l+1}]$ määritelty semidiskreettiä muotoa oleva interpolaatio

$$ \hat{w}(x, y) = {}^{i}H_1(y)\ {}^{i}\mathcal{W}(x) + {}^{i}H_2(y)\ {}^{i}\mathcal{R}(x) + {}^{i}H_3(y)\ {}^{i+1}\mathcal{W}(x) + {}^{i}H_4(y)\ {}^{i+1}\mathcal{R}(x). \quad (23) $$

Kun koordinaattiviivalla $y = y_i$ määriteltyjä funktioita ${}^{i}\mathcal{W}(x)$ interpoloidaan C_1-jatkuvana, niin osavälillä $x \in [x_j, x_{j+1}]$ saadaan approksimaatio

$$ {}^{i}\mathcal{W}(x) \approx {}^{j}\tilde{H}_1\ {}^{i}\mathcal{W}(x_j) + {}^{j}\tilde{H}_2\ {}^{i}\mathcal{W}_{,x}(x_j) + {}^{j}\tilde{H}_3\ {}^{i}\mathcal{W}(x_{j+1}) + {}^{j}\tilde{H}_4\ {}^{i}\mathcal{W}_{,x}(x_{j+1}), \quad (24) $$

jossa ${}^{j}\tilde{H}_l = {}^{j}\tilde{H}_1(x)$ ovat kyseiselle osavälille j määritellyt Hermiten interpolaatiopolynomit. Samaan tapaan funktiolle ${}^{i}\mathcal{R}(x)$ voidaan kirjoittaa

$$ {}^{i}\tilde{\mathcal{R}}(x) \approx {}^{j}\tilde{H}_1\ {}^{i}\mathcal{R}(x_j) + {}^{j}\tilde{H}_2\ {}^{i}\mathcal{R}_{,x}(x_j) + {}^{j}\tilde{H}_3\ {}^{i}\mathcal{R}(x_{j+1}) + {}^{j}\tilde{H}_4\ {}^{i}\mathcal{R}_{,x}(x_{j+1}). \quad (25) $$

Osa-alueella $(x, y) \in {}^{ij}\Omega_2 = [x_j, x_{j+1}] \times [y_i, y_{i+1}]$ voidaan nyt antaa diskretoitua semidiskreettiä mallia vastaava interpolaatio

$$ \hat{\tilde{w}}(x, y) = \sum_{k=1}^{4} \sum_{l=1}^{4} {}^{i}H_k(y)\, {}^{j}\tilde{H}_l(x)\ {}^{l}p_k, \quad (26) $$

jossa esimerkiksi solmupsiteessä (x_j, y_i) määritellyt neljä vapausastetta ovat

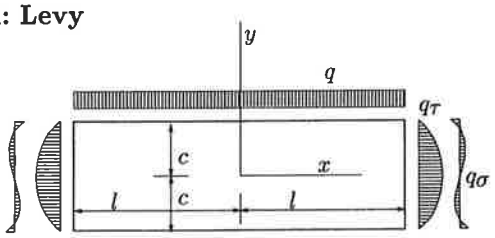
$$ \begin{aligned} {}^{1}p_1 &= {}^{i}\mathcal{W}(x_j) = w(x_j, y_i), \\ {}^{1}p_2 &= {}^{i}\mathcal{W}_{,x}(x_j) = w_{,x}(x_j, y_i), \\ {}^{2}p_1 &= {}^{i}\mathcal{R}(x_j) = w_{,y}(x_j, y_i), \\ {}^{2}p_2 &= {}^{i}\mathcal{R}_{,x}(x_j) = w_{,xy}(x_j, y_i). \end{aligned} \quad (27) $$

Kaikkiaan tarvitaan $4 \times 4 = 16$ vapausastetta. Näin generoitu, osa-alueittain ${}^{ij}\Omega_2$ määritelty siirtymäinterpolaatio vastaa taipuman interpolaatiota BSF- eli **B**ogner-**F**ox-**S**chmit-laattaelementein.

ESIMERKIT

Seuraavassa esitellään kaksi yksinkertaista esimerkkiä, joissa ensimmäisessä tarkastellaan päädyistään tuettua kapeahkoa levyä ja toisessa I-poikkileikkauksellista kannattajaa, joka on kuormitettu pienemmän jäykkyyden suunnassa.

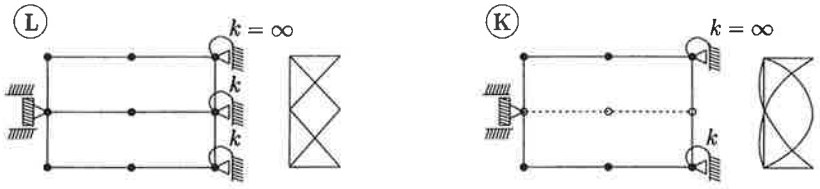
Esimerkki: Levy



$l = 12$ yks
$c = 8$ yks
$q = 1$ N/yks
$E = 10^{10}$ N/yks^2
$\nu = 0.3$
(arvot Hakala [3])

Kuva 2: Levy ja sen kuormitus

Tarkastellaan levytehtävää kuvan 2 tapauksessa, jossa suorakaiteen muotoisen kapeahkon levyn yläreunassa ($y = c$) on tasainen kuorma q. Levyn päädyissä ($x = \pm l$) vaikuttavat pinnan suuntainen kuorma q_τ sekä pintaa vastaan kohtisuora kuorma q_σ, joka ei aiheuta levyn päätyyn nollasta poikkeavaa normaalivoima- eikä momenttiresultanttia. Analyyttinen ratkaisu on löydettävissä (Timoshenko: [7]) yhdistelemällä polynomimuotoisia osaratkaisuja siten, että jännitysjakauma levyn reunoilla tasapainottaa ulkoisen kuormituksen. Itseasiassa päätykuormat ($x = \pm l$) ja q_τ määritellään niin, että niiden muoto vastaa jännitysjakaumia, jotka otaksuttu jännitysfunktio synnyttää. Siirtymät $u(x, y)$ ja $v(x, y)$ x- ja y-akselin suuntaan integroidaan venymistä, jotka riippuvat jännityksistä Hooken lain mukaan. Lisäksi otaksutaan, että levyn keskikohdan vaakasiirtymä $u(0, \cdot)$ häviää (symmetria) ja että vaakakeskilinjan ($y = 0$) taipuma $v(x, 0)$ saa arvon nolla päädyissä ($x = \pm l$). Näin menetellen määritetty keskipisteen taipuma valitaan referenssisiirtymäksi $\delta_{ref} = v(0, 0)$, johon VI-elementtien avulla laskettuja taipuman arvoja verrataan. Neljän numeron tarkkudella annettuna se on kuvan 2 levylle $\delta_{ref} = -0.1269 \times 10^{-7}$ pituusyksikköä.



Kuva 3: Laskentamallit, joissa molemmissa on yksi VI-elementti. Pystysiirtymän v kuvaamiseen ei-parametrisessa interpolaatiossa käytetyt kantafunktiot on piirretty mallin oikealle puolelle.

Käytetään hyväksi symmetriaa ja muodostetaan laskentamalli vain puolikasrakenteelle. Kuvan 3 molemmissa tapauksissa se on kuvattu yhdellä VI-elementillä. Tarkastellaan ensin näistä vasemmanpuolista (**L**), jossa on kolme solmuviivaa $\underline{i = 1}$: $y = -c$, $\underline{i = 2}$: $y = 0$ ja $\underline{i = 3}$: $y = c$, jotka jakavat levyn kahteen seinämänosaan. Kerroinfunktiot ${}^i\mathcal{U}(x)$, ${}^i\mathcal{V}(x)$ ja ${}^i\mathcal{R}(x)$ edustavat nyt vaakasiirtymää, pystysiirtymää ja rotaatiota levyn tason normaalin ympäri eli

$$ {}^i\mathcal{U}(x) \leftrightarrow u(x, y_i), \qquad {}^i\mathcal{V}(x) \leftrightarrow v(x, y_i) \quad \text{ja} \quad {}^i\mathcal{R}(x) \leftrightarrow \phi(x, y_i). $$

Annetaan niille kaikille kvadraattinen interpolaatio solmuviivoilla (solmupisteet on piirretty solmulinjoille mustina ympyröinä). Siirtymäkomponenttia u interpoloidaan mielivaltaisella y-akselin suuntaisella koordinaattiviivalla ($x = x_k = vakio$) C_1-jatkuvien Hermiten polynomien avulla kerroinfunktioiden arvoista ${}^i\mathcal{U}(x_k)$ ja ${}^i\mathcal{R}(x_k)$ ($i = 1, 2, 3$). Komponenttia v interpoloidaan puolestaan paloittain lineaarisesti kerroinfunktioiden ${}^i\mathcal{V}(x_k)$ ($i = 1, 2, 3$) arvoista. Sekä $u(x, y)$:lle että $v(x, y)$:lle muodostuva interpolaatio on epäsymmetrinen eli erilainen x- ja y-akselin suunnissa.

Oikeanpuolinen VI-elementti (**K**) eroaa edellisestä ensinnäkin siinä, että siinä on kaksi varsinaista solmuviivaa ($y = \pm c$), joilla interpoloidaan kaikkia kerroinfunktioita ${}^i\mathcal{U}(x)$, ${}^i\mathcal{R}(x)$ ja ${}^i\mathcal{V}(x)$. Tämän lisäksi kerroinfunktiota ${}^i\mathcal{V}(x)$ interpoloidaan edelleen "ylimääräisellä" solmuviivalla $y = 0$, sillä sitä tarvitaan komponentin v ei-parametrisen interpolaation muodostamiseen (levyn korkeuden pituisella välillä). VI-elementti (K) sisältää tällöin vain yhden seinämänosan, jossa v:n interpolaatio on samanlainen kuin yhdeksän pisteen Lagrangen elementissä ja u:n interpolaatio on analoginen mallin **L** seinämänosissa oleville u:n interpolaatioille. Reunaehto taipumalle päädyn keskellä voidaan antaa, sillä ylimääräisellä solmuviivalla on solmupiste juuri oikeassa kohdassa.

VI-elementein suorituissa laskelmissa käytetyt reunaehdot nähdään kuvasta 3. Symmetriasyistä pystylinjan $x = 0$ solmupisteiden vaakasiirtymät ja kiertymät estetään. Levyyn kohdistuvat reunakuormat q, q_τ ja q_σ korvataan ekvivalenteilla solmupistekuormilla. Taulukoista 1 ja 2 nähdään, kuinka keskipisteen taipuman δ suhde referenssitaipumaan δ_{ref} muuttuu, kun VI-elementtien jakoa tihennetään yhdestä neljään, mutta samalla ei-parametrinen interpolaatio pidetään ennallaan (ei lisätä seinämänosia). Havaitaan, että yhdellä elementillä saavutetaan kuvan 3 tyyppiä **L** olevalla VI-elementillä taipuman arvo, joka eroaa referenssitaipumasta reilut kolme prosenttia ja että neljällä elementillä taipuma eroaa referenssitaipumasta vielä vajaan prosentin. Jos sama tarkastelu tehdään tyyppiä **K** olevalle elementille, niin lähestytään varsin nopeasti referenssiarvoa, joka neljällä elementillä laskettuna on neljän numeron tarkkuudella sama kuin referenssitaipuma. Tehtiin myös tarkastelu, jossa lisättiin seinämänosia (=tihennettiin solmuviivojen väliä), mutta varsinaisten VI-elementtien määrä pidetään vakiona. Tällöin havaittiin, että seinämänosien lisäämisellä ei juurikaan ollut vaikututusta, jos siirtymäkomponentin v ei-parametrinen interpolaatio oli kvadraattinen (**K**), sillä tässä nimenomaisessa levyratkaisussa kvadradraattinen muoto dominoi v:n jakaumaa. Jos taas v:n ei-parametrinen interpolaatio oli lineaarinen (**L**), niin saadaan aikaan vain hidas parannus ja solmuviivoja olisi lisättävä runsaasti, jotta yhtä hyvä tulos kuin kvadraattisen interpolaation tapauksessa saavutettaisiin. Päätellään, että kvadraattinen interpolaatio (**K**) y-akselin suunnassa komponentille v tuo tarkastellun levytehtävän tapauksessa merkittävän parannuksen saatuihin tuloksiin, sillä tällöin päästään harvemmalla verkolla lähemmäs oikeaa ratkaisua kuin lineaarisen interpolaation tapauksessa (**L**).

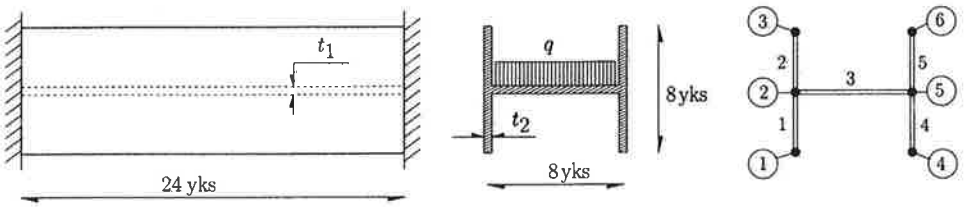
n	1	2	3	4
δ/δ_{ref}	0.9674	0.9896	0.9909	0.9912

Taulukko 1: Suhde δ/δ_{ref}, kun laskelmissa käytetään VI-elementtiä tyyppiä **L**.

n	1	2	3	4
δ/δ_{ref}	0.9751	0.9981	0.9997	1.000

Taulukko 2: Suhde δ/δ_{ref}, kun laskelmissa käytetään VI-elementtiä tyyppiä **K**.

Esimerkki: I-palkki



Kuva 4: I-kannattaja ja sen kuormitus, solmuviivat ja seinämänosat

Tarkastellaan kuvan 4 mukaista päistään jäykästi kiinnitettyä, kimmoisesta materiaalista ($E = 210000\text{N/yks}^2$, $\nu = 0.3$) valmistettua I-poikkileikkauksellista kannattajaa. Kannattajan vaakaseinämään vaikuttaa alaspäin suuntautunut, tasainen kuorma $q = 0.25\text{N/yks}^2$. Olkoon vaakaseinämän paksuus $t_1 = 0.8$ yks. ja pystyseinämämien paksuus t_2 mitta, jota vaihdellaan. Muut mitat ilmenevät kuvasta 4. Vaakaseinämälle valittu paksuuden arvo $t_1 = 0.8$ yks. täyttää Vlasovin ohutseinämäisen kannattajan poikkileikkauksen hoikkuudelle asettaman ehdon

$$\frac{t}{d} \leq 0.1,$$

jossa on t on seinämän paksuus ja d $(= b = h)$ on poikkileikkauksen karakteristinen mitta.

Sijoitetaan solmuviivat kuvan 4 osoittamalla tavalla poikkileikkauksen nurkkiin, jolloin VI-elementteihin tulee viisi seinämänosaa. Kaikkia kerroinfunktioita (siirtymiä ja rotaatioita) interpoloidaan Hermiten C_1-jatkuvien polynomien avulla siten, että syntyy tasainen jako kahdeksaan VI-elementtiin. Interpolaatio poikittaissuunnassa muodostetaan siten, että seinämänosat VI-elementeissä toimivat käytännössä erillisinä elementteinä. Kannattajan akselin suuntainen siirtymäkomponentti u ja seinämää vastaan kohtisuora siirtymäkomponentti w kuvataan molemmat tässä ei-parametrisessa interpolaatiossa Hermiten C_1-jatkuvien polynomien avulla. Poikkileikkauskoordinaatin tangentin suuntaista komponenttia v interpoloidaan puolestaan lineaarisena. Sekä u:lle että w:lle muodostuu seinämänosissa bikuubinen siirtymäyrite, mutta v:n interpolaatiosta tulee epäsymmetrinen.

Jos pystyseinämät puuttuvat ($t_2 = 0$), niin rakenteen vaste muodostuu yksin vaakaseinämän toiminnasta taivutettuna rakenteena (laattana). Jos taas pystyseinämillä on äärellinen paksuus, ne kantavat osan kuormasta toimimalla pääasiassa levyinä. Tutkitaan pysty- ja vaakaseinämien ottamien rasitusten suhdetta kasvattamalla pystyseinämien paksuutta vähitellen nollasta aina kaksinkertaiseksi vaakaseinämän paksuuteen verrattuna. Määritetään näissä tapauksissa jänteen kesikohdalla solmulinjan 2 (tai 3)

taipuman δ_i suhde referenssitaipumaan δ_{ref}, joka on vastaava keskikohdan taipuma silloin, kun vaakaseinämä kantaa kaiken kuorman yksin. Taulukosta 3 nähdään suhteen δ_i/δ_{ref} kehitys, kun pystyseinämien paksuus saa tapauksissa $i = 1..6$ arvot $\underline{i = 1}$: $t_2 = 0$, $\underline{i = 2}$: $t_2 = 0.01t_1$, $\underline{i = 3}$: $t_2 = 0.5t_1$, $\underline{i = 4}$: $t_2 = t_1$ ja $\underline{i = 5}$: $t_2 = 2t_2$.

i	1	2	3	4	5
δ/δ_{ref}	1.00	0.505	0.0204	0.0106	0.00544

Taulukko 3: Kannattajan taipuma jänteen keskellä verrattuna referenssitaipumaan.

Jos levyinä toimivien seinämien paksuus on sadasosa laattaosan paksuudesta ($\underline{i = 2}$: $t_2 = 0.01t_1$), niin rakenteen keskikohdan taipuma on enää puolet siitä, mitä taipuma oli siinä tapauksesta, jossa levyosat puuttuivat ($\underline{i = 1}$: $t_2 = 0$). Jos taas levyinä toimivien seinämien paksuus on puolet laattaosan paksuudesta, niin laskettu taipuma on enää parisen prosenttia referenssitaipumasta ja edelleen, kun pysty- ja vaakaosien seinämänpaksuudet ovat yhtäsuuret ($\underline{i = 4}$: $t_2 = t_1$), niin tämä suhde on enää noin sadasosan luokkaa. Voidaan todeta tässäkin tapauksessa se monille ohutseinämäisille kannattajille tyypillinen piirre, että seinämien taivutuksen osuus rakenteen vasteesta on pieni. Se voidaankin eräissä tapauksissa jättää kokonaan huomioon ottamatta rakennetta analysoitaessa. Tämä tapahtuu tiputtamalla seinämänosaa vastaan kohtisuoralle siirtymäkomponetille w kirjoitettu yrite pois.

Valitulla elementtijaolla laskettu referenssitaipuma ($t_2 = 0$) on $\delta_{ref} \approx 0.023414$ yksikköä. Tämä on vähän pienempi kuin taipuman arvo siinä tapauksessa, jossa Poissonin vakio häviää ($\nu = 0$), jolloin laatan taipuma jänteen puolivälissä on sama kuin vastaavalla palkilla eli $\delta = qL^4/(384EI) \approx 0.024107$. Jälkimmäisen laskemiseksi tarvitaan vain yksi BSF-elementti, joka antaa saman tuloksen kuin palkkiteoria. Tätä voidaan pitää yksinkertaisena testinä laskentaan käytetyn ohjelman oikeellisuudesta

YHTEENVETO

Semidiskreettiin siirtymäyritteeseen perustuvia ratkaisuja käytetään useimmiten silloin, kun tehtävän määrittelyalue on suorakaiteen muotoinen tai sitten sitä rajoittavat saman keskipisteen omaavat ympyräviivat ja niitä vastaan kohtisuorassa olevat (säteen suuntaiset) päädyt. Nämä muodot vastaavat perustapauksia, jollaisiin VI-elementit on tarkoitettu ja joita edellä tekstissä on yksinomaan käsitelty. Sitä, miten elementtien muodon poikkeaminen näistä perusmuodoista (distortio) vaikuttaa saatuihin tuloksiin eräissä erikostapauksissa, kuten aksiaalisuuntaan nähden vinojen tukilinjojen esiintyessä, tutkitaan parhaillaan.

Motiivi diskretoida ohutseinämäinen kannattaja VI-elementein löytyy ennenkaikkea pyrkimyksestä kevyeen rakennemalliin. Tällä on merkitystä, kun rakenne tulee ratkaista yhä uudelleen jonkin tekijän muuttuessa esimerkiksi rakenteen optimoinnin tai yleensä erilaisten iteratiivista menettelyä vaativien tarkastelujen yhteydessä. Myös-

kin suuntien separointiin perustuvan mallin selkeyttä voi pitää menetelmän hyvänä puolena.

LÄHTEET

[1] Cheung Y.K. (1976). *Finite Strip Method in Structural Analysis,* Pergamon Press, Oxford

[2] Crandall S. H. (1956). *Engineering Analysis, A Survey of Numerical Procedures,* McGraw-Hill, New York

[3] Hakala M. K. (1980). *Lujuusopin elementtimenetelmä,* Otapaino, Espoo

[4] Kantorovich L. V. ja Krylov V. I. (1958). *Approximate Methods of Higher Analysis,* P. Noordhoff, Ltd., Groningen

[5] Paavola J. (1990). *A Study of Curved Thin-Walled Girders,* Acta Polytechnica Scandinavica, Civil Engineering and Building Construction Series No. 94, Helsinki

[6] Richards T.H. (1977). *Energy Methods in Stress Analysis,* Ellis Horwood, New York

[7] Sokolnikoff I.S. (1956). *Mathematical Theory of Elasticity,* McGraw-Hill Book Company, Inc., New York

[8] Timoshenko S. P. ja Goodier J. N. (1970). *Theory of Elasticity,* McGraw-Hill Book Company, Inc., New York

[9] Vlasov V.Z. (1961). *Thin-Walled Elastic Beams,* Israel Program for Scientific Translations, Jerusalem Pergamon Press, Oxford

[10] Washizu K. (1980). *Variational Methods in Elasticity & Plasticity,* Pergamon Press, Oxford

FINITE ELEMENT APPROXIMATION OF BIOT'S CONSOLIDATION PROBLEM BY USING BUBBLE FUNCTIONS

J. HARTIKAINEN
Laboratory of Theoretical and Applied Mechanics
Helsinki University of Technology
P.O.Box 1100
FIN-02015 HUT, FINLAND
e–mail: juha.hartikainen@csc.fi

ABSTRACT

The Galerkin formulation stabilized with bubble functions is presented for the Biot's equations governing consolidation of incompressible fluid saturated porous medium. The bubble functions satisfying the equilibrium equations of the Biot's problem and vanishing on the element boundaries are determined by using Taylor series method. Numerical results are presented to demonstrate the performance of the method.

1 INTRODUCTION

The Biot's consolidation problem [1] models settlement of saturated porous medium as fluid percolates out of the pores. The porous medium is considered as a mixture of isotropic elastic solid skeleton and incompressible fluid. The model is based on the conservation laws of mass and linear momentum, where the inertia terms are neglected, on the isotropic Hooke's law and on the Darcy law of fluid flow in porous medium.

The consolidation problem include an incompressibility constraint on the initial condition, so that approximation of the Biot's problem leads to solve an incompressible elasticity or Stokes problem as the time $t \longrightarrow 0^+$. As a consequence, the standard Galerkin method results in spurious oscillations at the early stage of consolidation due to violation of the Babuška-Brezzi stability condition [2] unless a rather restrictive condition for the initial time step is satisfied [3]. Stable finite element formulations can be achieved by using particular pair of displacement and pressure interpolations which meet the stability criterion (see e.g. [4]). Convergence of pressure interpolations can be improved further by employing certain smoothing or post-processing techniques [5], [6]. The so-called Petrov-Galerkin and Galerkin-least-squares formulations (see e.g. [7], [8]) have also been shown to be stable and convergent. They circumvent restrictions of the stability condition allowing low equal-order interpolations for displacements and pressure.

In this presentation, the finite element subspaces of displacements spanned by piecewise linear polynomials are enriched with bubble functions which are chosen to satisfy the equilibrium equations of the Biot's problem in each element and to vanish on the element boundaries. This so-called residual-free-bubble approach (see e.g. [9], [10], [11]) ensures the stability allowing equal-order pair of interpolations as well. The bubble functions are approximated for triangle elements by the application of Taylor series method.

2 PROBLEM STATEMENT

Let a mixture of elastic skeleton and incompressible fluid occupy a bounded domain $\Omega \in \mathbb{R}^2$ with a smooth boundary $\Gamma = \Gamma_g \cup \Gamma_h$ and an outward unit normal $\vec{n}$. The Biot's consolidation problem with given boundary and initial conditions can be formulated as follows: find the displacements $\vec{u}$ of skeleton and the excess pore fluid pressure $\wp$ such that for all time $t \in (0, T]$

$$-\nabla \cdot \sigma^* + \nabla \wp = \vec{0} \tag{1}$$

$$\frac{\partial}{\partial t}(\nabla \cdot \vec{u}) + \nabla \cdot \vec{v} = 0 \qquad \text{in } \Omega \tag{2}$$

$$\sigma^* = 2\mu\epsilon + \lambda \operatorname{tr}(\epsilon)\, I \tag{3}$$

$$\vec{v} = -\kappa \nabla \wp \tag{4}$$

$$\vec{u} = \vec{0}, \quad \wp = 0 \qquad \text{on } \Gamma_g, \tag{5}$$

$$\vec{n} \cdot \sigma^* = \vec{q}, \quad \vec{n} \cdot \vec{v} = 0 \qquad \text{on } \Gamma_h, \tag{6}$$

$$\nabla \cdot \vec{u} = 0 \qquad \text{in } \Omega \text{ at } t = 0. \tag{7}$$

Above, the effective stress tensor σ^* is defined in terms of the strain tensor ϵ by the isotropic Hooke's law (3), where μ and λ are the Lamé parameters of skeleton. The strain tensor is the symmetric part of the displacement gradients: $\epsilon = \frac{1}{2}\left(\nabla \vec{u} + \nabla^{\mathsf{T}} \vec{u}\right)$. The flow of fluid through the skeleton is governed by the Darcy's law (4), where κ is the hydraulic conductivity of the porous medium.

3 FINITE ELEMENT FORMULATION

Let the space-time $\Omega \times T$ be divided into time slabs $\Omega \times]t_n, t_{n+1}[$ with boundaries $\Gamma \times]t_n, t_{n+1}[$, where $n = 0, \ldots, N-1$, and in turn, let each time slab be partitioned into space-time elements $\Omega^e \times]t_n, t_{n+1}[$, $e = 1, \ldots, n_{el}$, where n_{el} is the number of elements. By employing the time-discontinuous Galerkin method (see e.g. [12]), the space-time finite element formulation of the problem above reads as follows: find $\vec{u}_h$ and $\wp_h$ such that

$$\sum_{e=1}^{n_{el}} \int_{t_n}^{t_{n+1}} \int_{\Omega^e} \left[\epsilon(\vec{w}_h) 2\mu \epsilon(\vec{u}_h) + \operatorname{tr}(\epsilon(\vec{w}_h)) \lambda \operatorname{tr}(\epsilon(\vec{u}_h)) + \vec{w}_h \cdot \nabla \wp_h\right] \mathrm{d}\Omega \mathrm{d}t \tag{8}$$

$$= \sum_{e=1}^{n_{el}} \int_{t_n}^{t_{n+1}} \int_{\Gamma_h^e} \vec{w}_h \cdot \vec{h}_\sigma \mathrm{d}\Gamma \mathrm{d}t \qquad \text{for all } \vec{w}_h,$$

$$\sum_{e=1}^{n_{el}} \int_{t_n}^{t_{n+1}} \int_{\Omega^e} \left(-\tfrac{\partial}{\partial t} q_h \nabla \cdot \vec{u}_h + \nabla q_h \cdot \kappa \nabla \wp_h\right) \mathrm{d}\Omega \mathrm{d}t + \tag{9}$$

$$+ \sum_{e=1}^{n_{el}} \int_{\Omega^e} \left[q_h(t_{n+1}^-)\, \nabla \cdot \vec{u}_h(t_{n+1}^-) - q_h(t_n^+)\, \nabla \cdot \vec{u}_h(t_n^-)\right] \mathrm{d}\Omega$$

$$= \sum_{e=1}^{n_{el}} \int_{t_n}^{t_{n+1}} \int_{\Gamma_h^e} -q_h h_v \mathrm{d}\Gamma \mathrm{d}t \qquad \text{for all } q_h.$$

The second integral in (9) enforces weakly the continuity of the solution in time from one space-time slab to the next, so that the problem can be solved in one space-time slab at a time.

The finite element functions $\wp_h$ and q_h are assumed to be piecewise linear polynomials whereas the finite element functions for displacements are composed of piecewise linear polynomials and bubble functions, i.e.

$$\vec{u}_h = \vec{u}_l + \vec{u}_b. \tag{10}$$

Since the functions $\vec{u}_h$ can be split uniquely by (10) and on the condition that the bubble functions vanish on the element boundaries, the Galerkin formulae (8) and (9) can be rewritten as follows: find $\vec{u}_h = \vec{u}_l + \vec{u}_b$ and $\wp_h = \wp_l$ such that

$$\sum_{e=1}^{n_{el}} \int_{t_n}^{t_{n+1}} \int_{\Omega^e} \left\{\epsilon(\vec{w}_l) 2\mu\epsilon(\vec{u}_l) + \mathrm{tr}(\epsilon(\vec{w}_l))\lambda\, \mathrm{tr}(\epsilon(\vec{u}_l)) + \vec{w}_l \cdot \nabla \wp_l + \right. \tag{11}$$

$$\left. - \nabla \cdot [2\mu\epsilon(\vec{w}_l) + \lambda\, \mathrm{tr}(\epsilon(\vec{w}_l))\, I] \cdot \vec{u}_b \right\} \mathrm{d}\Omega \mathrm{d}t = \sum_{e=1}^{n_{el}} \int_{t_n}^{t_{n+1}} \int_{\Gamma_h^e} \vec{w}_l \cdot \vec{h}_\sigma \mathrm{d}\Gamma \mathrm{d}t \qquad \text{for all } \vec{w}_l,$$

$$\sum_{e=1}^{n_{el}} \int_{t_n}^{t_{n+1}} \int_{\Omega^e} \left[\epsilon(\vec{w}_b) 2\mu\epsilon(\vec{u}_l + \vec{u}_b) + \mathrm{tr}(\epsilon(\vec{w}_b))\lambda\, \mathrm{tr}(\epsilon(\vec{u}_l + \vec{u}_b)) + \right. \tag{12}$$

$$\left. + \vec{w}_b \cdot \nabla \wp_l \right] \mathrm{d}\Omega \mathrm{d}t = 0 \qquad \text{for all } \vec{w}_b,$$

$$\sum_{e=1}^{n_{el}} \int_{t_n}^{t_{n+1}} \int_{\Omega^e} \left(-\tfrac{\partial}{\partial t} q_l \nabla \cdot \vec{u}_l + \nabla q_l \cdot \kappa \nabla \wp_l + \tfrac{\partial}{\partial t} \nabla q_l \cdot \vec{u}_b\right) \mathrm{d}\Omega \mathrm{d}t + \tag{13}$$

$$+ \sum_{e=1}^{n_{el}} \int_{\Omega^e} \left[q_l(t_{n+1}^-)\, \nabla \cdot \vec{u}_l(t_{n+1}^-) - q_l(t_n^+)\, \nabla \cdot \vec{u}_l(t_n^-) - \nabla q_l(t_{n+1}^-) \cdot \vec{u}_b(t_{n+1}^-) + \right.$$

$$\left. + \nabla q_l(t_n^+) \cdot \vec{u}_b(t_n^-) \right] \mathrm{d}\Omega = \sum_{e=1}^{n_{el}} \int_{t_n}^{t_{n+1}} \int_{\Gamma_h^e} -q_l h_v \mathrm{d}\Gamma \mathrm{d}t \qquad \text{for all } q_l.$$

The equations (11) and (13) form the method to compute the improved approximations for the Biot's consolidation problem due to the effect of bubble functions while the equation (12) can be considered the Galerkin formula for the following problem dealing with the bubble functions. Given $\vec{u}_l$ and $\wp_l$ find $\vec{u}_b$ such that for all $t \in]t_n, t_{n+1}[$

$$\nabla \cdot [2\mu\epsilon(\vec{u}_b) + \lambda\, \mathrm{tr}(\epsilon(\vec{u}_b))\, I] = -\nabla \cdot [2\mu\epsilon(\vec{u}_l) + \lambda\, \mathrm{tr}(\epsilon(\vec{u}_l))\, I] + \nabla \wp_l \qquad \text{in } \Omega^e, \tag{14}$$

$$\vec{u}_b = \vec{0} \qquad \text{on } \Gamma^e. \tag{15}$$

The exact solution of the problem (14), (15) for any given μ and λ would give the optimal bubble functions in terms of the functions $\vec{u}_l$ and $\wp_l$. However, in general, this can be done merely in an approximate way. Hence, a simple procedure is introduced to determine the bubble functions in the following. The idea is to construct difference formulae for the differential equations (14) by using Taylor series method. The bubble functions are expanded at number of m points on the element boundary in Taylor series about the point P within the element, i.e.

$$\vec{u}_b(x_{\Gamma^e}) = \sum_{\alpha} \frac{1}{\alpha!} D^{\alpha} \vec{u}_b(x_P)(x_{\Gamma^e} - x_P), \tag{16}$$

where the following notation is used: $x = (x_1, x_2)$, $\alpha = (\alpha_1, \ldots, \alpha_n)$, $|\alpha| = \alpha_1 + \cdots + \alpha_n$, $\alpha! = \alpha_1!\alpha_2! \cdots \alpha_n!$ and $D^{\alpha} = \frac{\partial |\alpha|}{\partial x_1^{\alpha_1} \cdots x_2^{\alpha_n}}$. Then the derivatives D^{α} are determined in terms of the values of the functions $\vec{u}_b(x_P)$ and $\vec{u}_b(x_{\Gamma^e})$. In order to determine the derivatives up to order of n at least number of $m = \sum_{i=1}^{n+1} i$ boundary points are needed.

For an example, one of the simplest pair of difference equations approximating the differential equations (14) for linear triangle element is

$$\sum_{i=1}^{3} \left[(2\mu + \lambda) \frac{(L^i_{x_1})^2}{L^i} + \mu \frac{(L^i_{x_2})^2}{L^i} \right] \tilde{u}_{b_1} + \sum_{i=1}^{3} \mu \frac{L^i_{x_1} L^i_{x_2}}{L^i} \tilde{u}_{b_2} = -\tfrac{1}{2} \sum_{i=1}^{3} L^i_{x_1} \wp^i + \mathcal{O}(h), \tag{17}$$

$$\sum_{i=1}^{3} \mu \frac{L^i_{x_1} L^i_{x_2}}{L^i} \tilde{u}_{b_1} + \sum_{i=1}^{3} \left[\mu \frac{(L^i_{x_1})^2}{L^i} + (2\mu + \lambda) \frac{(L^i_{x_2})^2}{L^i} \right] \tilde{u}_{b_2} = -\tfrac{1}{2} \sum_{i=1}^{3} L^i_{x_2} \wp^i + \mathcal{O}(h), \tag{18}$$

where L^i are the common area coordinates (see Figure 1). The difference approximations of the bubble functions are denoted by $\vec{\tilde{u}}_b = (\tilde{u}_{b_1}, \tilde{u}_{b_2})$ and the nodal fluid pressure values by $\wp^i$. $\mathcal{O}(h)$ is the formal truncation error and $L^i_{x_1}$ and $L^i_{x_2}$ are the components of the gradient DL^i.

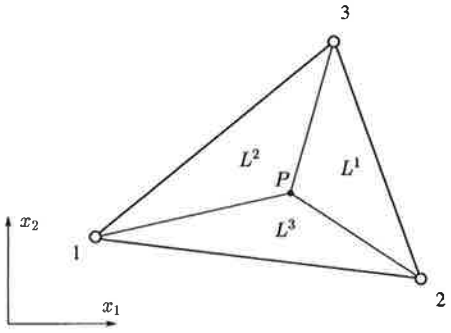


Figure 1: Triangle element. Area coordinates.

3 EXAMPLE

A strip loaded water saturated finite clay layer on a rigid impervious base is a common two-dimensional test case. The problem definition is shown in Figure 2 and the following material parameters are used: the elasticity modulus $E = 5$ MPa, the Poisson's ratio $\nu = 0.3$, the permeability $k = 10^{-9}$ ms^{-1} and the weight of water $\gamma = 10$ kNm^{-3}. The Lame parameters and the hydraulic conductivity are determined by the relations: $\mu = \frac{E}{2(1+\nu)}$, $\lambda = \frac{\nu E}{(1+\nu)(1-2\nu)}$ and $\kappa = \frac{k}{\gamma}$. A uniform mesh of quadrangles with size of $h = H/10$ subdivided into two triangle elements is employed and constant interpolations in time are used. The water pressure distributions after the initial time step $\Delta t_o = 10$ s obtained with the standard Galerkin method and the bubble stabilized method are shown in Figures 3 and 4.

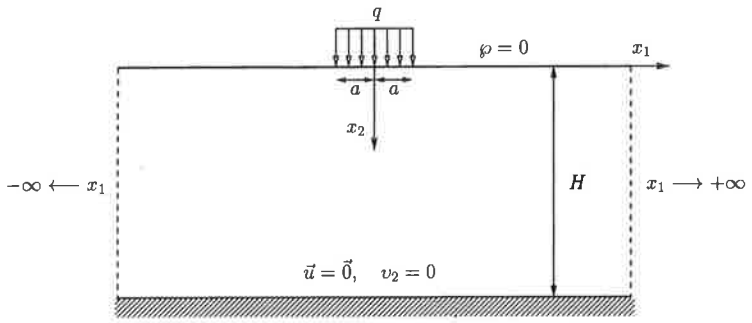


Figure 2: Two-dimensional consolidation problem. A strip loaded water saturated finite clay layer on a rigid impervious base. ($H = 5$ m, $a = 0.5$ m, $q = 100$ kNm^{-2}).

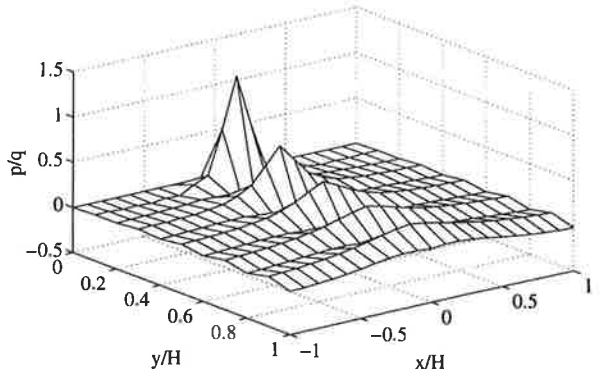


Figure 3: Water pressure distribution at time $t = 10$ s. Standard Galerkin method. ($x = x_1, y = x_2$).

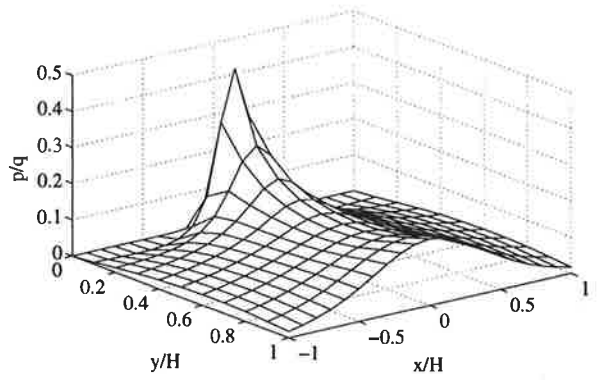


Figure 4: Water pressure distribution at time $t = 10$ s. Bubble stabilized method. ($x = x_1$, $y = x_2$).

3 CONCLUSIONS

A bubble stabilized finite element formulation for the Biot's consolidation problem has been proposed. Stable method is obtained without any numerical tricks and additional mesh-dependent terms. Preliminary numerical results show that the oscillations of the pressure field can be eliminated completely. However, further studies are needed to assure the performance of the method.

REFERENCES

[1] M. A. Biot. General theory of three-dimensional consolidation. *J. Appl. Phys.*, 12:155–164, 1941.

[2] F. Brezzi. On the existence, uniqueness and approximation of saddle-point problems arising from lagrange multipliers. *R.A.I.R.O. Anal. Numér.*, 8(R-2):129–151, 1974.

[3] P. A. Vermeer and A. Verruijt. An accuracy condition for consolidation by finite elements. *Int. J. Num. Anal. Meths. Geom.*, 5:1–14, 1981.

[4] F. Brezzi and M. Fortin. *Mixed and hybrid finite element methods*, volume 15 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, New-York, 1991.

[5] M. B. Reed. An investigation of numerical errors in the analysis of consolidation by finite elements. *Int. J. Num. Anal. Meths. Geom.*, 8:243–257, 1984.

[6] M. A. Murad and A. F. D. Loula. Improved accuracy in finite element analysis of biot's consolidation problem. *Comput. Meths. Appl. Mech. Engrg.*, 95:359–382, 1992.

[7] T. J. R. Hughes, L. P. Franca, and M. Balestra. A new finite element formulation for computational fluid dynamics: V. circumventing the babuška-brezzi condition: A stable petrov-galerkin formulation of the stokes problem accommodating equal-order interpolations. *Comput. Meths. Appl. Mech. Engrg.*, 59:85–99, 1986.

[8] A. Truty. A galerkin least-squares fe-formulation for consolidation. In *ECCM '99*, München, Germany, 1999. CD-ROM Proceedings of European Conference on Computational Mechanics.

[9] F. Brezzi, Franca, T. J. R. L. P., Hughes, and A. Russo. $b = \int g$. *Comput. Methods Appl. Mech. Engrg.*, 145:329–339, 1997.

[10] L. P. Franca and A. Russo. Approximation of the stokes problem by residual-free macro bubbles. *East-West J. Num. Anal.*, 4:265–278, 1996.

[11] F. Brezzi, T. J. R. Hughes, A. Russo, and E. Süli. A priori error analysis of residual-free bubbles for advection-diffusion problems. *SIAM J. Numer. Anal.*, 36(6):1933–1948, 1999.

[12] T. J. R. Hughes and G. M. Hulbert. Space-time finite element methods for elastodynamics: formulations and error estimates. *Comput. Meths. Appl. Mech. Engrg.*, 66:339–363, 1988.

GEOSPHERE TRANSPORT OF RADIONUCLIDES IN SAFETY ASSESSMENT OF SPENT FUEL DISPOSAL

Petri Jussila

Helsinki University of Technology
Laboratory of Theoretical and Applied Mechanics
P.O.Box 1100
Fin-02015 HUT Finland

ABSTRACT

Radionuclide transport in rock is complicated. The flow paths are difficult to characterise and the phenomena involved are various. Critical paths along which radionuclides can reach the biosphere are considered in modelling. The calculations of the present study are performed for an analytical single nuclide model taking into account 1D advective transport along a fracture, 1D diffusion within the porous rock surrounding the fracture, retardation within rock, and radioactive decay. The results are compared to those obtained by VTT and presented in TILA-99 safety assessment report.

1 INTRODUCTION

One of the basic problems of the final disposal of nuclear waste is how to ensure isolation of the waste from biosphere for great time spans. Moving groundwater and the prevailing chemical conditions can induce degradation of the fuel disposed and transport of the released radionuclides to the biosphere via a fracture network in the geosphere. From among the wide field of research the problem of radionuclide transport through the geosphere is considered. The actual situation is complex, the flow paths in the geosphere are difficult to characterise and there are various phenomena involved. In a mathematical model, a critical path along which radionuclides can reach the biosphere is considered. The worst predictable cases and the effect of the parameters can be studied with the help of the model although it simplifies the actual situation considerably. The presentation is based on a study [1] utilising a widely used model of radionuclide transport in porous media. The calculation results have been compared to those obtained by Technical Research Centre of Finland (VTT) published in TILA-99 safety assessment report [2]. The study is associated with a research project of the Radiation and Nuclear Safety Authority (STUK) to utilise analytical models in safety and performance assessment for geological disposal of spent nuclear fuel.

2 TRANSPORT MODEL

The system consists of a fracture through which water moves between two identical semi-infinite saturated porous rock matrices (Figure 1.). There is a steady nuclide source of constant volume C_0 at $x = y = 0$. The following processes are considered:

- advective transport along the fracture,
- mechanical dispersion in the fracture,
- diffusion within the fracture, in the direction of the fracture axis,
- diffusion from the fracture into and within the matrix,
- adsorption on the fracture surfaces,
- adsorption within the matrix and
- radioactive decay.

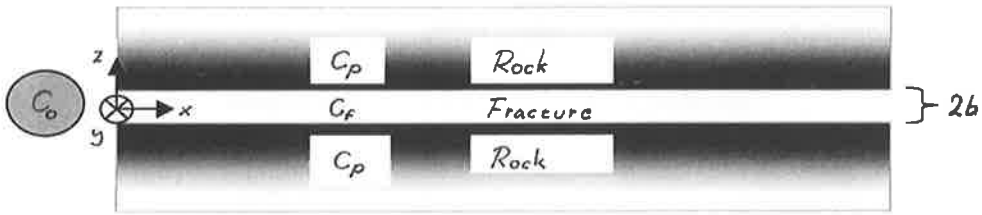


Figure 1. The geometry of the transport problem

Nuclides can either be dissolved in water and mobile or adsorbed to a solid phase and immobile. The number of variables is reduced by using a modelling tool called a *linear equilibrium isotherm*, according to which the ratio of adsorbed nuclide inventory and the dissolved inventory is defined as a distribution coefficient [1]. Assumptions used in 1D transport model for a single nuclide are the following.

- The fracture is thin and rigid with stationary characteristics.
- The fracture and the porous rock are saturated.
- The water velocity in the fracture is constant and directed along the fracture (x-axis).
- At the origin of the fracture there exists a contaminant source of constant strength.
- Decay chains are not taken into account; the model is for a single nuclide.
- Water and rock characteristics do not depend on position.

Assumptions on the geometry and hydraulic properties are the following.

- The width (aperture $2b$) of the fracture is much smaller than its length.
- Transverse diffusion and dispersion within the fracture assure complete mixing across the fracture width at all times.
- The permeability of the porous matrix is very low and transport in the matrix occurs mainly by molecular diffusion.
- Transport along the fracture is much faster than transport within the matrix.

This results in the simplification of the basically 2D system to two orthogonal, coupled 1D systems [1]:

$$\frac{\partial C_f(x,t)}{\partial t} = \frac{D_f}{R_f}\frac{\partial^2 C_f(x,t)}{\partial x^2} - \frac{v}{R_f}\frac{\partial C_f(x,t)}{\partial x} - \lambda C_f(x,t) + \frac{2}{2b}\frac{\varepsilon D_p}{R_f}\frac{\partial C_p(x,z,t)}{\partial z}\bigg|_{|z|=b}; |z| = b, \tag{1}$$

$$\frac{\partial C_p(x,z,t)}{\partial t} = \frac{D_p}{R_p}\frac{\partial^2 C_p(x,z,t)}{\partial z^2} - \lambda C_p(x,z,t), \quad |z| \geq b, \tag{2}$$

with the boundary and initial conditions

$$C_f(x,t) = C_p(x,z,t); \quad |z| = b, \tag{3}$$

$$C_f(x=0,t) = C_0,\ C_p(x,z,t=0) = 0,\ \lim_{x\to\infty} C_f(x,t) = 0,\ \lim_{z\to\pm\infty} C_p(x,z,t) = 0, \tag{4}$$

where C_f is the concentration of dissolved nuclides in fracture fluid (mol/m^3), C_p is the concentration of dissolved nuclides in the rock matrix fluid (mol/m^3), R_f is the surface retardation coefficient in the fracture (-), R_p is the matrix retardation coefficient (-), D_f is the hydrodynamic dispersion coefficient (m^2/s), D_p is the diffusion coefficient in the pore structure of rock matrix (m^2/s), v is the water velocity (m/s), ε is the porosity of the rock (-), and λ is the decay constant of a nuclide (1/s). The retardation coefficients are defined as

$$R_f = 1 + \frac{2}{2b}K_a, \qquad R_p = 1 + \frac{\rho_R(1-\varepsilon)}{\varepsilon}K_d, \tag{5}$$

where K_a is the area based distribution coefficient (m^3/m^2), K_d is the volume based distribution coefficient (m^3/kg), and ρ_R is the density of solid rock (kg/m^3).

The model is the same as presented in [3, p. 557], except the inclusion of the absolute value of the z-co-ordinate, which takes into account the fact that diffusive loss occurs to both the fracture walls. The solution of the 1D problem (1) in the fracture for a single nuclide with no hydrodynamical dispersion, i.e., with $D_f = 0$, is [1]

$$\frac{C_f(x,t)}{C_0} = 0, \quad t < t_0,$$

$$\frac{C_f(x,t)}{C_0} = \frac{1}{2}e^{-\lambda t_0}\left[e^{-\frac{\sqrt{\lambda}t_0}{A}}\operatorname{erfc}\left(\frac{t_0}{2A\sqrt{t-t_0}} - \sqrt{\lambda(t-t_0)}\right) + e^{\frac{\sqrt{\lambda}t_0}{A}}\operatorname{erfc}\left(\frac{t_0}{2A\sqrt{t-t_0}} + \sqrt{\lambda(t-t_0)}\right)\right], \quad t > t_0, \tag{6}$$

where we have the abbreviations

$$t_0 = \frac{R_f}{v} x, \qquad A = \frac{2b}{2} \frac{R_f}{\varepsilon \sqrt{R_p D_p}}. \tag{7}$$

3 VTT APPROACH

In the VTT studies [2, p. 114] the transport model is reduced for illustrational purposes. The aim has been to study releases at a certain distance, say, $x = L$ (m), from the source of the contaminant. A stable nuclide solution for the fracture can be reduced to [1]

$$\frac{C_f(L,t)}{C_0} = \operatorname{erfc}\left(\frac{t_w(L)}{2b} \sqrt{\varepsilon D_e R_p} \cdot \sqrt{1/t} \right) = \operatorname{erfc}\left(u(L) \cdot t^{-1/2} \right), \quad t > R_f t_w(L), \tag{8}$$

where $D_e = \varepsilon D_p$ is the effective diffusion coefficient from fracture to the matrix (m^2/s) and $t_w = L/v$ is the groundwater transit time (s or a). This approximation also removes the fracture surface distribution coefficient K_a from the model. The u-parameter in (8) describing the properties of a transport route for a given species

$$u(L) = \sqrt{\varepsilon D_e R_p} \cdot \frac{t_w(L)}{2b} = \sqrt{\varepsilon D_e R_p} \cdot \frac{WL}{Q}, \tag{9}$$

where W is the width of the flow channel (m) and Q is the flow rate in the channel (m^3/s), is a product of two terms. The first one is nuclide specific and takes into account the effects of matrix diffusion on the transport. The second factor called the transport resistance describes the groundwater flow distribution in the route.

The parameters for modelling purposes (except the decay constant λ) are involved in (9) and (5): the matrix related D_e, K_d, ρ_R, ε, and flow channel related transport resistance WL/Q. The density of rock is always $\rho_R = 2700$ kg/m^3.

Table 1. Distribution coefficients (K_d) in the rock matrix (m^3/kg) [2, Table 11-9, p. 118] for the nuclides involved in the calculations of the present study

Element	Conservative non-saline reducing	Conservative saline reducing
C	0.0001	0.0001
Cl	0	0
Ni	0.1	0.005
Se	0.0005	0.0001
Sr	0.005	0.0001
Zr	0.2	0.2
Nb	0.02	0.02
Tc	0.05	0.05
Pd	0.001	0.0001
Sn	0.001	0.0001
I	0	0
Cs	0.05	0.01

Values of K_d are defined for five sets of conditions in TILA-99. Only two cases of these are used in the reference scenarios of TILA-99, namely, *conservative values in reducing*

conditions in non-saline waters and *conservative values in reducing conditions in saline waters*. The values used for those cases for the nuclides involved in the present study are presented in Table 1. In TILA-99, the penetration depth of rock is 10 cm and the altered rock zone within 1 cm from the fracture is given different values than the intact rock. For the porosity and effective diffusion coefficient values used see [1] or [2, Table 11-10, p. 120]. Various transport resistance values have been used in TILA-99. The "median" value, involved in the following is $WL/Q = 5{\cdot}10^4$ a/m. With a volume aperture of $2b = 5{\cdot}10^{-4}$ m, it corresponds to a groundwater transit time $t_w = 25$ a.

The transport analyses in TILA-99 are performed with the FTRANS code [4], the input parameters L, v, $2b$ of which are fixed in TILA-99 in such a way that the chosen value for transport resistance is obtained.

4 CALCULATIONS

Different approaches – implications to choice of parameter values

Both the approaches of TILA-99 and the present study employ a geosphere transport model in which the same phenomena are incorporated. Main differences between the approaches are collected in Table 2.

Table 2. The main differences between the calculation model used in this report and the FTRANS code used in TILA-99 by VTT

This report; equation (6)	**TILA-99; FTRANS**
Analytical model	Numerical model – discretisations
Infinite domain with single rock parameter values	Two finite domains with differing rock parameter values
$R_T = WL/Q = t_w/2b$ is given by means of t_w and $2b$	$R_T = WL/Q = t_w/2b$ is given by means of L, Q (v) and $2b$ [2, p. 120]
Single nuclide	Decay chains

The most considerable difference between the approaches is the doubly finite domain used in TILA-99 and the infinite domain used in the present study. VTT has used different parameter values for each domain in TILA-99 and only single values can be used in the approach of the present study.

According to a correspondence with T. Vieno and H. Nordman of VTT, the nuclides can be divided into two categories according to their values of distribution coefficient K_d. The parameter values for the 0-1 cm zone of the rock can be used for the retarding nuclides for which $K_d > 1{\cdot}10^{-4}$, because they are mainly retarded in this zone. On the other hand, for the weakly or non-retarding nuclides for which $K_d \leq 1{\cdot}10^{-4}$ the parameter values for the 1-10 cm zone of the rock can be used. This division is rough, of course, and the behaviour of an output does not depend solely on the K_d value but also, e.g., on the shape of the input. [5]

The rock parameter and K_d values in this study are chosen to be the same as in TILA-99. The transport resistance gets the value $WL/Q = t_w/2b = 5{\cdot}10^4$ a/m in the base cases.

Sensitivity analysis for transport resistance

To investigate the influence of the transport resistance on the result the value of it is varied for some nuclides in a particular case. In TILA-99 the transport resistance is given the values $WL/Q \in [5 \cdot 10^3, 5 \cdot 10^4]$ a/m [2, p. 132]. When the aperture is kept constant ($2b = 5 \cdot 10^{-4}$ m), this corresponds to a transit time range $t_w \in [2.5 \cdot 10^0, 2.5 \cdot 10^1]$ a. In this study, we have chosen two values used in TILA-99 corresponding to the "median" $WL/Q = 5 \cdot 10^4$ a/m and the "very high flow in saline conditions" $WL/Q = 1 \cdot 10^4$ a/m cases. Additionally, two values are chosen beyond this range but within the range of the measurement and simulation estimates of TILA-99 to assess the effect of this parameter. The variation is done by keeping the aperture constant and varying the transit time (Table 3.).

Table 3. The transport resistance values WL/Q (a/m) and the corresponding transit time range t_w (a) used in the sensitivity analysis of the present study

WL/Q (a/m)	t_w (a)
$1.0 \cdot 10^3$	0.50
$1.0 \cdot 10^4$	5.0
$5.0 \cdot 10^4$	25
$5.0 \cdot 10^5$	50

Chosen cases

In TILA-99 [2, p. 135-147], a detailed analysis of two scenarios is presented:

1. SH-sal50, i.e., small initial hole in the canister, median flow and transport data, saline groundwater and
2. DC-ns50, i.e., canister disappearing at 10 000 years, median flow and transport data, non-saline groundwater.

The results are for a single canister containing 2.14 tU spent fuel from the Olkiluoto reactors. Because our analytical model can not handle decay chains, we have chosen some single nuclides from among the nuclides involved in these cases. These nuclides are

- C-14, Cl-36, Ni-59, Ni-63, Se-79, Sr-90, Nb-94, Tc-99, Sn-126, I-129, Cs-135 and Cs-137 in SH-sal50 scenario and
- C-14, Cl-36, Ni-59, Se-79, Zr-93, Nb-94, Tc-99, Pd-107, Sn-126, I-129 and Cs-135 in DC-ns50 scenario.

For the half lives of the nuclides involved in this study see [1] or [2, p. 21].

For variation of t_w the results are calculated for three nuclides of different type: non-retarding Cl-36, retarding Ni-59 and quickly decaying Sr-90. The variant cases are based on the SH-sal50 scenario.

The calculation program

The solution (6) used in the present study is for a constant infinite inlet condition, i.e., for a step input. An output for a finite rectangular input pulse with a duration DT(a) can be

constructed by calculating the difference of outputs of two identical step inputs shifted DT from apart [6]. An output for a more arbitrary input can be constructed by approximating the input by rectangular pulses of given height (release rate (Bq/a)) and width (duration (a)). The program used receives a release rate vector to the geosphere and the corresponding time vector as an input, treats this data as individual rectangular inputs, calculates the individual outputs of these inputs and gives the final result output as the sum of the individual outputs.

5 RESULTS

The significant result outputs with the corresponding TILA-99 results of the SH-sal50 and DC-ns50 scenarios are presented in Figure 2. and in Figure 3., respectively. For more results see [1]. The result outputs of Cl-36, Ni-59, and Sr-90, when t_w is varied in cases based on the SH-sal50 scenario are presented in Figure 4.

The results for the SH-sal50 and DC-ns50 scenarios are quite similar to those in TILA-99. Reasons for the deviations are now discussed on the basis of the differences in the calculation approaches.

The parameter values of intact rock were used for weakly and non-retarding nuclides whereas the parameter values of more porous altered rock zone were used for more retarding nuclides. I.e., for weakly and non-retarding nuclides the rock was generally more solid and for retarding nuclides generally more porous when compared to the TILA-99 calculations. Consequently, the weakly and non-retarding nuclides should diffuse and be retarded less and the retarding nuclides should diffuse and be retarded more in the present approach compared to the TILA-99 calculations. The expected result is, that the outputs of the weakly and non-retarding nuclides

- C-14, Cl-36, Se-79, Sr-90, Pd-107, Sn-126 and I-129 in SH-sal50 scenario and
- C-14, Cl-36 and I-129 in DC-ns50 scenario

should be larger and/or less retarded than the corresponding TILA-99 outputs and the outputs of the retarding nuclides

- Ni-59, Ni-63, Nb-94, Tc-99, Cs-135 and Cs-137 in SH-sal50 scenario and
- Ni-59, Se-79, Zr-93, Nb-94, Tc-99, Pd-107, Sn-126, and Cs-135 in DC-ns50 scenario

are expected to be smaller and/or more retarded than the corresponding TILA-99 outputs. This theoretical expectation is fulfilled by all the results in Figure 2. and Figure 3.

This effect is further assessed by interchanging the rock parameter values, i.e., by giving weakly retarding nuclides the values of altered rock and the retarding nuclides those of intact rock. In Figure 5. this is done for some nuclides of various retardation type. It can be seen that the TILA-99 results occur in between the previous results and the results for interchanged rock parameter values, which was also expected. The results of TILA-99 for sorbing species, e.g., Ni-59 in Figure 5., differ approximately as much from the results of the base cases and the cases where the rock parameter values are interchanged between the rock zones. This implies, that the rough estimation that the nuclides for which $K_d > 1 \cdot 10^{-4}$ see essentially only the altered rock zone is, indeed, too rough.

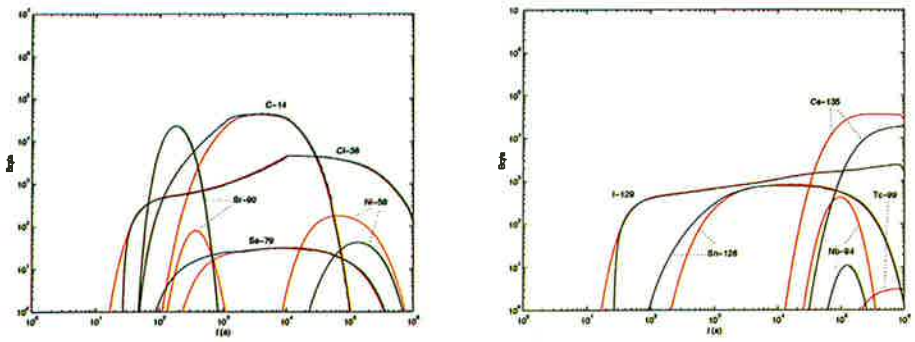


Figure 2. Release rates of fission and activation products in SH-sal50 scenario

Red: TILA-99 output; **Black**: Result output of the present study

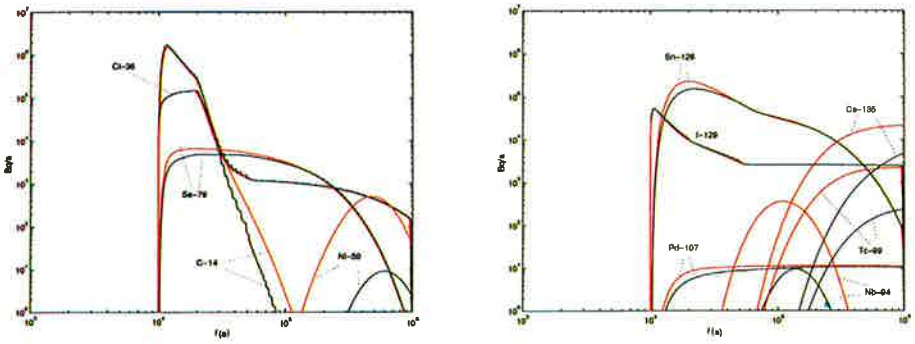


Figure 3. Release rates of fission and activation products in DC-ns50 scenario

Red: TILA-99 output; **Black**: Result output of the present study

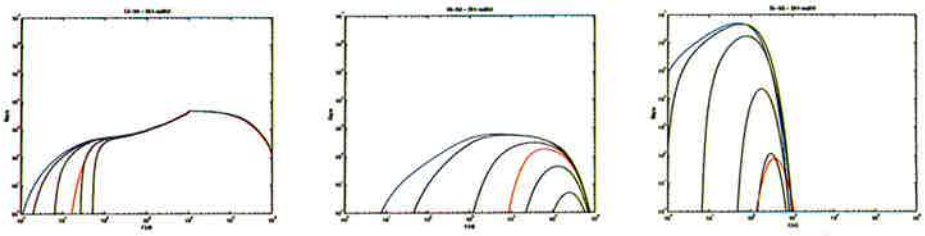


Figure 4. Variation of transit time in cases based on the SH-sal50 scenario

Blue: TILA-99-input; **Red:** TILA-99-output; **Black** (from left to right in each figure): t_w = 0.5, 5, 25, 50 a

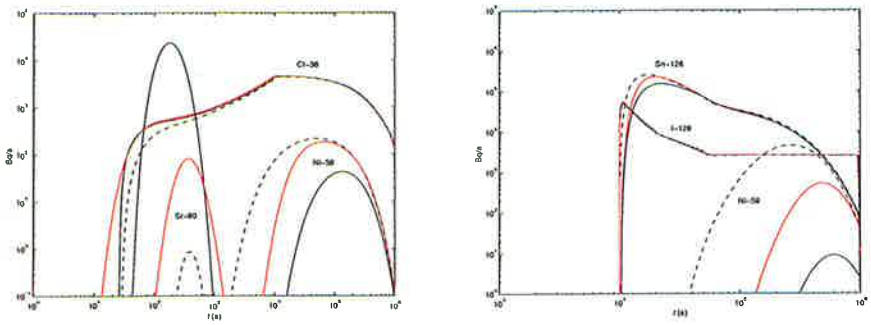


Figure 5. Release rates of Cl-36, Ni-59 and Sr-90 in the SH-sal50 scenario (left) and Ni-59, Sn-126 and I-129 in the DC-ns50 scenario (right)

Red: TILA-99 output; **Black**: Present result; Dashed: Present result with the rock parameters interchanged between the rock zones

For weakly and non-sorbing nuclides the effect of the rectangular approximation of the input is the most noticeable. If the calculation lattice is dense, the outputs of the individual rectangular input pulses are also quite edgy and do not overlap considerably. Consequently, the summed result outputs are also edgy, see, e.g., I-129, Cl-36, and C-14 in Figures 2. and 3. On the contrary, for sorbing nuclides the outputs of the individual rectangular input pulses get significant values in most cases only after $t = t_w$+DT years after the beginning of the respective individual input pulse. Consequently, the individual output pulses overlap and the summed result outputs are smooth.

The value of t_w has a considerable effect for some nuclides in the sensitivity analysis (Figure 4.). The effect is the most significant in the case of quickly decaying Sr-90. The effect is the smallest for the non-retarding Cl-36, for which only the starting point of the output is shifted as t_w.

6 CONCLUSION AND DISCUSSION

Calculations were done for a single radionuclide transport model for which advective transport along the fracture, diffusion from the fracture into and within the rock matrix, retardation within the matrix, and radioactive decay were taken into account. The results were compared to the results of the same calculation cases obtained by VTT and presented in TILA-99 safety assessment report. In addition, the effect of the value of transport resistance was assessed by varying the groundwater transit time.

The main difference in the approaches of TILA-99 and the present study is the use of a numerical model with two finite domains with different rock parameter values in TILA-99 and an analytical model with one infinite and homogenous domain with single rock parameter values in the present study. The differences of the results of these models can to a large extent be explained by this main difference in the modelling approaches. In addition, the results of the present study are also similar to those presented in TILA-99. Consequently, confidence on the results presented in TILA-99 is increased.

The effect of the rock parameter values is found to be significant for some cases in the present study. Variation of the transport resistance by varying the groundwater transit time affects the result in some cases considerably. The results are the most sensitive for the nuclides that have a small half life compared to the transit time. The simple model is the least applicable to these cases, however. Decreasing of the transit time too much is unrealistic, because actual groundwater velocities are not very large. On the other hand, increasing of the transit time conflicts with the model assumption of the transport being one-dimensional in the directions of the fracture and of the matrix.

Modelling natural phenomena involves large uncertainties in the actual nature of the phenomena as well as in the values of the appropriate parameters. Some of the main differences between the transport model used and the reality are the mathematical characterisation of the flow route in rock as a smooth and straight fracture and the modelling of the complicated chemical processes causing retardation with the help of a distribution coefficient. The model does not represent the actual situation well but using it provides understanding of the key characteristics involved.

7 REFERENCES

[1] P. Jussila: *Geosphere transport of radionuclides in safety assessment of spent fuel disposal.* Master's Thesis, Helsinki University of Technology, Department of Engineering Physics and Mathematics, Espoo, 1999.

[2] T. Vieno, H. Nordman: *Safety assessment of spent fuel disposal in Hästholmen, Kivetty, Olkiluoto and Romuvaara. TILA-99.* POSIVA-99-07, Posiva Oy, Helsinki 1999.

[3] D.H. Tang, E.O. Frind, E.A. Sudicky: *Contaminant Transport in Fractured Porous Media: Analytical Solution for a Single Fracture.* Water Resources Research, Vol. 17, No. 3, Pages 555-564, June 1981.

[4] FTRANS. *A two-dimensional code for simulating fluid flow and transport of radioactive nuclides in fractured rock for repository performance assessment.* Intera Environmental Consultants Inc., Report ONWI-426, Houston, 1983.

[5] T. Vieno, H. Nordman. VTT Energy. E-mail correspondence, August 1999.

[6] A. Hautojärvi: *Matriisidiffuusion vaikutus rakokulkeutumisessa.* Helsinki, Valtion teknillinen tutkimuskeskus, Ydinvoimatekniikan laboratorio, Työraportti TOKA-4/89. (In Finnish)

NORMAALISTI KONSOLOIDOITUNEEN SAVEN POTENTIAALIYHTÄLÖ

K-H. KORHONEN
M. LOJANDER
Teknillinen korkeakoulu
Rakennus- ja ympäristötekniikan osasto
PL 2100 TKK

TIIVISTELMÄ

Paperissa on käsitelty normaalisti konsolidoituneen ja anisotrooppisen Otaniemen saven plastisen potentiaalin yhtälön määrittämistä varten tehtyjen kolmiakselikokeiden tuloksia. Näytteet on tutkittu Teknillisen korkeakoulun Pohjarakennuksen ja maamekaniikan laboratoriossa. Paperissa osoitetaan, että Otaniemen saven muodonmuutosyhtälöt sekä plastisen potentiaalin yhtälö voidaan kehittää populaatiomekaniikassa yleisesti sovelletun lineaarisen systeemin teorian perusteella. Maanäytteen ja –kerroksen rakeet ja huokoset tulkitaan toistensa kanssa kosketuksissa oleviksi populaatioiksi. Maa-alkion jäykkyydessä (leikkausvastuksessa) tapahtuvat muutokset edustavat populaatioiden geoteknisiä muutoksia. Populaatiomekaniikan aikatermi korvataan leikkausmuodonmuutoksella. Kuvasta 1 todetaan , että lineaarisen systeemin teorian perusteella kehitetty muodon-muutosyhtälö (6), $\nu=\nu(\varepsilon_s)$, approksimoi hyvin koetuloksia. Muodonmuutosyhtälöä (7) , $\eta=\eta(\varepsilon_s)$, kuvassa 2, ei sitävastoin voida pitää riittävän tarkkana. Otaniemen saven muodonmuutosominaisuuksien matemaattinen mallintaminen on tavanomaista hankalampaa. Kuormitetun savikerroksen plastista käyttäytymistä simuloivissa malleissa tarvittava potentiaaliyhtälö ja myötöfunktio voidaan kuitenkin , kuten lähes aina, kehittää lineaarisen systemin teorian yhtälöstä (4).

1. LINEAARINEN SYSTEEMI

Muodonmuutosyhtälö

Lineaarista systeemiä sovelletaan mm. populaatiomekaniikassa kahden tai useamman toisiinsa vaikuttavan populaation koossa tapahtuvien muutosten selittämiseen. Maamekaniikassa maakerroksen rakeet ja huokoset voidaan tulkita populaatioksi, joiden vaikutus toisiinsa määrää jännitysten ohella kerroksen mekaanisen käyttäytymisen. Kun lineaarisella systeemillä pyritään selvittämään maakerroksen käyttäytymistä on selitystä täydennettävä niin, että kontinuumimekaniikan "reunaehdot" täyttyvät. Tässä yhteydessä lineaarisella systeemillä yritetään määrittää maamekaniikan malleissa tarvittavia materiaaliyhtälöitä, jotka ovat usein empiirisiä. Esim. Hooken laki sekä Coulombin ja Darcyn lait ovat empiirisiä yhtälöitä, jotka edustavat malleissa ko. materiaalin ominaisuuksia. Lineaariseen systeemiin perustuvat yhtälöt sovitetaan kontinuumimekaniikan mukaiseen malliin erillisessä, myöhemmin julkaistavassa artikkelissa.

Lineaarinen systeemi on yksinkertaisin matemaattinen malli, jolla voidaan ottaa huomioon kahden populaation vaikutus toisiinsa. Systeemin lähtöyhtälöt ovat yleensä nopeusyhtälöitä (dN/dt). Maamekaniikan reologisissa teorioissa, joita on kehitetty mm. muodonmuutosten nopeuden ja rakenteiden painumisnopeuden laskemista varten, lähdetään myös nopeusyhtälöistä. Tässä yhteydessä pyritään kuitenkin selvittämään alkioksi tulkittuun maanäytteeseen kohdistuvien jännitysten ja niiden aiheuttamien muodonmuutosten välisiä relaatioita. Avoimessa kolmiakselikokeessa muodonmuutosten ja jännitysten muutosten nopeuksien vaikutukset, dε/dt ja dσ/dt, pyritään eliminoimaan siten, että muodonmuutokset tapahtuvat erittäin hitaasti ns. nollanopeudella eli $(d\varepsilon / dt) \rightarrow 0$. Populaatiomekaniikan nopeustermi (*dN/dt*) korvataan maa-alkion jäykkyydessä tapahtuvilla muutoksilla; $\frac{d\eta}{d\varepsilon_s} ja \frac{dv}{d\varepsilon_s}$. Maa-alkion tilavuudenmuutoksen inkrementin ja leikkausmuodonmuutoksen inkrementin suhde $\dot{\varepsilon}_v / \dot{\varepsilon}_s$ on ns. dilataatiokulma ja se edustaa tilavuuden muutoksen aiheuttamaa leikkausvastusta. Tekninen tilavuudenmuutos vaihdetaan maan mekaanisia malleja kehitettäessä ominaistilavuuden muutokseksi $-\dot{\varepsilon}_v = -\dot{v}/v$. Maa-alkion ja maakerroksen muodonmuutokset riippuvat leikkausjännityksen ja normaalijännityksen suhteesta τ/σ. Kun operoidaan kriittisen tilan mallin formalismilla, vaihdetaan τ/σ jännitysdeviaattorin $q = (\sigma'_1 - \sigma'_3)$ ja hydrostaattisen jännityksen $p' = 1/3(\sigma'_1 + \sigma_2 + \sigma_3)$ suhteeksi $\eta = q/p'$. Jännityssuhteen inkrementin $\dot{\eta}$ ja leikkausmuodonmuutoksen inkrementin suhde $\dot{\eta}/\dot{\varepsilon}_s$ voidaan rinnastaa Hooken lakia noudattavien materiaalien leikkausmoduuliin (leikkaus-vastukseen). Dilataatiokulma $\dot{\varepsilon}_v/\dot{\varepsilon}_s$ ja η ovat sellaisenaan vertailukelpoisia. Esim. kriittisen tilan mallin plastisen potentiaaliyhtälön muodostamiseen sovelletaan jännitysdilataatioyhtälöä (1). [1].

$$\frac{\dot{\varepsilon}_v}{\dot{\varepsilon}_s} = -\frac{\dot{v}}{v\dot{\varepsilon}_s} = M - \frac{q}{p} = M - \eta \qquad (1)$$

M on CSM:n lujuusparametri

Maa-alkion muodonmuutoskäyttäytymistä simuloivan lineaarisen systeemin lähtöyhtälössä (2) inkrementit on vaihdettu differentiaaleiksi. Derivaatat $dv/d\varepsilon_s$ ja $d\eta/d\varepsilon_s$ edustavat edellä tarkoitettuja jäykkyyksiä.

$$\frac{dv}{d\varepsilon_s} = -av + b\eta \qquad (2a)$$

$$\frac{d\eta}{d\varepsilon_s} = cv - d\eta \qquad (2b)$$

v on ominaistilavuus
a, b, c ja *d* ovat vakioita.

Yhtälön (2) vakioiden etumerkit voivat olla joko + tai -. Yhtälöllä (2) on kolme ratkaisua; $v = v(\eta)$, $v = v(\varepsilon_s)$ ja $\eta = \eta(\varepsilon_s)$, joita sanotaan muodonmuutosyhtälöiksi. Tässä yhteydessä tarkastellaan kuvien 1 ja 2 mukaisia relaatioita edustavia yhtälöitä $v = v(\varepsilon_s)$ ja $\eta = \eta(\varepsilon_s)$. Muodonmuutosyhtälöllä $v = v(\eta)$ on lähinnä teoreettinen merkitys "kontrolliyhtälönä". Koetuloksia approksimoivaa yhtälöä $v=v(\eta)$ voidaan soveltaa ko olosuhteisiin soveltuvan mallin tunnistamiseen. Alkuperäisessä kriittisen tilan mallissa $v=v(\eta)$ on lineaarinen [1].

Yhtälön (2) ratkaisu riippuu karakteristisen yhtälön (3) juurien laadusta, jotka voivat olla reaalilukuja, kompleksilukuja tai imaginaarisia lukuja. Potentiaaliyhtälöissä tulevat kysymykseen kaikki edellä luetellut tyypit. Muodonmuutosyhtälöissä ja reologisissa yhtälöissä karakteristisen yhtälön (3) juuret ovat joko negatiivisia tai positiivisia reaalilukuja.

$$\lambda^2 - (a+d)\lambda + (ad - bc) = 0 \tag{3}$$

Yhtälöä (3) muodostettaessa on otettu huomioon, että kaikki vakiot $a,\ b,\ c,\ d > 0$ ovat positiivisia. Karakteristiseen yhtälöön vakiot on sijoitettava etumerkkeineen.

2. PLASTISEN POTENTIAALIN YHTÄLÖ

Lineaarinen systeemi soveltuu useiden maatyyppien potentiaaliyhtälön muodostamiseen. Kriittisen tilan mallin formalismilla kirjoitettu lähtöyhtälö on tällöin (4). Derivaatta $dq/d\varepsilon_s$ voidaan rinnastaa Hooken lakia noudattavien materiaalien leikkausmoduuliin.

$$\frac{dq}{d\varepsilon_s} = Ap + Bq \tag{4a}$$

$$\frac{dp}{d\varepsilon_s} = Cp + Dq \tag{4b}$$

$A,\ B,\ C,\ D$ ovat maaparametreja, joko + tai -.
p ja q ovat jännityskomponentteja, kuva 4

Yhtälö (4) ja sen ratkaisut ovat yksikäsitteisesti voimassa määrätyllä jännityspolulla. Kaikki yhtälön (4) maaparametrit määritetään yhden avoimessa tilassa leikatun näytteen koetulosten perusteella. Potentiaaliyhtälö g saadaan yhtälöstä (4) kun siitä muodostetaan integroimalla yhtälö $q= q(p)$ ja edelleen yhtälö (5).

$$g = g(p, q, p_i) \tag{5}$$

p_i on potentiaaliyhtälön kuvaajan sijainnin p-q-tasolla määräävä muuttuja.

Yhtälö (4) voidaan esittää myös niin, että leikkausmuodonmuutos ε_s korvataan monotonisesti kasvavalla apumuuttujalla, yleensä aikatermillä t. Populaatiomekaniikan lineaaristen systeemien lähtöyhtälöt ovat nopeusyhtälöitä reaaliajassa.

3. OTANIEMEN SAVEN MUODONMUUTOSYHTÄLÖT

Kuvassa 1 on esitetty Otaniemen saven avoimen kolmiakselikokeen leikkausvaiheen aikana havaittu leikkausmuodonmuutoksen ja ominaistilavuuden relaatio , jota on approksimoitu yhtälöllä (6a). Yhtälö on muodostettu integroimalla ryhmä (2).Kuvissa 1, 2 ja 3 esitetyt koetulokset on saatu kokeesta, jossa jännityspolku on p-q -tasolla ollut: $\dot{q} / \dot{p} = 3$.

$$\nu = \nu_c + \frac{\Delta \nu_c}{(\lambda_2 - \lambda_1)}\left[(a + \lambda_2)e^{\lambda_1 \varepsilon_s} - (a + \lambda_1)e^{\lambda_2 \varepsilon_s}\right] \tag{6a}$$

v_c on ominaistilavuus kriittisessä tilassa

$\Delta \nu_c = \nu_o - \nu_c$, tilaparametri, kun $\dot{p} = 0$

λ_1 ja λ_2 karakteristisen yhtälön (3) juuria, Tässä tapauksessa $\lambda_1 < 0$ ja $\lambda_2 < 0$

a parametri yhtälössä (2a)

Kun yhtälön (6a) kuvaajaa sovitettiin kuvan 1 mukaisiin tuloksiin todettiin, että parametri b yhtälössä (2a) on b = 0. Yhtälö (2a) supistuu tällöin muotoon (6b), jolla on ratkaisu (6c).

$$\frac{d\nu}{d\varepsilon_s} = -a\nu \tag{6b}$$

$$\nu = \nu_c + \Delta \nu_c \cdot e^{-a\varepsilon_s} = 2{,}381 + 0{,}47 \cdot e^{-4{,}7\varepsilon_s} \tag{6c}$$

$$a = \lambda_2 = -4{,}7$$

Kuvasta 1 todetaan, että yhtälö (6c) approksimoi hyvin koetuloksia. Erityisesti on todettava, että yhtälö (6) approksimoi koetuloksia arvioidun ja laskennallisesti korjatun $\Delta \nu_c$:n alueella. Käyrän yläpäästä poistettiin ylikonsolidoitunut alue; $\Delta \nu \approx 0{,}04$. Kriittistä tilaa vastaava ominaistilavuus arvioitiin hyperbelin avulla.

Kuvaan 2 on piirretty leikkausmuodonmuutoksen ε_s ja jännityssuhteen η välinen relaatio koetulosten perusteella, joita approksimoitiin yhtälöryhmästä (2) integroimalla saadulla yhtälöllä (7a).

$$\eta = M - \frac{M}{(\lambda_2 - \lambda_1)}\left[(d + \lambda_2)e^{\lambda_1 \varepsilon_s} - (d + \lambda_1)e^{\lambda_2 \varepsilon_s}\right] \tag{7a}$$

$$\left(\frac{d\eta}{d\varepsilon_s}\right)_{\varepsilon_s = 0} = dM \tag{7b}$$

λ_1 ja λ_2 ovat myös yhtälön (6) parametreja

Parametria d yritettiin määrittää havaintojen ja yhtälön (7b) avulla. Vaikka kuvan 2 mittakaava suurennettiin, osoittautui $\varepsilon_s - \eta$-käyrän initiaalimoduulin (7b) määrittäminen epätarkaksi. Parametrin d numeerinen arvo laskettiin valitsemalla useita havaittuja arvopareja ja sijoittamalla ne yhtälöön (7a). Parametrien λ_1 ja λ_2 arvot oli aikaisemmin laskettu myös Dalmatovin [2] esittämällä menetelmällä. Parametrien arvot vaihtelivat seuraavasti: $a \approx \lambda_2 \approx 3{,}5...5{,}5$, $d \approx \lambda_1 \approx 11{,}5...13.0$.

4. OTANIEMEN SAVEN POTENTIAALIYHTÄLÖ

Otaniemen saven potentiaaliyhtälön funktionaalista muotoa pyrittiin selvittämään jännitysdilataatioyhtälön perusteella. Kuvaan 3 on piirretty dilataatiokulman $d\varepsilon_v / d\varepsilon_s$ ja jännityssuhteen η välinen relaatio, jota voidaan plastisella alueella, $\eta \approx 0.6...1.1$, approksimoida suoralla. Suoran kaltevuuskulma on hieman pienempi kuin lujuusparametri M = 1.1. Erityisesti on todettava, että dilataatiokulma $d\varepsilon_v / d\varepsilon_s \neq 0$ kun η = M. Kriittisen tilan määritelmä ei toteudu. Kuvasta 1 todetaan, että $\varepsilon_s - v$ relaatiolla ei ole vaakasuoraa tangenttia, joten näyte on murtunut huomattavasti ennen perinteellistä kriittistä tilaa. Murtokohta voidaan todeta kuvasta 2, jossa $d\eta / d\varepsilon_s = 0$, kun $\varepsilon_s \approx 28\%$. Vastaava η:n arvo on tulkittu lujuusparametriksi η = M = 1.10. Kuvaan 3 piirretyn suoran alueella ts. kun $\eta \approx 0.6...1.1$ jännitys-dilataatioyhtälö voidaan kirjoittaa kriittisen tilan mallin mukaiseen muotoon (1). Vastaava potentiaaliyhtälö on (8).

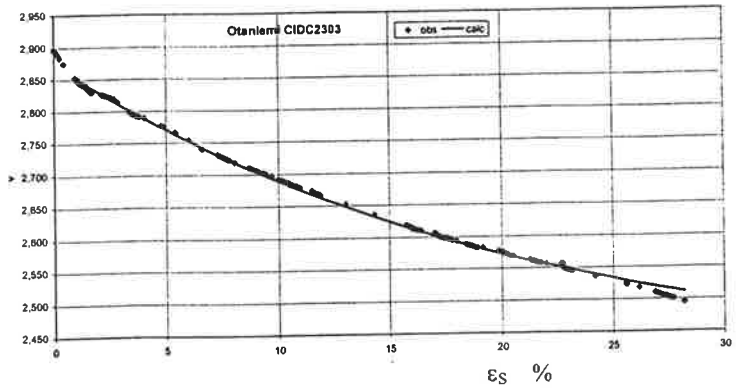
$$\frac{d\varepsilon_v}{d\varepsilon_s} = -\frac{dq}{dp} = M - \eta \qquad \text{(1 bis)}$$

$$g = \eta - M \ln \frac{p_i}{p} \qquad (8)$$

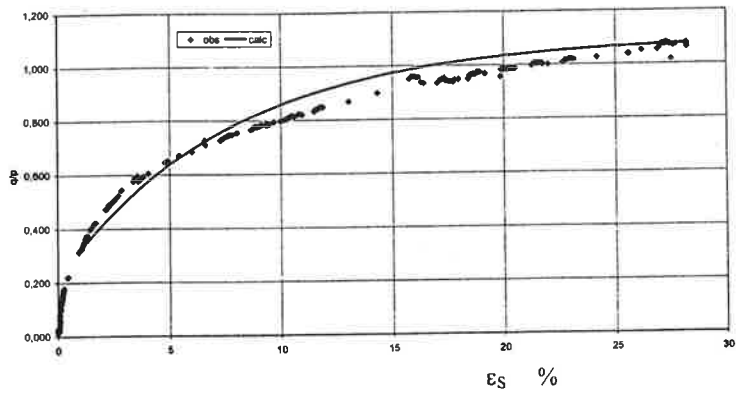
g on potentiaaliyhtälö
p_i parametri,joka määrää potentiaaliyhtälön kuvaajan hetkellisen paikan p-q -tasolla.

5. MYÖTÖPINTA

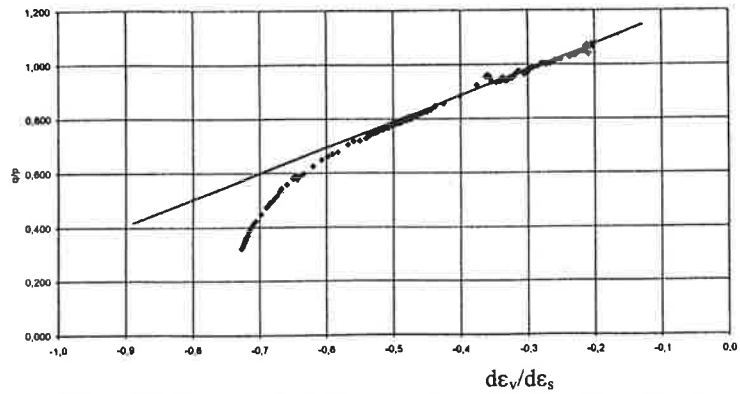
Vanhimmissa maan mekaanisissa malleissa myötöpinnan geometrinen muoto on yleensä otaksuttu. Myötöpintaa on approksimoitu tavallisesti ellipsillä tai spiraalilla. Erityisesti kanadalaiset tutkijat (Tavenas, Leroueil & al. 1987 [3]) ovat 1970-luvulta alkaen julkaisseet laajan kokoelman empiirisesti määritettyjä myötöpintoja. Useissa tapauksissa myötöpinnat ovat lähes ellipsejä. Aineistossa on myös tapauksia, joissa myötöpintaa ei voida approksimoida millään yksinkertaisella funktiolla (kartioleikkauksilla tai spiraaleilla). Suomessa tehtyjen kokeiden tulokset ovat samansuuntaisia [4]. Kuvassa 4 on Otaniemen anisotrooppisen saven empiirinen myötöpinta. Anisotrooppisuus on α=0,5[5]. Myötöfunktio voidaan muodostaa analogiaperiaatteella yhtälöstä (4).



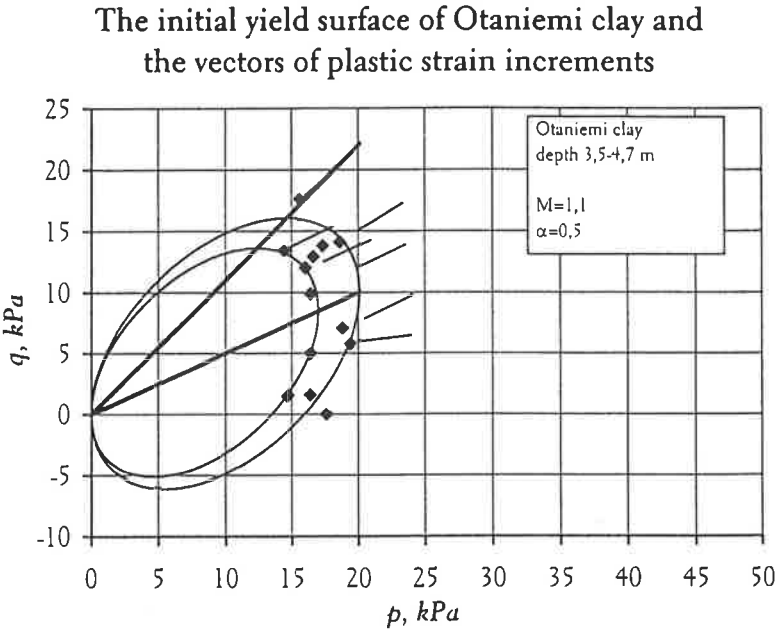
Kuva 1. Leikkausmuodomuutoksen ε_s ja ominaistilavuuden v välinen relaatio. Laskennallinen relaatio yhtälöllä (6), jossa b=0.



Kuva 2. Leikkausmuodonmuutoksen ε_s ja jännityssuhteen $\eta=q/p$ välinen relaatio. Laskennallinen relaatio yhtälöllä (7).



Kuva 3. Dilataatiokulman $d\varepsilon_v/d\varepsilon_s$ ja jännityssuhteen $\eta=q/p$ välinen relaatio, yhtälö (8)



Kuva 4. Otaniemen anisotrooppisen saven empiirinen myötöpinta, f [5].

6. OTANIEMEN SAVEN MEKAANINEN KÄYTTÄYTYMINEN

Kuvissa 1...4 esitetyt koetulokset muodostavat kokonaisuuden, jossa on useita merkittäviä poikkeamia saven "normaalista" käyttäytymistavasta, kun kriittisen tilan malli tulkitaan normaali-tapaukseksi. Kuvan 1 koetulokset voidaan selittää yhtälöllä (6c), jossa on yksi exp-termi, kuten Kelvinin reologisessa mallissa. Yhtälö approksimoi hyvin koetuloksia ε_s:n arvosta $\varepsilon_s = 0$ murtotilaan saakka. Kuvan 1 tuloksissa ei ole havaittavissa epäjatkuvuuskohtia eikä mitään sellaista, joka voitaisiin otaksua elastisen ja plastisen alueen rajaksi. Saman kolmiakselikokeen tulosten perusteella piirretyt kuvat 2 ja 3 vastaavat toisiansa ja niiden perusteella saven mekaaninen käyttäytyminen jakautuu sitä vastoin selvästi kahteen osaan. Kun ε_s vaihtelee välillä $\varepsilon_s = 0 \ldots \varepsilon_s \approx 6\%$ vaikuttaa käyttäytyminen elastoplastiselta. Kun η ylittää arvon $\eta \approx 0{,}6$ ja ε_s arvon $\varepsilon_s \approx 6\%$ muuttuu käyttäytyminen selvästi plastiseksi ja noudattaa kriittisen tilan mallin jännitys-dilataatioyhtälöä (1) ja potentiaaliyhtälöä (8). Saven empiirisesti määritetty myötöpinta kuvassa 4 on vino ellipsi, jonka anisotrooppisuus on $\alpha = 0.50$. Myötöpinta on tyypillinen anisotrooppiselle savelle. Saven mekaaninen käyttäytyminen on selvästi ei assosioitua, ts. $g \neq f$. Elastisen ja plastisen alueen lävitse ilman minkäänlaista epäjatkuvuuskohtaa ulottuva yhtälön (6c) kuvaaja kuvassa 1 jää tässä yhteydessä vaille kontinuumimekaniikan mukaista selitystä.

LÄHTEET

1. Schofield, A. N., Wroth, C. P. Critical State Soil Mechanics. (1968). McGraw-Hill, London.

2. Dalmatov, B., Bang, D. Savimaan deformoituvuusominaisuuksien määrittäminen ajan funktiona. Maamekaniikka, maapohjat ja perustukset. Leningrad 1968. Suomennos venäjänkielisestä paperista.

3. Tavenas, F., Lerouiel, S. State-of-the-art on laboratory and in Situ stress-strain-time behaviour of soft clays. Universite Laval. Sciences et Genie, Genie Civil, Août 1987.

4. Näätänen, A., Lojander, M. The selection of the parameters for the modified cam clay analysis. Proceedings of the third European Conference on Numerical Methods in Geotechnical Engineering – ECONMIG 94. Manchester/UK/7-9 September 1994. Numerical Methods in Geotechnical Engineering. Smith (ed.). Balkema, Rotterdam. pp.105-110.

5.Toivanen, T-L.Modelling of the aniosotropy of Otaniemi clay. Helsinki University of Technology. Soil Mechanics and Foundation Engineering. Master's Thesis 1999.

A simple estimate for sinking speed of a high level nuclear waste canister in a bentonite buffer

Antti Lempinen
Laboratory of Theoretical and Applied Mechanics
Helsinki University of Technology
P.O.Box 1100, FIN-02015 HUT

Abstract

The high level nuclear waste is planned to be disposed of in bedrock below water table. Thefore, it has to be isolated from ground water. The barrier material is planned to be bentonite, which is a swelling clay. In this paper a simple estimate for sinking speed for waste canister in compacted bentonite is given. The estimate is based on assumption that compressed bentonite behaves like very viscous liquid.

Notation

p	average total pressure
t	time
$\boldsymbol{\gamma}$	deviatoric strain tensor
$\boldsymbol{\epsilon}$	strain tensor
ε	relative dilation
$\boldsymbol{\xi}$	displacement vector
$\boldsymbol{\sigma}$	stress tensor
$\boldsymbol{\tau}$	deviatoric stress tensor

1 Introduction

High level nuclear waste is planned to be disposed of in bedrock surrounded with bentonite buffer, as sketched in Figure 1. The bentonite buffer should isolate

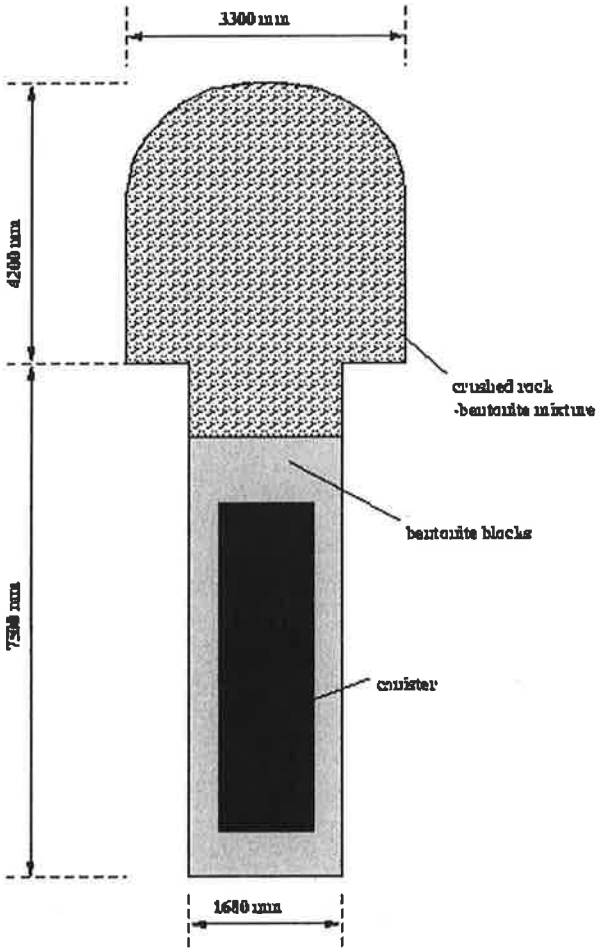


Figure 1: The dimensions of the bentonite buffer system [1].

the composite steel and copper waste canister from ground water. Of course, bentonite should be strong enough to carry the weight of the canister.

Whether a material is called a solid or a fluid is in some cases a matter of the time scale. For example, at room temperatures glass can be regarded as a solid, when observation time is hours. However, if observation time is centuries glass is a supercooled liquid.

The observation time of laboratory experiments on bentonite is months or shorter. High level nuclear waste should be isolated for a million years. The difference between these time scales constitutes clearly a problem, since bentonite being a solid within the former time scale does not imply it being a solid within the latter one.

A simple estimate of upper bound for gravitational sinking rate of a waste canister is presented here. It is based on assumption that viscous shear strain rate is too small to be observed in short creep tests.

2 Viscosity of compacted bentonite

The notation here is adopted from [2]. Additionally, deviatoric strain is defined by

$$\boldsymbol{\gamma} = \boldsymbol{\epsilon} - \frac{1}{3}\varepsilon \mathbf{1}$$

and deviatoric stress is defined by

$$\boldsymbol{\tau} = \boldsymbol{\sigma} + p\mathbf{1}$$

In [3], a series of creep tests with compressed MX-80 bentonite is presented. The measurements show that in time scales shorter than 10 days, the deviatoric strain rate $\dot{\gamma}_{zz}$ follows the formula

$$\dot{\gamma}_{zz} = \dot{\gamma}_0 \exp\left(\alpha \frac{\tau_{zz} - \tau_0}{\tau_\mathrm{f}}\right) \left(\frac{t}{t_\mathrm{r}}\right)^{-n} \tag{1}$$

where $\dot{\gamma}_0$ is the reference rate at time $t = t_\mathrm{r}$ with deviatoric stress $\tau_{zz} = \tau_0$, and τ_f is the deviatoric stress at failure. Formula (1) is purely empirical, and it cannot be valid with small stresses, since $\tau_{zz} = 0$ gives $\dot{\gamma}_{zz} \neq 0$. The parameter values evaluated from undrained tri-axial tests with $\tau_\mathrm{f} = 1.5$ MPa, $\tau_{zz}/\tau_\mathrm{f} = 0.5$ and $t_\mathrm{r} = 10^5$ s are

$$\begin{aligned} \dot{\gamma}_0 &= 4.4 \times 10^{-8}\,\mathrm{s}^{-1} \\ n &= 0.91 \\ \alpha &= 4.15 \end{aligned} \tag{2}$$

Integration of formula (1) gives

$$\gamma_{zz} = \gamma_0 + \frac{\dot{\gamma}_0 t_r}{(1-n)} \exp\left(\alpha \frac{\tau_{zz} - \tau_0}{\tau_f}\right) \left(\frac{t}{t_r}\right)^{(1-n)} \tag{3}$$

where γ_0 is the instantaneous response to shear stress, i.e. elastic shear strain. Theoretically, since $\dot{\gamma}_0 \to 0$ when $t \to \infty$ this kind of material is not a fluid, but since $\gamma \to \infty$ when $t \to \infty$ it is not a solid either. After one million years, formula (3) with parameters (2) and $\tau_{zz} = \tau_0$ gives $\gamma = \gamma_0 + 0.04$, which means that material is practically a solid.

However, extrapolation of (1) to long-time behaviour is not automatically justified. Here, it is simply assumed that compacted bentonite acts as a very viscous incompressible liquid, that is

$$\boldsymbol{\tau} = 2\eta\dot{\boldsymbol{\gamma}} \tag{4}$$

$$\dot{\varepsilon} = 0 \tag{5}$$

where η is dynamic viscosity. Formulas (1) and (4) can be combined, if the displacement vector is decomposed in short-time response $\boldsymbol{\xi}^*$ and long-time response $\boldsymbol{\xi}^V$

$$\boldsymbol{\xi} = \boldsymbol{\xi}^* + \boldsymbol{\xi}^V \tag{6}$$

Component $\boldsymbol{\xi}^*$ behaves according to (1) and $\boldsymbol{\xi}^V$ behaves according to (4) -(5) with very big viscosity

$$\eta > \frac{\tau_{zz}}{2\dot{\gamma}_0} \exp\left(-\alpha \frac{\tau_{zz} - \tau_0}{\tau_f}\right) \left(\frac{T}{t_r}\right)^n \tag{7}$$

where T is the time length of the creep test. In that case, formula (1) is valid with $t < T$ and formula (4) describes the behaviour when $t \gg T$.

Substitution of parameter values (2) and the time length of the measurements $T = 10^6$ s into relation (7) gives a lower bound for the viscosity of compressed MX-80 bentonite:

$$\eta > 2 \times 10^{14}\,\text{Pa s} \tag{8}$$

In Figure 2 it is shown, how the result of a creep test would look like, if

$$\dot{\gamma}_{zz} = \dot{\gamma}_0 \exp\left(\alpha \frac{\tau_{zz} - \tau_0}{\tau_f}\right) \left(\frac{t}{t_r}\right)^{-n} + \frac{1}{2\eta}\tau_{zz}$$

with $\eta = 4 \times 10^{14}$ Pa s and parameter values (2). It can be seen, that deviation from straight line would not be large.

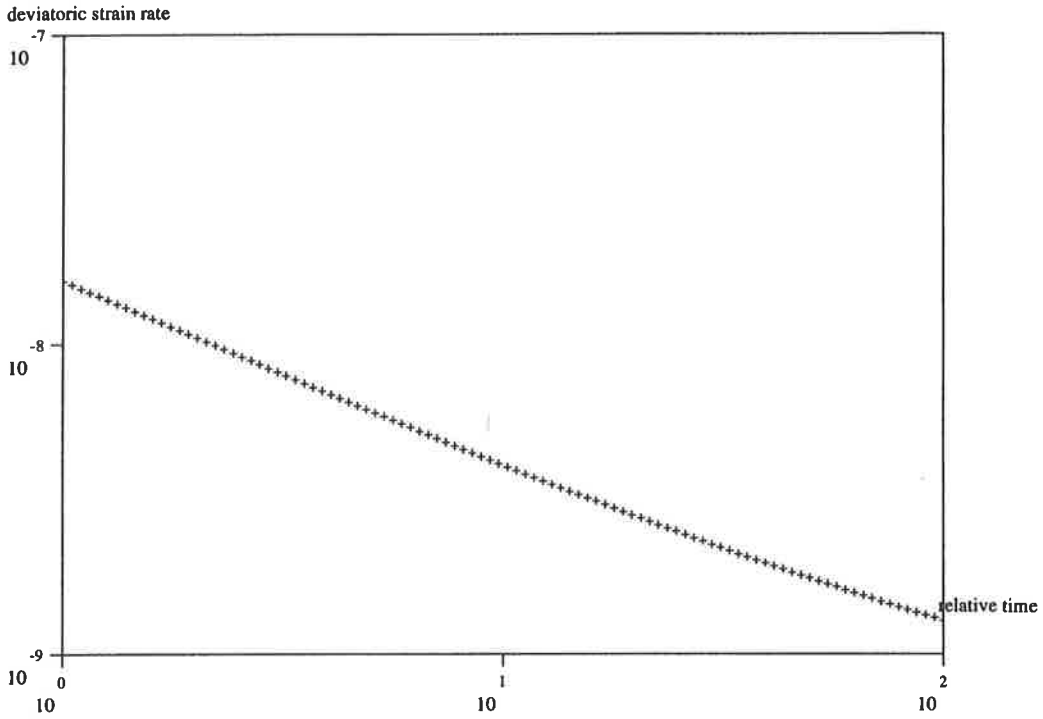


Figure 2: Theoretical result of a creep test, $\tau_{zz}/\tau_{\mathrm{f}} = 0.25$

3 Estimate of sinking rate

The well known Stokes formula gives the drag force F of a sphere of radius R moving with slow velocity u in fluid with viscosity η:

$$F = 6\pi\eta R u \tag{9}$$

Volume of a waste canister is $3.36\mathrm{m}^3$ and its mass is 20836 kg. When density of bentonite is 2000 kg m^{-3}, the net force acting on a canister, that is weight minus buoyancy force, is 138 kN. Using (9) with $\eta = 2 \times 10^{14}\,\mathrm{Pa\,s}$ and the radius of the canister $R = 0.491$ m gives a rough estimate of the sinking rate v of a waste canister in bentonite:

$$v < 7.5 \times 10^{-11}\,\mathrm{m\,s^{-1}} = 2.3\,\mathrm{mm\,a^{-1}}$$

Finite element calculation of the sinking rate with no friction on bentonite boundaries gives

$$\eta v = 3200\,\mathrm{N\,m^{-1}}$$

so that

$$v < 1.6 \times 10^{-11}\,\mathrm{m\,s^{-1}} = 0.5\,\mathrm{mm\,a^{-1}}$$

This means, that according to these simple estimates, it takes at least few thousand years for canisters to penetrate the bentonite buffer.

4 Conclusions

If the time scale of few thousand years is consider to be too short for the buffer performance, more experiments is needed to check the lower bound of the viscosity. However, it is difficult to design such experiments, since the observation time cannot be prolonged very much, and there seem to be no accelerated experiment methods which retain the essential properties of the material. Therefore, also calculations based on microscopic phenomena are necessary.

The estimates above are simple calculations of movement caused by gravitation only. Movements caused by other phenomena like swelling have to be considered more closely, and non-symmetric movements have to be taken into account.

References

[1] Posiva Oy. Final disposal of spent nuclear fuel in the finnish bedrock, technical research and developement in the period 1993-1996. Technical Report POSIVA-96-14, Posiva Oy, Helsinki, 1996. In Finnish.

[2] Antti Lempinen. Thermomechanical model for expansive clays: I. reversible processes. Technical Report 51, Helsinki University of Technology Department of Engineering Physics and Mathematics Laboratory of Theoretical and Applied Mechanics, Espoo, 1998.

[3] Lennart Börgesson. Modelling of buffer material behaviour. some examples of material models and performance calculations. Technical Report 88-29, Swedish Nuclear Fuel and Waste Management Co, Stockholm, 1988.

MODELLING OF ANISOTROPY OF FINNISH CLAYS

A. NÄÄTÄNEN, M. LOJANDER
Laboratory of Soil Mechanics and Foundation Engineering
Helsinki University of Technology
P.O. Box 2100
FIN-02015 HUT, FINLAND

ABSTRACT

Suitability of a rotational hardening elasto-plastic soil model for anisotropic soft soils called SCLAY-1 is tested for four Finnish clays. First part of the paper demonstrates that the initial anisotropic yield curve can be established easily from the knowledge of K_0-value of the soil (earth pressure coefficient at rest). The second part shows some test simulations. Triaxial tests of POKO-clay and Paimio clay have been simulated with SCLAY-1 and Modified Cam Clay (MCC) model and results will be compared with laboratory data.

1. INTRODUCTION

The anisotropy of soft soils has been in an interest of the Laboratory of Soil Mechanics and Foundation Engineering at HUT since 1980's [1-4]. Korhonen [1-2] presented a model called MAC (Model for Anisotropic Clays) with a yield curve formulation that took into account the initial anisotropy of clay. MAC, however, did not account for changes of yield curve orientation due to plastic straining. Anisotropy became again an active study at HUT after Wheeler from University of Glasgow (GU) published his formulation 1997 [5]. The Wheeler model included the same yield curve formulation as MAC combined with a rotational hardening law to take into account the anisotropy caused by plastic straining. HUT and GU decided to combine their efforts towards developing a realistic and still simple to use model for taking into account the influence of plastic anisotropy as a part of the project called 'Modelling of mechanical behaviour of anisotropic soft soils'. As an essential part of the model development is detailed comparison with high quality laboratory data, a laboratory-testing programme was designed to investigate the validity of the proposed model [6-7]. The model was tested with a large number of triaxial tests with Otaniemi clay, and as a result, the original model proposed by Wheeler [5] was slightly modified. New version of the model was named SCLAY-1 [8].
In this paper the applicability of SCLAY-1 is tested for some other Finnish clays. Results from triaxial consolidation tests of POKO-clay, specifically made for testing the model, are presented. In addition, a number of other test results on Finnish clays have been examined.

2. MODEL FORMULATION

The main equations describing the model formulation are given below and more detailed descriptions can be found in [6, 7]. SCLAY-1 describes the behaviour of normally consolidated or lightly overconsolidated soft soils, where plastic deformations dominate. It

takes into account both the initial anisotropy of the soil as well as the changes of the anisotropy following the plastic straining due to different loading conditions. Although the model has been generalised [8], equations are presented here in triaxial stress space for the sake of simplicity.

Elastic behaviour is assumed to be isotropioc similar to the Modified Cam Clay (MCC) model [9], while the yield function is a sheared ellipse:

$$f = (q - \alpha p')^2 - (M^2 - \alpha^2)(p'_m - p')p' = 0 \qquad (1)$$

where p' and q are the mean effective stress and deviatoric stress, respectively, M is the critical state value of stress ratio η (where $\eta=q/p'$) and the parameters p'_m and α define the size and the inclination of the yield curve respectively (Fig. 1). For the case $\alpha=0$ (isotropy), the yield curve reduces to the same form as in the Modified Cam Clay model. Associated flow rule is assumed.

The model incorporates two hardening laws. The first one describes changes in size of the yield curve and it is identical to that of MCC (Eq.2). The second hardening rule describes the changes of orientation of the yield curve produced by plastic straining (Eq. 3). This form was developed based on a large testing programme with Otaniemi clay [6, 7].

$$dp'_m = \frac{vp'_m \, d\varepsilon_v^p}{\lambda - \kappa} \qquad (2)$$

$$d\alpha = \mu \left[\left(\frac{3\eta}{4} - \alpha \right) < d\varepsilon_v^p > + \beta \left(\frac{\eta}{3} - \alpha \right) | d\varepsilon_d^p | \right] \qquad (3)$$

where plastic volumetric strains are currently attempting to drag α towards a target value of $3\eta/4$ and plastic shear strains are attempting to drag α towards a target value of $\eta/3$. μ and β are new soil constants. Constant μ controls the absolute rate at which α heads towards its current target value: the higher is μ, the faster the yield curve rotates. Constant β controls the relative effectiveness of plastic shear strains and plastic volumetric strains in the yield curve rotiaton.

3. ANISOTROPIC BEHAVIOUR OF CLAYS

3.1 Properties of studied clays

Four different clays were studied in order to see whether the proposed model is valid for other Finnish clays than Otaniemi clay. Only the POKO-clay has been tested as a part of this project. The test data for other clays have been collected from the various projects done at the Laboratory of Soil Mechanics and Foundation Engineering [4, 10-15]. As these tests were performed with different modelling purposes not all the aspects of the new model could be tested. The index properties of the clays studied are presented in Table 1.

3.2 Initial yield curve

If it can be assumed that the initial anisotropy of the normally or lightly overconsolidated soil deposit is connected to geological formation of the soil layer under K_0-consolidation a simple procedure can be used for determination of α. The method of calculating α based on an estimate of K_0 value have been used successfully for several clays namely Otaniemi

clay, Winnipeg clay, Mexico city clay, Bothkennar clay and Marjamäki clay [8]. The determination of α can be based on measured value of K_0 if suitable laboratory measurements are available. Alternatively it can be estimated via the simplified Jaky's formula. The procedure is following:

1. Determine the critical state value M from triaxial test data
2. Calculate the corresponding value of friction angle from $\sin\phi' = 3M/(6+M)$
3. Calculate K_0 value using simplified Jaky's formula: $K_0 = 1-\sin\phi'$
4. Calculate stress ratio η_{K0} at K_0-line: $\eta_{K0} = 3(1-K_0)/(1+2K_0)$
5. Calculate corresponding value of α_0: $\alpha_0 = (\eta_{K0}^2 + 3\,\eta_{K0} - M^2)/3$
6. Choose the p_m' value so that the yield curve fits the data best

Initial yield curves for the four clays studied have been established in this way (Fig. 1). The yield points have been determined from triaxial consolidation tests mainly from bilinear lnp'-v plots. The K_0-value was determined with simplified Jaky's formula following the precedure above. This method gave an α-value of 0.46 for Paimio clay while earlier a value of 0.45 had been proposed just by visual estimation [4]. For Vaasa clay this methodology yields to a α-value of 0.58 instead of previously introduced 0.57 [15]. The yield points of POKO-clay have been normalised with the vertical in situ stress, as the range of depth for the samples tested was quite high. The fitted yield curves (Fig. 1) match the yield points very well.

Table 1. Soil properties of the studied clays

Index property	Symbol	Paimio clay [11]	Haarajoki clay [13]	Vaasa clay [14]	POKO clay
Depth [m]	Z	6.2 - 6.7	5 - 7	4.8 - 5.3	9.5 - 11
Water content [%]	W	76	88 - 110	95 - 102	83 - 92
Liquid limit [%]	w_L	61	89 - 98	96 - 109	80 - 89
Plastic limit [%]	w_P	24	29 - 30	31- 33	30 - 31
Plasticity index	I_P	37	60 - 68	65 - 76	50 - 58
% of particles < 0.002 mm	Cl-%	80	66 - 73	30	62 - 70
Organic content [%]	Hm	0.6	1.8-2.2	5	0.9-1.2
Specific gravity	Gs	2.78	2.77	2.63	2.78
Undrained shear strength [kPa]	c_u	18	18 - 29	15	19 - 24
Sensitivity	S_t	50 - 94	30 - 46		7.5 - 13.5

4. TEST SIMULATIONS

4.1 Modelling the behaviour of POKO-clay

Series of triaxial consolidation tests were done to investigate the validity of SCLAY-1. The testing programme for POKO-clay followed the same ideas as used earlier for Otaniemi clay [6]. Each sample was first loaded along constant q/p' stress path. From this loading stage the initial yield point was verified (Fig. 1d) and the loading was continued to a new p' value equal to 1.5 to 2.8 times the initial yield value of p'. This was followed by unloading along same stress path and then the sample was reloaded along some different constant q/p' stress path to find the new yield value (Table 2).

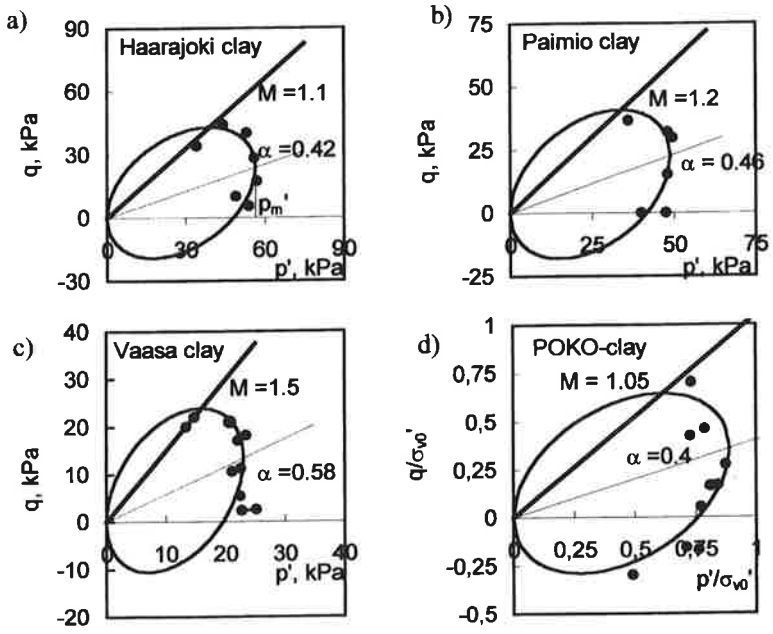


Figure 1. Initial yiels curves. a) Haarajoki clay b) Paimio clay [11] c) Vaasa clay d) POKO-clay.

Triaxial test results and model simulations with SCLAY-1 and Modified Cam Clay (MCC) models are presented here for three of these tests. The model constants for POKO-clay are following: λ=0.61, κ=0.03, M=1.05, ν'=0.2, α=0.4 and β=0.57 (the latter was estimated via K_0 as described in [7]). For p_m' and v the individual values corresponding to sample are used. As there is not yet a direct way for determining the parameter μ for given soil, three different values of μ have been used (10, 20 and 50). The stress path inclinations, maximum loads and the size of the initial yield curve for three simulated tests are presented in Table 2. For all simulations the initial yield curves for SCLAY-1 and MCC were fitted to coincide at K_0-line. This means that the normally consolidated value of in situ stress defines the size of the initial yield curve in both cases, which is a logical assumption for practical design problems.

Table 2. Triaxial tests of POKO-clay. Stress paths and yield points.

	CAE2737	CAE2770	CAD2751
p_m', kPa	45.6	53.8	53.8
Loading stage 1 q/p'	0.59	-0.22	0.91
Initial yield p_{y1}, kPa	42.1 – 45.6	49.3	38.5
Initial yield q_{y1}, kPa	24.6 – 26.7	10.9	37
p_{max1}, kPa	62.2	139.1	93.0
Loading stage 2 q/p'	-0.6	0.55	0.05
Second yield p_{y2}, kPa	35.9	96.4 - 99.7	95.3
Second yield q_{y2}, kPa	-21.2	53.0 – 54.8	5.1
p_{max2}, kPa	80.2	244.3	306.1

4.1.1 Rotation of the yield curve and direction of the plastic strain increments

Figure 2 presents the rotation of the SCLAY-1 yield curves during the tests. The yield curve marked with number 1 presents the initial yield curve and numbers 2 and 3 the yield curves after first loading and reloading, respectively. The orientation of the yield curves after first loading are based on the maximum load value and the second yield point determined after reloading stage (Table 2). Fitting the yield curve equation through two points gives the explicit value of α. The largest yield curves marked with 3, and a dashed line, have been drawn based on the target values for the final stress paths followed. These target values of α are not necessarily yet reached, and the third yield curves only indicate the orientation towards which the yield curves are rotating. Figure 2 shows also the directions of plastic strain increments at yielding for both stress paths. In case of associative flow rule the plastic strain increment vectors should be normal to a yield curve. The plastic strain increment vectors have been determined from the load step following the yield according to the ideas presented by Graham et al. [16]. As the loading was stepwise, there may be some gaps, so that the orientation may not always correspond to the situation immediately at yield.

In test CAE2737 the first loading follows closely to K_0-loading (η_{K0} =0.64) and hence no significant rotation occurs (Fig. 2a). The SCLAY-1's target value for α at η=0.59 loading is 0.39. Fitting the yield curve through maximum loading point and yield point gives a slightly smaller α value of 0.33. On reloading the yield curve starts to rotate anticlockwise due to plastic straining towards a new target value of α= –0.49. The plastic strain increment vector for first loading clearly supports the normality. This is the stage of loading, which did not involve much rotation. The plastic strain increment vector after the second yield is not normal to yield curve. However, if we take into account the fact that as soon as the material yields the plastic straining starts to drag the yield curve towards a new orientation, in this case anticlockwise to more isotropic orientation, we should in reality draw the vector against a yield curve with this changed orientation.

Figure 2b shows again the loading paths and the yield curves for test CAE2770. First stage involved loading with η=-0.22 and the observed rotation for the yield curve could be defined with α value of –0.12...-0.15 while the target value is –0.17. Two values of α is shown as it was not possible to determine the yield point unambiguously and a range of possible minimum and maximum values is given for yield stress (Table 2). After unloading, the reloading was done at η=0.55, which gives the target value of α=0.37. In this case the plastic strain increment vector after initial yield in not normal to the yield curve, but as previously, the yield curve is already rotating towards it's new orientation with α=-0.12...-0.15. Comparing the direction of plastic strain increment vector to this rotated yield curve would suggest normality. The same applies for the second yield.

Test CAD2751 started with loading at η=0.91. Based on the yield point determined from the second loading stage, the new yield curve orientation after the first loading stage can be defined with α value of 0.46, which is almost exactly the same as SCLAY-1 target value α= 0.47. For the reloading the target value of α for η=0.05 loading is 0.04. Rotated yield curves for test CAD2751 are presented in Figure 2c. The direction of the initial plastic strain increment does not support normality but, as the stress path is close to critical state line, the plot is very sensitive to the value of M. There might be some natural variation in the soil and just a slightly bigger value of M would have given an improved

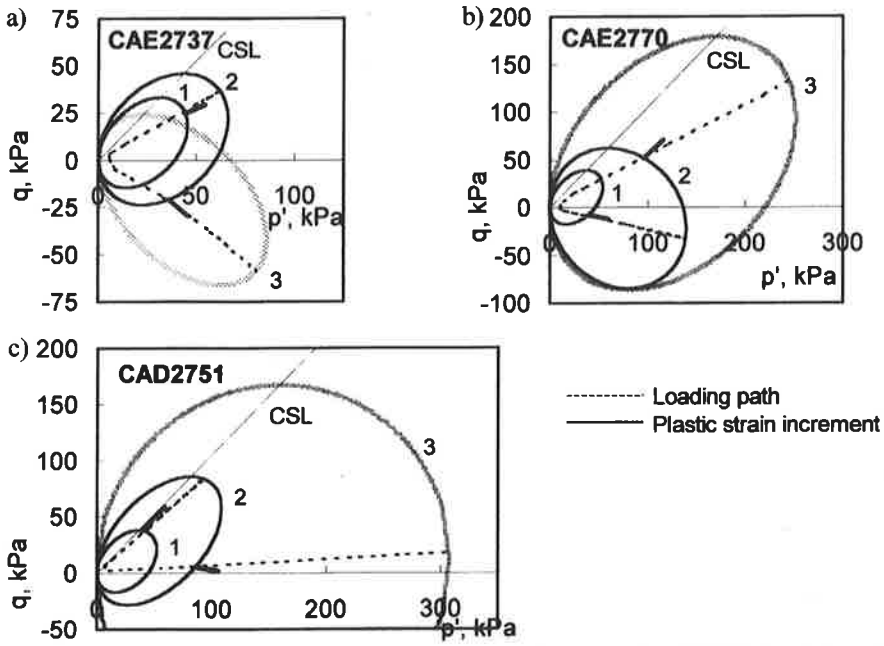


Figure 2. Loading paths and initial and rotated yield curves for POKO-clay triaxial tests.

agreement with normality. After second yield on the reloading path, the increment vector direction would suggest more isotropic yield curve based on obeying normality, but again the yield curve is rotating towards isotropy.

4.1.2 Stress-strain behaviour

Figure 3 presents the measured and calculated stress-strain behaviour for all three tests. Predictions of CAE2737 (Fig. 3a) show that both models predict the observed yield point very well as could be expected for loading close K_0-line. Significant differences occur in predictions for the yield point in reloading. SCLAY-1 is more realistic than MCC, which predicts a far too high yield stress. Also the pattern of straining (Fig. 3b) is well predicted with SCLAY-1, while MCC overestimates deviatoric strains. Figure 3b would indicate that a suitable value for parameter μ could be between 20 and 50.

Figures 3c and d show the measured and predicted stress-strain behaviour for test CAE2770. In this case SCLAY-1 predicts the initial and the second yield point very well while MCC model clearly overpredicts them both. The ε_v-ε_d plot in Figure 3d shows that SCLAY-1 predictions with μ values 20 or 50 match the observed behaviour nicely. With MCC simulation the miss-match is considerable.

CAD2751 involved initial loading close to critical state. Figures 3e and f show the measured and predicted behaviour. The first yield point is equally well predicted with both models. SCLAY-1 gives better estimate for the second yield than MCC. Both models are unable to predict the magnitude of strains. The volumetric strains observed are much lager than predicted and the calculated shear strains are unrealistically high. The reason for this will be discussed later.

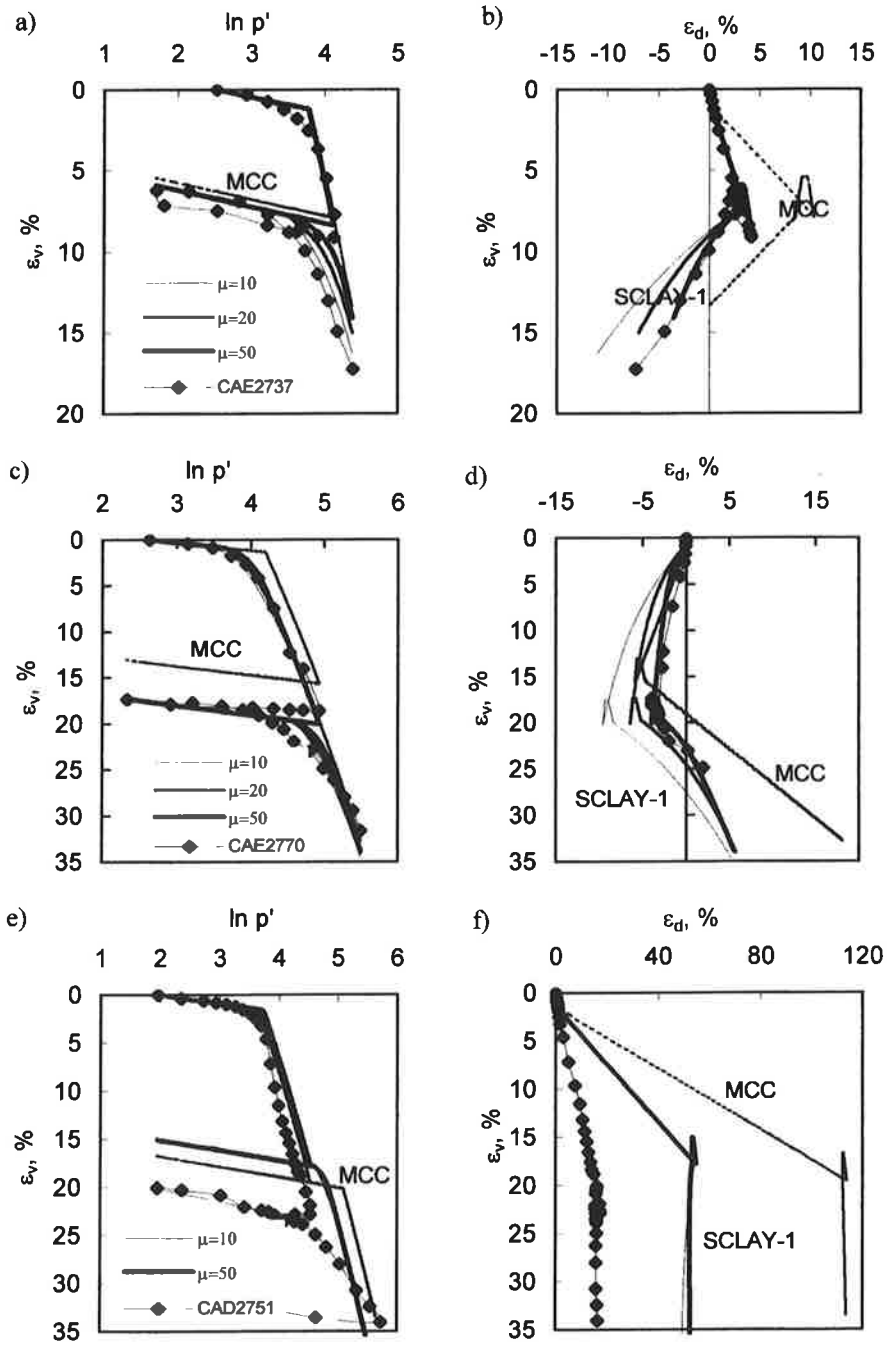


Figure 3. Simulations of POKO-clay. a) Test CAE2737. b) Test CAE2737. c) Test CAE2770. d) CAE2770. e) CAD2751. f) CAD2751.

4.2 Modelling the behaviour of Paimio clay

Triaxial tests of Paimio clay were done for a different project [4,11,14]. They, however, offer a possibility to test the model for a very sensitive clay (see Table 1) and one of those tests is simulated here. The model constants for Paimio-clay are following: λ=0.9, κ=0.05, M=1.2, ν'=0.1, α=0.45 and β=0.75. For p_m' and v the values of 54.6 kPa and 3.42 were used. Two values of the parameters μ have been used for simulations (20 and 50).

As there was a series of three tests, which all had been loaded to almost same maximum isotropic load, the results from all these tests were used to determine to new rotated yield curve (Fig. 4). Figure 4 presents the initial SCLAY-1 and MCC yield curves, which were again fitted to coincide at K_0-line. The test simulated was first loaded almost isotropically (η=0.03) to maximum point shown with black square and after unloading reloaded again along η=0.5 stress path. The yield points determined from three tests after the first loading are also shown. Figure 4b shows the development of α as the simulation proceeds. After the first yield α starts to reduce towards a target value of 0.02. During the unloading there is no change. On the reloading stage, after the second yield, α starts to increase towards new target value of 0.33. The test results suggest that the yield curve is rotated to a nearly isotropic orientation during the first loading, which would suggest a μ value of 50 or more. SCLAY-1 predicts the first yield point better than MCC, but gives the same value for the second yield. Because of too high yield stress at first loading, MCC predicts far too small volumetric strains. It could be expected that measured strains after first loading are bigger than simulated as the test was designed for studying creep. From the test results it is not easy to determine the second yield point, and in Figure 4a a range for yield stress is shown.

Figure 4d shows the predicted and measured values of volumetric and deviatoric strains. During the test itself no negative shear strains were measured, and in that respect the SCLAY-1 prediction does not follow the observed pattern of straining during the first loading. On reloading stage, however, the behaviour is predicted better, and clearly closer to measured values than the MCC prediction.

5. CONCLUSIONS

The way of determining α value on the basis of K_0 seems to be valid for several clays. This can significantly reduce the number laboratory tests needed for determining parameters for practical boundary value problems.

The directions of the plastiç strain increments seem to conform the assumtion of normality. The observed plastic strain increment directions of POKO-clay show the same kind of behaviour than Otaniemi clay [17].

On the light of Otaniemi clay a default value of μ=20 was suggested. While it might be acceptable for POKO-clay, for a highly sensitive Paimio clay higher value could be better. More data is needed for given any recommendations for possible default values of μ.

SCLAY-1 model gives much more realistic estimate of the soil behaviour than MCC. Only the behaviour of test involving first loading close to critical state could not be simulated accurately. Unexpectedly large volumetric strains occurred during the first loading at the test. Same kind of behaviour was observed with Otaniemi clay [7] and this could be attributed to the destructuration of the sensitive clay under certain loading paths.

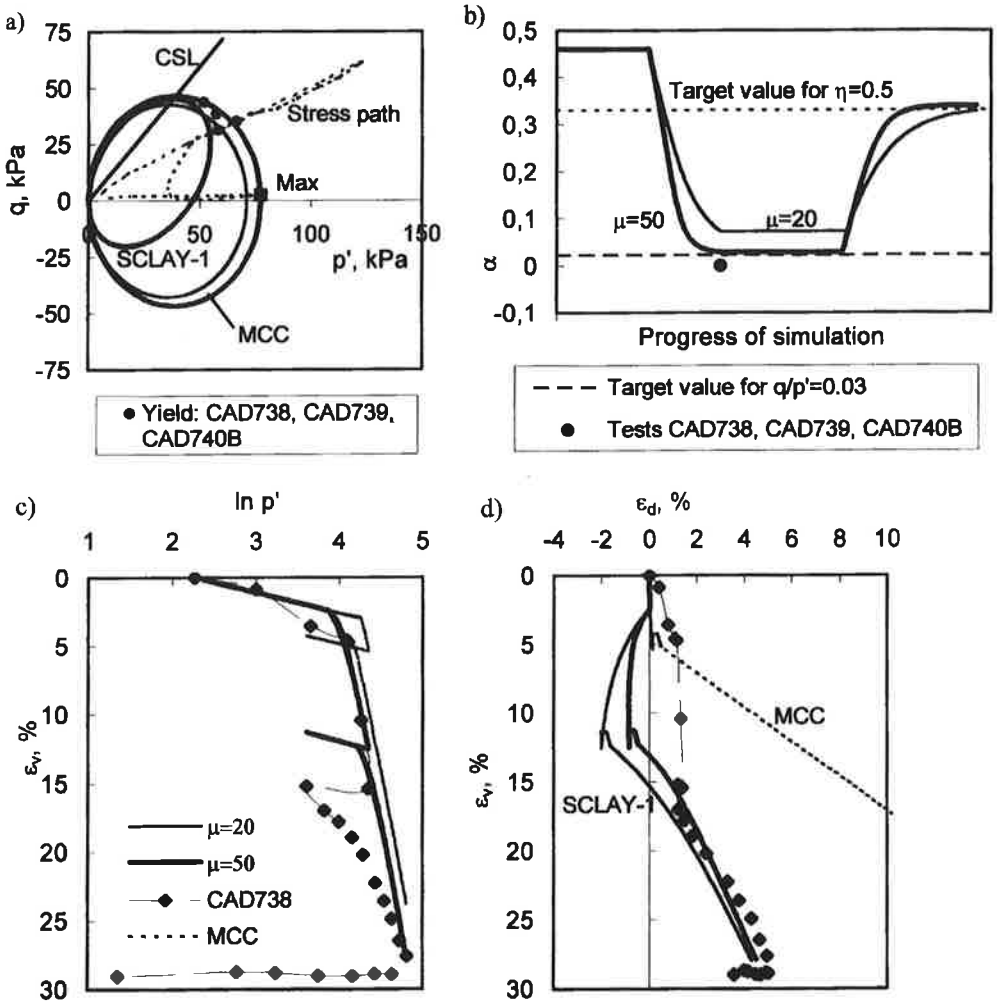


Figure 4. Simulation of Paimio clay.

ACKNOWLEDGEMENTS

This research was supported by the Finnish Academy and a Marie Curie Research Training Grant from the EU.

REFERENCES

1. Korhonen, K-H. & Lojander, M. 1987. Yielding of Perno clay. *Proc. 2nd Int. Conf. on Constitutive Laws for Engineering Materials*, Tucson, Arizona 1987. Vol. 2: 1249-1255. Elsevier.
2. Korhonen, K-H., Lojander, M., 1988.Mechanical Model for Anisotropic Clay. *Nordiska Geoteknikermöte,* Oslo, 1988.

3. Lojander, M. 1988. The parameters of the Mechanical Model of Anisotropic Clay. *Nordiska Geoteknikermöte,* Oslo, 1988: 154-157.
4. Korhonen, K-H., Karstunen, M., Lojander, M., 1991.The Yielding of anisotropic cohesive soils. X ECSMFE, Florence 26-30 May 1991. A.A. Balkema. Rotterdam. Pp. 237-241.
5. Wheeler, S. 1997. A rotational hardening elasto-plastic model for clays. *Proc. of the XIV ICSMFE,* Hamburg, Vol. 1: 431-434. Rotterdam: A. A. Balkema
6. Näätänen, A., Wheeler, S., Karstunen, M., Lojander, M. 1999.Experimental investigation of an anisotropic hardening model for soft clays. *Proc. 2nd Int. Symp. on Pre-failure Deformation Characteristics of Geomaterials*, Torino, Italy, 1999. Rotterdam: A.A.Balkema. Pp. 33-40.
7. Wheeler, S., Näätänen, A., Karstunen, M. 1999. Anisotropic hardening model for normally consolidated soft clays. *Proc. NUMOG VII,* Graz, Austria.
8. Wheeler, S., Näätänen, A., Karstunen, M., Lojander, M. Anisotropic volumetric and rotational hardening model for soft clays. Draft 15.12.1999 written to be submitted for Géotechnique.
9. Roscoe, K.H. & Burland, J.B. 1968. On the generalised stress-strain behaviour of 'wet' clay. *Engineering Plasticity.* 553-609. Cambridge University Press.
10. Aalto, A., Rekonen, R., Lojander, M. 1998. The calculations on Haarajoki test embankment with the finite element program Plaxis 6.31. *Proc. 4th European Conf. on Numerical Methods in Geotechnical Engineering*, Udine, Italy: 151-160. New York; Springer-Verlag. Pp.37-46.
11. Karstunen, M. 1991. The visco-plastic model of Paimio clay. MSc thesis. Helsinki University of Technology. (In Finnish).
12. Näätänen, A., Vepsäläinen, P., Lojander, M. 1998. Finite element calculations on Haarajoki test embankment. *Proc. 4th European Conf. on Numerical Methods in Geotechnical Engineering*, Udine, Italy: 151-160. New York; Springer-Verlag. Pp.151-160.
13. Saarelainen, K. 1998. Determination and verification of the time-settlement calculation parameters. MSc thesis. Helsinki University of Technology. (In Finnish).
14. Vepsäläinen, p., Arkima, O., Näätänen, A., Lojander, M., 1991. The Trial Embankments in Vaasa and Paimio, Finland. X ECSMFE, Florence 26-30 May 1991. A.A. Balkema. Rotterdam.
15. Karstunen, M., Näätänen, A., Lojander, M. 1995. The effect of natural anisotropy of clays. *Proc. 11th ECSMFE,* Copenhagen.Vol. 6: 6.83-6.88. Danish Geotechnical Society, dgf-Bulletin 11.
16. Graham. J., Noonan, M.L., Lew, K.V. 1983. Yield states and stress-strain relations in natural plastic clay. *Canadian Geotech. J. 20: 502-516.*
17. Toivanen, T-L. 1999. Modelling of the anisotropy of Otaniemi clay. MSc thesis. Helsinki University of Technology.

KERAAMISEN TUOTTEEN DEFORMOITUMINEN POLTTOPROSESSIN AIKANA

Paavo Hassinen[1], Juha Paavola[1], Pekka Valjento[1],
Pentti Hytönen[2], Risto Mannonen[3] ja Raino Ranta[3]

ABSTRAKTI

Keraamisen esineen valmistusprosessissa oleellisen vaiheen muodostaa raakamuokatun tuotteen polttaminen. Polton aikana hauras savimateriaali muuttuu sintraantumisreaktioiden vaikutuksesta kovaksi keraamiseksi materiaaliksi. Samalla esine deformoituu ja siihen syntyy pysyviä muodonmuutoksia. Tämän esityksen tarkoituksena on pyrkiä valottamaan niitä fysikaalis-kemiallisia ilmiöitä, joita keraamin polttoprosessissa tapahtuu ja esittämään näkemyksiä deformaatioiden synnystä sekä niiden ennalta arvioimisesta.

JOHDANTO

Keraamien valmistusprosessissa tapahtuva ennalta hankalasti arvioitavissa oleva muodonvääristyminen on osoittautunut varsin haitalliseksi ja kustannuksia aiheuttavaksi tekijäksi koko tuotannolle. Esineen suunnittelussa muodonvääristymistä on vaikea ennakoida, koska pienelläkin geometrian muutoksella esineen suunnitteluvaiheessa saattaa olla suuri vaikutus deformaation muodostumiseen. Ongelman ratkaisussa suunnittelija joutuu arvioimaan lopputulokseksi toivottuun esineen geometriaan ennakon, joka sitten polttoprosessin aikana deformaation edetessä häviää ja tuotteen muoto saavuttaa tavoitellun lopputuotteen geometrian. Tämä kokemusperäinen, 'yritys-erehdys'-tyyppinen muodon ennakoiminen aiheuttaa luonnollisesti merkittäviä kustannuksia lisääntyvänä työmääränä, muotti- ja materiaalikustannuksina jne. Nyt esitettävät ajatukset on kehitetty esiprojektissa, jonka lopullisena tavoitteena on kehittää numeerinen laskentamenetelmä ja sen yhteyteen keraamin materiaalimalli siten, että etukäteen suoritettavalla epälineaarisella analyysilla voitaisiin määrittää tuotteen tarvittava alkugeometria.

Varsinaisesti keraamisten talousesineiden polton aikaisia muodonmuutoksia käsitteleviä tieteellisiä artikkeleita on julkaistu varsin vähän. Aihetta sivuavia vanhojakin artikkeleita kyllä löytyy. Niissä käsitellään kuitenkin enimmäkseen menetelmiä, joilla pystytään arvioimaan savipohjaisen materiaalin muodonmuutosherkkyyttä - esimerkiksi materiaalin huokoisuuden vaikutusta deformoitumiseen ja pyroplastisen indeksin merkitystä. Modernien artikkeleiden aiheet liittyvät usein teknillisen keramiikan valmistamiseen tai tiilten poltta-

[1] Teknillinen korkeakoulu, Rakenteiden mekaniikan laboratorio, Pl 2100, 02015 TKK
[2] Hackman Designor Oy Ab
[3] Taideteollinen korkeakoulu, Keramiikka- ja lasisuunnittelun osasto, Hämeentie 135 C, 00560 Helsinki

miseen, joista suoraan hyödynnettävää tietoa keraamisten talousesineiden valmistukseen ei ole saatavilla.

VALMISTUSPROSESSI

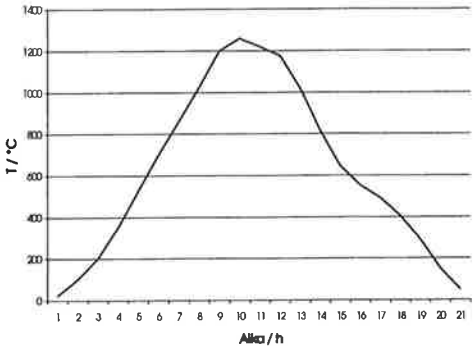
Käytännössä tyypillisen keraamisen tuotteen raakamateriaali, granulaatti, toimitetaan valmiiksi sekoitettuna kuivatuotteena tehtaalle. Raakamateriaali laivataan säkitettynä tuotantolaitokselle ja tämän jälkeen se seisoo jonkin aikaa tehtaan varastoissa. Materiaali koostuu pallomaisista, osittain ontoista, läpimitaltaan noin 0.1 ... 0.5 mm:n kokoisista granulaateista ja sen kosteuspitoisuus on pieni, noin 2 ... 2.5%. Materiaalin koostumus pysyy muuttumattomana, tosin tuotekehityksen tuloksena eri seoskomponenttien osuuksia joskus hallitusti muunnellaan tavoitteena lopputuotteen osalta tiettyjen ominaisuuksien parantaminen. Tuotteen valmistuksessa granulaatti suihkutetaan ilmanpaineen avulla, 3.5...4 bar metallimuottia vasten ja tämän jälkeen puristetaan toiselta puolelta muottia vastaan muotin toimiessa passiivisena alustana. Puristuspaine on hyvin korkea, noin 280 bar. Valmiiksi prässätty tuote irrotetaan mekaanisesti muotista. Muotin tulee olla valmistettu siten, että sen muodossa on huomioitu polton aikana tapahtuvat deformaatiot eli lopputuotteen tavoittena olevan geometrian saavuttamiseksi tarvittava ennakkomuoto.

POLTTOPROSESSI

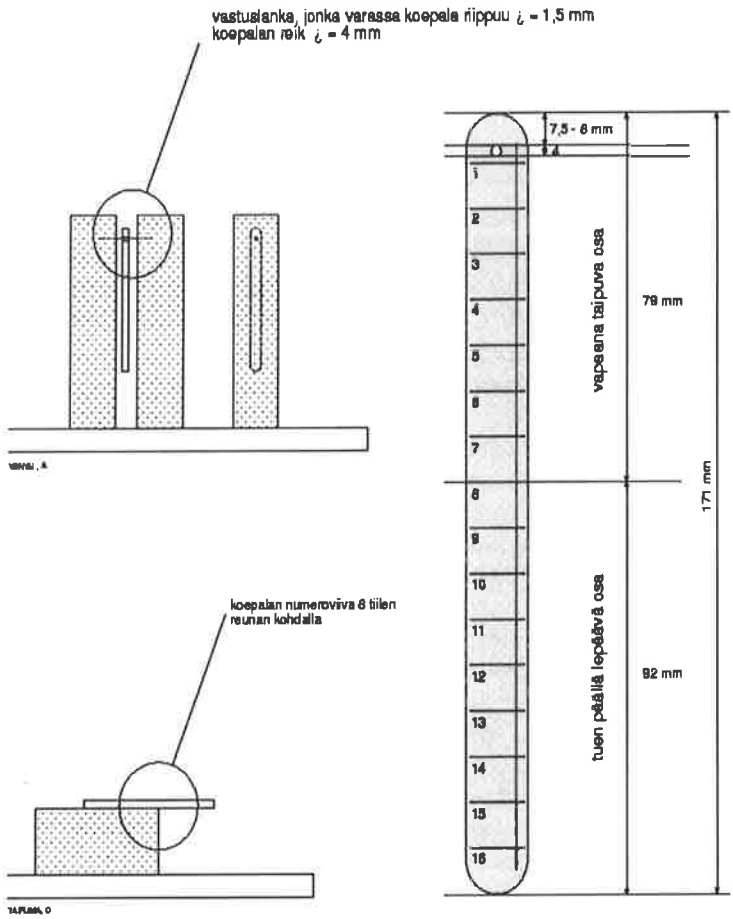
Poltto tuotantoprosessissa suoritetaan yleisesti tunneliuunissa, jonka läpi prässätyt raakakappaleet kulkevat. Koko prosessi kestää useita tunteja (jopa noin 20 tuntia, kuva 1), josta maksimilämpötilalle (> 1200°C) koekappaleet altistetaan noin kolmen tunnin ajan. Kappaleita ei kuormiteta kokeen aikana eikä myöskään tueta, joten omapaino on ainoa mekaaninen kappaleisiin kohdistuva rasitus. Polttotapahtumaa ylläpidetään syöttämällä uuniin lämpöenergiaa. Poltonohjaus eli polttoaineen syöttö suoritetaan siten, että uunissa tietyissä mittauspisteissä rekisteröity keskimääräinen lämpötila noudattaa haluttua aika-lämpötila-riippuvuutta. Lämmönsiirtyminen uunista koekappaleeseen tapahtuu sekä säteilemällä, kuljettumalla että johtumalla. Se on jatkuva prosessi, jonka vaikutuksesta uunissa vallitseva lämpötilajakauma on epätasainen ja ajan suhteen muuttuva, epästationaarinen. Raakamateriaali reagoi herkästi lämpötilamuutoksiin. Lämpötilan kohoaminen käynnistää ja voimistaa materiaalin rakennetta muuttavia fysikaalisia ilmiöitä. Samalla kemiallinen reaktioaktiivisuus kasvaa. Polton edetessä olosuhteet uunissa vaihtelevat siten, että lopputuotteen ominaisuuksien kannalta oleelliset kemialliset reaktiot ja faasimuu-tosilmiöt pääsevät tapahtumaan juuri oikeissa lämpötiloissa.

Savipohjainen raakamateriaali sisältää aina vettä. Mekaanisesti sitoutunut vesi höyrystyy lämpötila-alueella 100 ... 200°C. Kemiallisesti sitoutunut vesi (savien ja kaoliinien kidevesi) poistuu noin 450 ... 600°C:n lämpötilassa. Orgaaniset epäpuhtaudet palavat pois lämpötila-alueella 300 ... 700°C. Massan sintraantuminen (sulaminen ja tiivistyminen) alkaa noin 1000°C:n lämpötilassa. Sekä veden poistuminen että raakamassan eri komponenttien sulaminen aiheuttavat kutistumista. Kutistumisen voidaan otaksua tapahtuvan homogeenisesti. Kutistuminen, joka on suuruudeltaan noin 10 - 15 %, ja joka tapahtuu pääasiassa lämpötilan 1000°C yläpuolella muodostaa polton aikaisesta kokonaismuodon-muutoksesta suurimman osan. Kutistuminen pystytään varsin hyvin ennakoimaan esineen suunnittelussa.

Varsinkin lautasmaisilla ohuilla esineillä ongelmalliseksi muodostuu muodon eli alkuperäisgeometrian vääristyminen. Polton aikana maksimilämpötila on noin 1270 °C. Tässä lämpötilassa sintraantumisreaktiot ovat jo suurimmaksi osaksi tapahtuneet ja massa on paikoitellen lähes sulassa tilassa. Esineen oman painon merkitys kuormituksena on tässä tilanteessa suuri. Lautasmaisten esineiden ohuissa osapinnoissa materiaali on lähes kauttaaltaan pehmentynyt, joten näissä osissa muodonmuutokset muodostuvat suuriksi. Paksummissa osissa materiaali pysyy ainakin osittain jäykkänä, jol-



Kuva 1. Tunneliuunin aika-lämpötilajakauma



Kuva 2. Koesauvan mitat ja asennukset

loin muodonvääristymät jäävät verrattain pieniksi. Korkeassa lämpötilassa esineen ja sen alustan väliset kitkavoimat synnyttävät helposti deformoitumista varsinkin tukipinnan läheisyydessä.

SAUVAMAISET KOEKAPPALEET

Projektissa valmistettiin sauvamaisia koekappaleita, kuva 2. Raakasauvojen pituus oli 171 mm, ja poikkileikkausmitat olivat 15.22 mm x 8.90 mm. Sauvojen valmistuksessa pyrittiin mahdollisimman tarkoin identtiseen menetelmään teollisen tuotantoprosessin kanssa. Prässääminen tehtiin kuitenkin mekaanisella manuaalipuristimella, mikä mahdollisesti aiheutti jossain määrin normaalituotannosta eriävän materiaalirakenteen ja lisäksi hajontaa eri koekappaleiden ominaisuuksien välillä. Kun sauvat mekaanisesti irroitettiin muoteista, sauva oli kaareutunut siten, että muotin puoli muodosti koveran ja puristuspuoli puolestaan kuperan pinnan. Säilytettäessä tämän jälkeen sauvoja vaakasuoralla alustalla joitakin päiviä, sauvan alkutaipumat hävisivät ja sauvan akseli oli kutakuinkin suora. Prosessi synnyttää jonkinlaisen alkujännitystilan sauvaan.

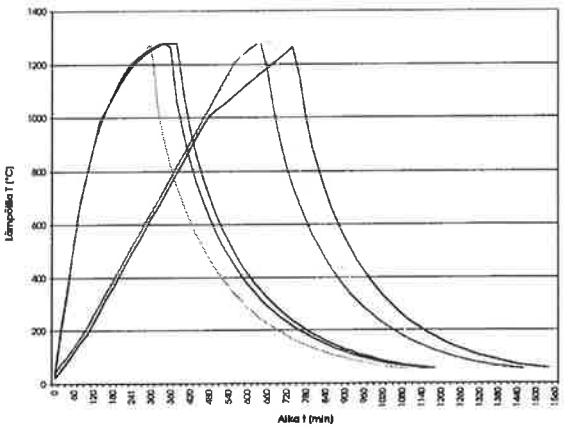
KOESAUVOJEN POLTTOKOKEET

Taideteollisessa korkeakoulussa suoritettiin esitutkimusta varten sarja polttokokeita edellämainituille sauvamaisille koekappaleille. Polttaminen tapahtui pienessä koeuunissa. Kullakin polttokerralla pyrittiin noudattamaan varsinaiseen tuotantoon käytetyn tunneliuunin aika-lämpötilariippuvuutta haluttuun maksimilämpötilaan asti. Kokeita tehtiin siten, että maksimilämpötilat vaihtelivat lämpötila-alueella 500 - 1300 °C. Maksimilämpötilan ollessa välillä 1200 - 1300 °C tehtiin yksi koe 20 °C:een välein, ts. $T_{max} = 1300$ °C, 1280 °C, 1260 °C jne., välillä 1000 - 1200 °C 40 °C:een välein ja lämpötilavälillä 500 - 1000 °C 100 °C:een välein. Koejärjestelyt oli valittava tällaisiksi, jotta saataisiin tietoa lämpötilahistorian vaikutuksesta sauvan deformaatioihin, sillä kyseiseen uuniin ei ollut mahdollista asentaa mittalaitteita, joilla olisi kokeen aikana voitu mitata deformaation kehittymistä. Näinollen varioitavia parametreja kokeiden suorituksessa olivat maksimilämpötila, lämmittämisnopeus, maksimilämpötilan ylläpitoaika sekä jäähdyttämisnopeus. Tietoa deformaation kehittymisestä saatiin mittaamalla eri lämpötiloista jäähdytettyjen koekappaleiden mitat ja massa kunkin kokeen jälkeen. Mittausten helpottamiseksi koekappaleisiin oli painettu suorapainomenetelmää käyttäen poltonkestävällä väripigmentillä numeroitu mitta-asteikko. Asteikon viivojen väli oli 1 cm.

Polttokokeissa käytettiin kolmea keskenään erilaista koejärjestelyä, joilla pyrittiin selvittämään sauvamaisten koekappaleiden muodonmuutoksien kehittymistä. Näissä osa sauvoista oli ripustettu toisen päänsä läpi pujotetun vastuslangan varaan ja tuettu uunin rakenteisiin. Tällöin oman painon aiheuttama kuormitus sai aikaan aksiaalisen vetorasituksen. Osa oli tuettu noin puolelta pituudeltaan vaakasuoralla tasolla siten, että ne toimivat ulokesauvoina ja tukemattomaan päähän syntyi oman painon kuormituksesta taipumia, kuva 2. Näissä kokeissa koesauvat asennettiin vaihtoehtoisesti kahteen eri asemaan siten, että prässäyksessä muottia vastaan ollut pinta oli joko yläpintana tai alapintana. Kokonaan vaakasuoralla tasolla tuettuja sauvoja käytettiin kokeissa, joilla yritettiin selvittää kitkan vaikutusta muodonmuutosten suuruuteen. Tällöin koekappaleet sijaitsivat suoraan uunilevyn pinnalla tai vaihtoehtoisesti "alumiinioksidi-patjan" päällä. Koekappaleita ei erikseen kuormitettu kokeen aikana. Poltettujen sauvojen mittojen ja

massojen lisäksi mitattiin ulokesauvojen vapaan pään taipuma-arvot.

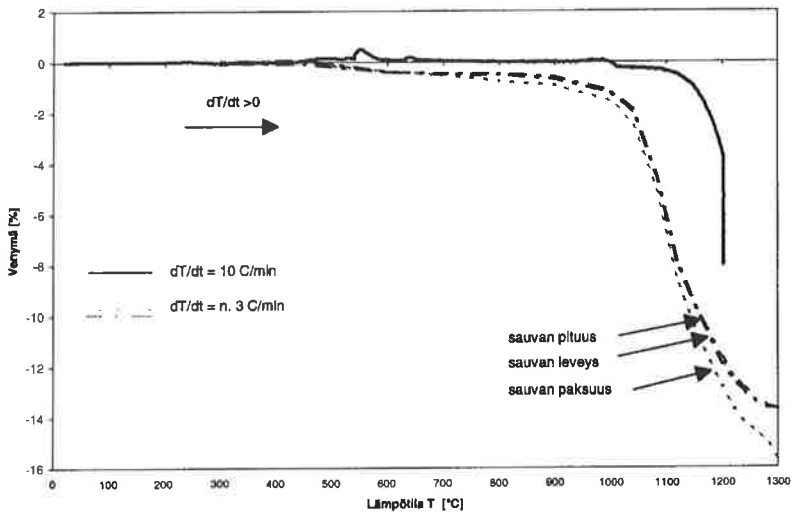
Uunin lämpötilaa nostettiin kahdella eri nopeudella; hidas poltto noin 2 °C/min ja nopea poltto noin 8 °C/min. Maksimilämpötilan ylläpitoaika oli joko 0, 30 tai 60 minuuttia. Kuvassa 3 on esitetty sarja uunin lämpötilajakaumista kokeista, joissa $T_{max} = 1280°C$. Esitetty sarja käsittää kuusi erillistä koetta, joissa jokaisessa on kolme koesauvaa. Tämä edustaa yhtä tyypillistä koesarjaa.



Kuva 3. Erään koesarjan lämpötilahistoria

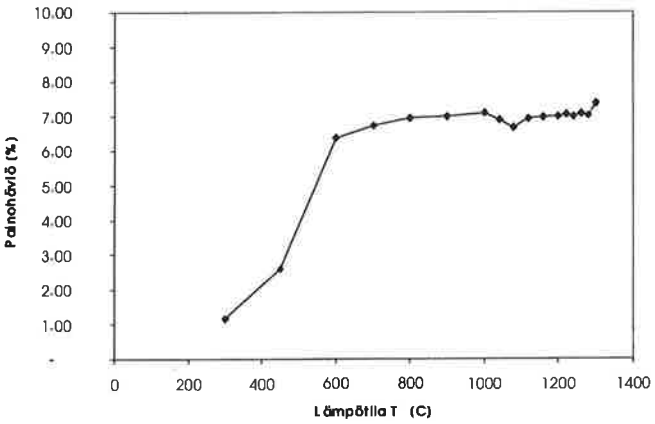
HAVAINTOJA KOETULOKSISTA

Tässä kappaleessa esitetään johtopäätöksiä, joita on voitu tehdä kokeiden jälkeen suoritettujen mittausten perusteella. Kaikkien koekappaleiden osalta voidaan todeta, että pääosa kutistumasta tapahtuu 1000°C:n yläpuolella. Tätä ennen on havaittavissa, lämpötilaan noin 500 - 600 °C saakka koekappaleiden lämpölaajenemista. Erikoista tuloksissa on se, että kutistuminen on jokseenkin identtistä sauvojen pituus ja poikkisuunnissa. Sen sijaan paksuuden suunnassa kutistuminen on joitakin prosenttiyksikköjä voimakkaampaa, kuva 4 katkoviivat. Ero kuvannee materiaaliominaisuuksien anisotrooppisuutta, joka syntyy, kun sauvat puristetaan raakamassasta. Koekappa-



Kuva 4. Koekappaleiden keskimääräisen kutistuman riippuvuus lämpötilasta, ehyt viiva kuvaa kuumavetouunissa testatun sauvan kutistumaa.

leiden painohäviö kasvaa jyrkimmin lämpötilan ollessa 300 - 600 °C, tämän jälkeen paino pysyy lähes muuttumattomana, kuva 5. Epäilemättä saadut mittaustulokset ainakin osittain riippuvat sauvojen ainepaksuudesta.



Kuva 5. Koekappaleiden keskimääräinen painohäviö.

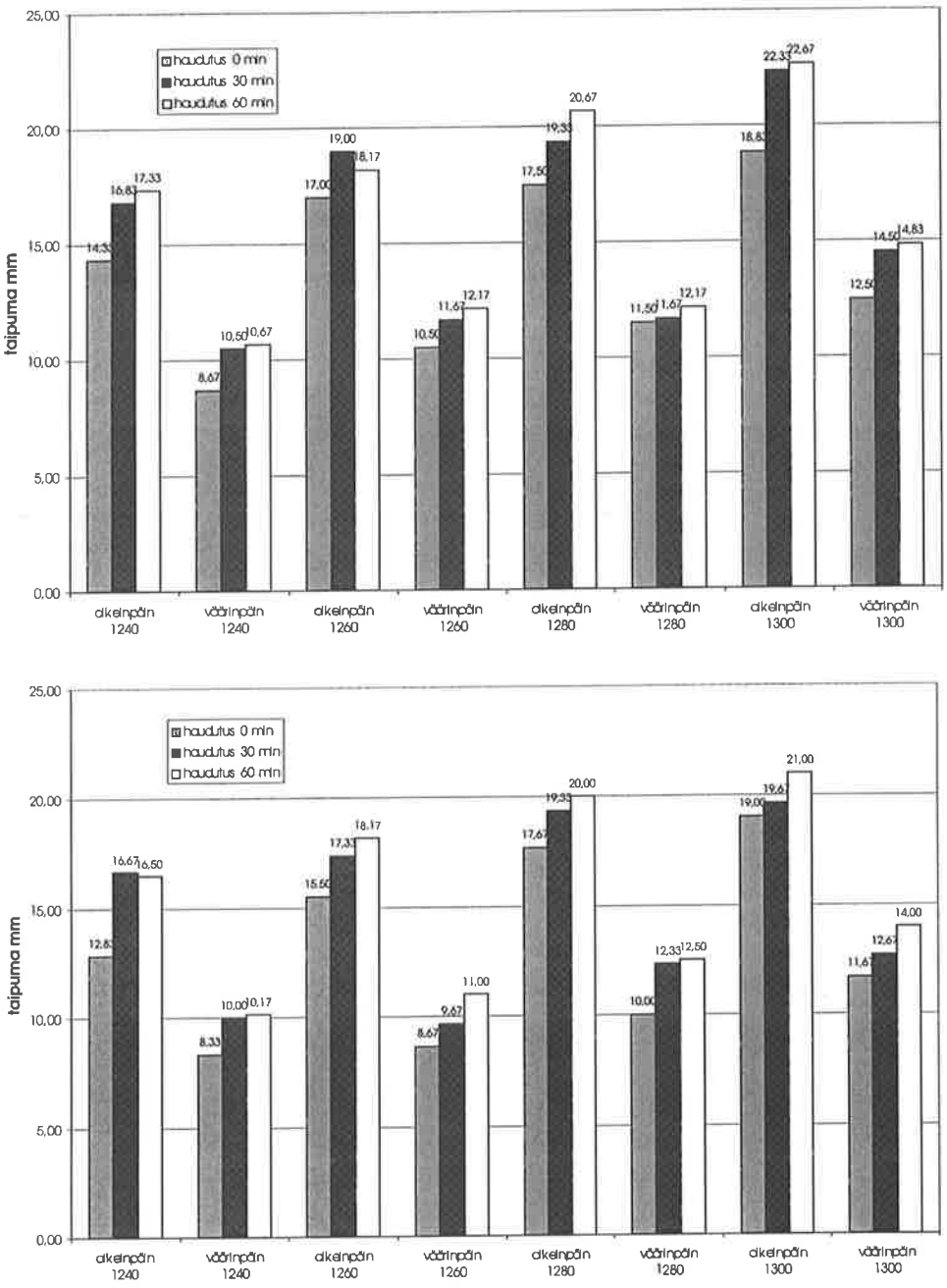
Kuvassa 6 on esitetty ulokkeen vapaan pään taipumat eri koesarjoista. Kuvasta on nähtävissä lämpiämisnopeuden ja maksimilämpötilalle altistamisajan vaikutus taipuma-arvoihin. Oleellisin havainto on ulokkeen pään taipuman erisuuruus: kun koesauva oli asennettu uuniin prässäyksessä muottia vasten ollut pinta yläpintana oli taipuma noin 60% niiden sauvojen taipumasta, joissa muottia vasten ollut pinta oli alapintana. Kuvassa 7 on hahmoteltu taipumaviivan muoto kummassakin tapauksessa. Taipumaviiva viittaisi jälkimmäisessä tapauksessa epätasaiseen kutistumiseen, mutta ripustetuissa tai suoraan alustalle asennetuissa koekappaleissa ei tällaista epäsymmetristä kaareutumista ole havaittavissa.

Ripustettujen sauvojen osalta venymien merkitys kutistumisen rinnalla jää merkityksettömäksi, eikä sitä pystytä koetuloksista erottamaan. Myöskään kitkan merkityksellä alustan ja koekappaleen välillä niissä tapauksissa, joissa alustan pintamateriaali oli erilainen, ei ole oleellista merkitystä.

Koesauvoista mitattiin polttokokeen jälkeen myös huokoisuus. Kuvassa 8 on esitetty eri lämpötiloista jäähdytettyjen sauvojen huokoisuus kokeen maksimilämpötilan funktiona. Tämä on hyvin sopusoinnussa kutistumakäyrien kanssa kuva 4.

KOESAUVOJEN KUTISTUMISKOKEET

Joidenkin raakasauvojen kutistumista tutkittiin käyttäen TKK:n Rakennus- ja ympäristötekniikan osastolla olevaa kuumavetouunia. Uunissa saavutettavissa oleva suurin lämpötila on 1200 °C. Oleellinen ero edellisiin kokeisiin on se, että koekappaleen yhden akselin suuntaista muodonmuutosta voidaan mitata jatkuvasti kokeen aikana. Kuvassa 4 on esitetty erään tyypillisen koekappaleen venymän muuttuminen lämpötilan mukana (ehyt viiva). Koekappale oli pystyasennossa ilman ulkoista kuormaa, jolloin jännityksiä syntyi ainoastaan omanpainon vaikutuksesta. Lämpötilan muutosnopeus lämmitysvaiheessa oli 10 °C/

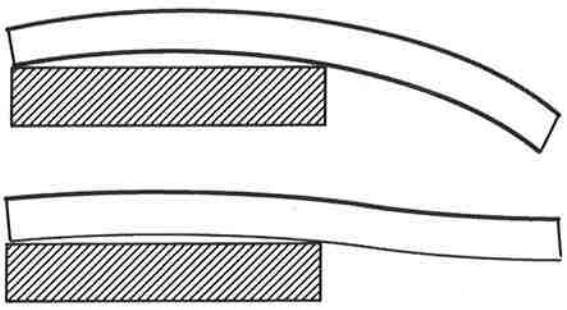


Kuva 6. Ulokesauvojen vapaan pään taipumat eri kokeissa. a) Lämpötilan nousunopeus 2 °C/min ja b) 8 °C/min. Kunkin pylväsryhmän ensimmäinen pylväs kuvaa maksimilämpörasituksen kestoaikaa t = 0 min, toinen pylväs kestoaikaa 30 min ja kolmas 60 min.

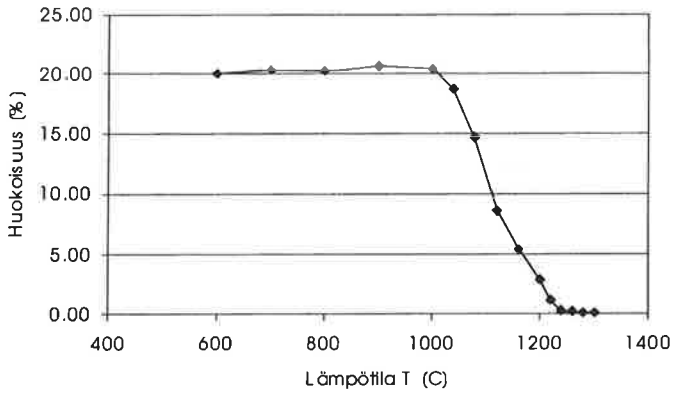
min. Kuvasta havaitaan, että koekappale laajenee lämpötilaan noin 700 °C saakka ja kutistuu voimakkaasti, kun lämpötila kohoaa yli 1000 °C.

KOESAUVOJEN TAIVUTUSKOKEET

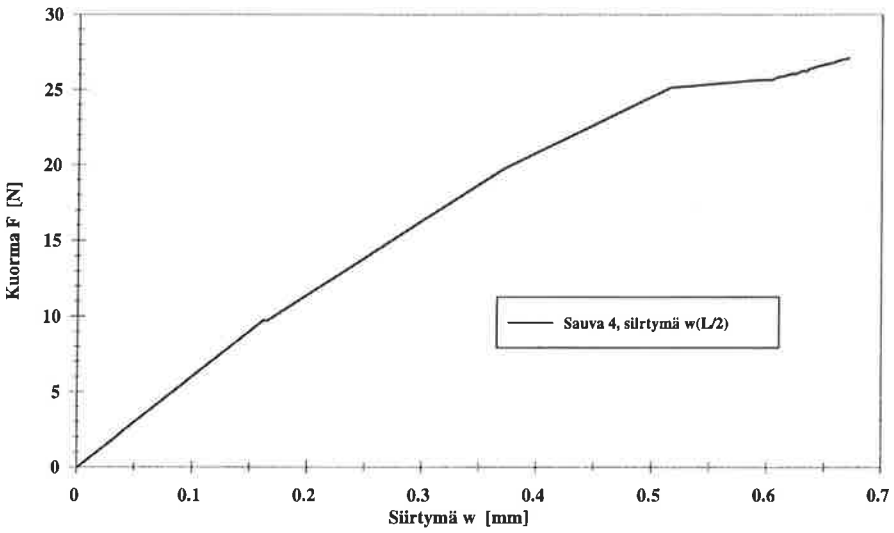
Koesauvojen epäsymmetrisen käyttäytymisen vuoksi tutkittiin raakasauvan taivutusvetolujuutta huoneen lämpötilassa, kun sauva oli asennettu eri asentoihin: muottia vasten ollut pinta oli yläpintana, alapintana tai sauva oli kyljellään. Raakamateriaalin alkukimmokertoimen ja taivutusvetolujuuden määrittämiseksi sauvat kuormitettiin kolmannespisteisiin sijoitetuilla viivakuormilla murtotilaan asti. Sauvojen voima-siirtymäkuvaaja on hieman kaareva, mikä kertoo materiaalin plastisoitumisesta ja mahdollisesta säröilystä. Sauvat murtuivat hauraan aineen tavoin (kuva 9). Taivutuskimmokertoimen arvo raaka-sauvan materiaalille on luokkaa 2400 N/mm^2 ja taivutusvetolujuuden arvo 3 ... 4 N/mm^2. Eri tavoilla orientoitujen sauvojen materiaalin alkukimmokertoimen arvoissa ei havaittu merkittäviä eroja. Mikroskooppikuvauksilla pyrittiin vielä havainnoimaan mahdollisesta eri pinnoilla esiintyvästä erilaisesta materiaalirakenteesta, mutta havaittavia eroja ei onnistuttu löytämään.



Kuva 7. Taipumaviivojen muodot.



Kuva 8. Koesauvan huokoisuus lämpötilan funktiona.



Kuva 9. Kolmannespisteissä viivakuormilla kuormitetun raakasauvan voima-siirtymä kuvaaja. Jännemitta L=150 mm.

LASKENTA-ANALYYSI

Polttoprosessin numeerinen simulointi on käytännössä varsin vaativa tehtävä. Poltonaikainen muodonmuutos syntyy usean eri tekijän vaikutuksesta. Luotettavien tulosten saavuttamisen edellytyksenä tulee olla selkeä kuva keraamin rakenteesta ja lisäksi käsitys kaikista niistä fysikaalis-kemiallisista ilmiöistä ja reaktioista, joita polttoprosessin aikana itse keraamissa tapahtuu. Erityisesti on kiinnitettävä huomiota eri ilmiöiden toisiinsa kytkeytymiseen, samoin myös niiden syntyjärjestyksellä on merkitystä. On myös huomattava, että polton aikaisiin ilmiöihin ei vaikuta yksinomaan tarkasteluhetkellä vallitseva lämpötila, vaan myös koko siihenastinen lämpötilahistoria. Polton pitkäkestoisuudesta ja syntyvien muodonmuutosten suuruudesta johtuen, pienetkin virheet lähtötiedoissa saattavat vaikuttaa merkittävästi tulosten oikeellisuuteen.

Laskennallista tarkastelua varten tulee raakakeraamin materiaaliominaisuudet tuntea riittävän tarkasti. Jotta numeerisella laskentamallilla pystyttäisiin oikein ennakoimaan rakenteen ja materiaalin käyttäytymistä, materiaalimallin tulisi ottaa huomioon eri osatekijöiden vaikutus:

$$\varepsilon = \varepsilon^{el} + \varepsilon^{th} + \varepsilon^{pl} + \varepsilon^{cr} + \varepsilon^{sw} + \varepsilon^{ph}$$

Tässä kokonaismuodonmuutos koostuu kimmoisten, termisten, plastisten ja viskoelastisten muodonmuutosten lisäksi materiaalin paisumisesta ja tapahtuvista faasimuutoksista.

Raakamateriaali on rakenteeltaan ilmeisesti jonkin verran epähomogeenista ja materiaaliominaisuudet siten erilaisia rakenteen eri osissa ja eri suunnissa. Veden poistuminen tapahtuu lämpötiloissa, joissa materiaali ei ole vielä pehmentynyt niin paljoa, että esimerkiksi kitkan vaikutuksesta esineen muoto vääristyisi oleellisesti. Uunin maksimilämpötilassa sintraantumisreaktiot ovat jo suurimmaksi osaksi tapahtuneet ja esineen omanpainon merkitys kuormituksena on tässä tilanteessa suuri. Lautasmaisten esineiden

ja mitkä johdettuja lauseita. Varsinkin siirtyminen Newtonin mekaniikasta Lagrangen mekaniikkaan herättää tämänkaltaisia kysymyksiä.

Peruslakeihin pohjautuva fysikaalinen Newtonin hiukkasmekaniikka ja jäykän kappaleen mekaniikka eli niin sanottu vektoroitu mekaniikka on esitetty lukuisissa kirjoissa [Salmi 1996-1998; Kurki-Suoniot, 1990; Salonen 1989; Beer & Jonston 1976]. Nämä kirjat edustavat peruslakeihin nojautuvaa lähestymistapaa, missä kirjojen lähtökohtana on fysikaalisuus eli luonnontieteeseen ja havaintoihin pohjautuva lähtökohta.

Perusoletuksiin eli aksioomiin pohjautuvaa matemaattista lähestymistapaa edustaa Truesdellin kirja järkiperäisestä jatkumomekaniikasta (*rational continuum mechanics*) [Truesdell, 1977]. Siinä asetetaan ensin perusoletukset ja määritelmät, joiden pohjalta mekaniikassa käytetyt lauseet suoraan johdetaan. Truesdell erottaa toisistaan fysikaaliset oliot ja matemaattiset oliot ja asettaa matemaattisille olioille perusoletuksia eli aksioomia. Truesdellin määritelmän mukaan mekaniikka on matemaattisten mallien ääretön luokka, jolla pyritään mallintamaan luontoa jollakin tavalla tai näkökulmalla. Tätä perusoletuksiin nojautuvaa johtamistapaa voidaan pitää yhtenä tieteen tärkeimmistä tavoitteista. Tässä työssä pyritään esittämään Newtonin ja Lagrangen mekaniikan peruslait matemaattisesti ilman fysikaalista painolastia.

Desloge esittää kirjassaan klassiselle mekaniikalle viisi perusoletusta [Desloge, 1982, sivut 11, 12, 20, 56 ja 216]: inertiaalikehyksen (*inertial frame*) olemassaolon, suhteellisuusperiaatteen, kappaleen nopeuden ylärajattomuuden, kappaleiden vuorovaikutuslain sekä oletuksen järjestelmän sisäiselle voimasysteemille. Näiden perusoletusten pohjalta voidaan johtaa jatkavuuden lause (Newtonin I laki), voiman ja vastavoiman lause (Newtonin III), lisäksi voima määritellään Newtonin II lain avulla sekä massalle annetaan oma määritelmänsä, mikä perustuu kappaleiden törmäykseen.

Tässä työssä ensin määritellään käsitteet hiukkanen ja jäykkä kappale sekä kootaan yhteen mekaniikan peruslakeja: fysikaalisen Newtonin hiukkasmekaniikan ja jäykän kappaleen mekaniikan peruslait; matemaattisen Newtonin hiukkasmekaniikan peruslait; matemaattisen Lagrangen hiukkasmekaniikan peruslait. Matemaattisella mekaniikalla tarkoitetaan mekaniikkaa, josta fysikaalinen eli luonnonilmiöihin ja havaintoihin sidottu painolasti on riisuttu. Peruslaeilla tarkoitetaan sekä perusoletuksia eli aksioomia eli postulaatteja että perusmääritelmiä, jotka olennaisia mekaniikan oppirakenteen kannalta.

HIUKKANEN, JÄYKKÄ KAPPALE JA VAPAA KAPPALE

Hiukkasen määritelmä: Hiukkasella (*particle*) tarkoitetaan kappaletta, jonka mitat ovat kyseessä olevan tehtävän kannalta epäoleellisia.

Jäykän kappaleen määritelmä: Jäykällä kappaleella tarkoitetaan kappaletta, joka koostuu äärellisestä määrästä hiukkasia, joiden keskinäinen etäisyys pysyy muuttumattomana kuormitettiin kappaletta miten tahansa.

Vapaan kappaleen määritelmä: Vapaalla kappaleella tarkoitetaan kappaletta, joka ei ole vuorovaikutuksessa muiden kappaleiden kanssa [Kurki-Suoniot, 1990, s. 89]

FYSIKAALISEN NEWTONIN MEKANIIKAN PERUSLAIT

Hiukkasjärjestelmän ja jäykän kappaleen järjestelmän fysikaalisen Newtonin mekaniikan oppijärjestelmän perustana ovat seuraavat 7 peruslakia:

1. On olemassa absoluuttinen, fysikaalinen euklidinen paikka-avaruus ja absoluuttinen fysikaalinen aika.
2. Voima on fysikaalisen vektoriavaruuden alkio eli fysikaalinen vektorisuure, ts. voima toteuttaa suunnikassäännön. Voima on kiinteä eli vaikutuspisteessään vaikuttava vektorisuure.
3. Voiman siirtolaki: Jos voima, joka vaikuttaa jäykkään kappaleeseen, siirretään pitkin vaikutussuoraansa, sen ulkoinen vaikutus kappaleeseen pysyy muuttumattomana.
4. Jatkavuuden laki eli Newtonin I laki: Kaikki vapaat kappaleet liikkuvat tasaisesti ja suoraviivaisesti toistensa suhteen.
5. Dynamiikan peruslaki eli Newtonin II laki: Jos kappaleeseen vaikuttavien fysikaalisten voimien summa eli resultantti on $\vec{R}$, niin kappale saa fysikaalisen kiihtyvyyden $\vec{a}$, siten että on voimassa yhtälö
   $$\vec{R} = m_h \vec{a}$$
   Verrannollisuuskerrointa m_h sanotaan hitaaksi massaksi.
6. Voiman ja vastavoiman laki eli Newtonin III laki: Jos hiukkanen A vaikuttaa hiukkaseen B jollakin voimalla, niin hiukkanen B vaikuttaa aina takaisin hiukkaseen A hiukkasten yhdysjanan suuntaisella, yhtä suurella mutta vastakkaissuuntaisella voimalla.
7. Yleinen gravitaatiolaki eli Newtonin IV laki: Kaksi hiukkasta, joiden painavat massat ovat m_{p1} ja m_{p2} vetävät aina toisiaan puoleensa yhdysjanan suuntaisella voimalla, jonka suuruus on suoraan verrannollinen hiukkasten painaviin massoihin ja kääntäen verrannollinen niiden välisen etäisyyden r neliöön eli
   $$F = \gamma \frac{m_{p1} m_{p2}}{r^2}$$
   Verrannollisuuskerrointa γ kutsutaan gravitaatiovakioksi.

Edellä olevat peruslait on pitkälti lainattu kirjasta [Salmi, 1998, s. 17]. Perinteisesti neljäs peruslaki esitetään muodossa: Hiukkanen on levossa tai tasaisessa suoraviivaisessa liikkeessä aina, kun siihen ei vaikuta voimia tai siihen vaikuttavien voimien summa eli resultantti on nolla. Tätä kutsutaan hitauden laiksi. Neljäs peruslaki eli jatkavuuden laki on käsitteenmuodostuksen kannalta parempi, sillä lain avulla voidaan määrittää inertiaalikehys ja samalla joukko inertiaalikehyksiä, jotka ovat tasaisessa suoraviivaisessa liikkeessä toisiinsa nähden. Jatkavuuden laki tekee tasaisesta liikkeestä absoluuttisen eli havaitsijasta riippumattoman ominaisuuden sekä lisäksi asettaa ajalle ja avaruudelle euklidisuuden, homogeenisuuden ja isotrooppisuuden vaatimukset [Kurki-Suoniot, 1990, s. 89-91]. Toisaalta, jos hitauden laki korvataan jatkavuuden lailla, niin tasapainon käsite pitää erikseen määritellä, jolloin statiikka on dynamiikan erikoistapaus.

Dynamiikan peruslaki määrittelee hitaan massan kappalekohtaiseksi ominaisuudeksi, joten lakia ei voida käyttää voiman määrittelemiseen. Hitaan ja painavan massan samaistaminen eli ekvivalenttisuus on esitetty kirjoissa [Salmi, 1998, s. 87; Kurki-Suoniot, 1990, s. 91-92] sekä kinemaattisen kiihtyvyysvektorin ja fysikaalisen kiihtyvyysvektorin samaistaminen kirjassa [Salmi, 1996, liite I].

Kolmas ja viides peruslaki voidaan yhdistää, jos lait muotoillaan hieman toisin: Jos järjestelmän hiukkanen A vaikuttaa saman järjestelmän hiukkaseen B voimalla $\vec{F}_{BA}$, niin hiukkanen B vaikuttaa hiukkaseen A hiukkasten yhdysjanan suuntaisella, yhtä suurella mutta vastakkaissuuntaisella voimalla $\vec{F}_{AB}$, lisäksi jäykälle kappaleelle pätee yhteys $\vec{F}_{AB} = -\vec{F}_{BA}$. Tästä yhtäsuuruusehdosta seuraa voiman siirtolaki sekä se, että jäykkä kappale on sisäisesti tasapainossa. Voiman ajatellaan olevan toisen peruslain mukaisesti kiinteä vektorisuure, joka siis vaikuttaa avaruuden pisteessä eikä tätä vaikutuspistettä voida siirtää mielivaltaisesti muuttamatta järjestelmän liiketilaa.

Monasti myös ajatellaan Newtonin gravitaatiolain eli seitsemännen peruslain olevan vain eräs voimavaikutuskaava, kuten kitkavoimakaavat ja Hooken kaava, jolloin seitsemäs peruslaki voidaan poistaa. Tällöin on kuitenkin tehtävä lisäoletus, että hidas ja painava massa voidaan samaistaa.

MATEMAATTISEN NEWTONIN MEKANIIKAN PERUSLAIT

Hiukkasjärjestelmän matemaattisen Newtonin mekaniikan oppijärjestelmän perustana ovat seuraavat 6 perusoletusta (**O#**) ja 6 perusmääritelmää (**M#**):

O1. On olemassa kiinteä, riippumaton, tasa-aineinen, geometrialtaan euklidinen, matemaattinen paikka-avaruus ja riippumaton, tasainen matemaattinen aikaparametri. On olemassa koko avaruuden peittävä kiinteä karteesinen paikkakoordinaatisto, jota kutsutaan kiinteäksi inertiaalikoordinaatistoksi, jossa mekaniikan lait ovat samanarvoisia eli ekvivalentteja asemasta, suunnasta ja ajan hetkestä riippumatta ja jossa mekaniikan lait saavat yksikertaisimman muodon.

O2. Suhteellisuusperiaate: Kaikki mekaniikan lait, jotka ovat voimassa kiinteässä inertiaalikoordinaatistossa, ovat voimassa myös kaikissa koordinaatistoissa, jotka ovat tasaisessa suoraviivaisessa liikkeessä tämän kiinteän inertiaalikoordinaatiston suhteen, ja saavat täsmälleen saman muodon. Näitä koordinaatistoja myös kutsutaan inertiaalikoordinaatistoiksi.

O3. Hiukkasen nopeudella ei ole ylärajaa minkään koordinaatiston suhteen.

O4. Massa eli ainepaljous on hiukkaselle ominainen häviämätön ja yhteenlaskettava skalaarisuure. Massalla on sekä hitaus- että vetovoimaomaisuudet, jotka samaistetaan.

O5. Vuorovaikutuslaki: Jos kaksi tai useampi muutoin vapaata hiukkasta ovat vuorovaikutuksessa toistensa kanssa, niin suure

$$\sum_{i=1}^{n} m_i \mathbf{v}_i$$

saa täsmälleen saman arvon ennen vuorovaikutusta ja vuorovaikutuksen jälkeen, missä n on vuorovaikuttavien hiukkasten lukumäärä ja m_i on hiukkasen i massa ja $\mathbf{v}_i$ on hiukkasen i matemaattinen (kinemaattinen) nopeus.

O6. Jos hiukkasten lukumäärä on mekaanisessa järjestelmässä riittävän suuri, niin järjestelmä käyttäytyy kuin hiukkasparin keskinäiset vuorovaikuttavat voimat suuntautuisivat pitkin hiukkasparin yhdyssuoraa.

M1. **Liikemäärä määritelmä:** Kiinteässä inertiaalikoordinaatistossa hiukkasen massan m ja nopeuden $\mathbf{v}$ tulo määritellään yhtä suureksi kuin hiukkasen matemaattinen liikemäärä $\mathbf{p}$

$$\mathbf{p} = m\mathbf{v}$$

M2. **Voiman määritelmä** eli Newtonin II laki: Inertiaalikoordinaatistossa hiukkaseen vaikuttava matemaattinen voima **f** määritellään yhtä suureksi kuin hiukkasen liikemäärän muutosnopeus ajan suhteen eli aikaderivaatta

$$\mathbf{f} = \dot{\mathbf{p}}$$

tai samanarvoisesti

$$\mathbf{f} = m\mathbf{a},$$

missä **a** on hiukkasen matemaattinen (kinemaattinen) kiihtyvyys ($\mathbf{a} = \dot{\mathbf{v}}$) ja m on hiukkasen massa.

M3. **Sisäisten ja ulkoisten voimien määritelmä:** Rajatun ja eristetyn järjestelmän sisäisillä voimilla tarkoitetaan saman järjestelmän hiukkasten keskinäisiä vuorovaikuttavia voimia. Voiman ja vastavoiman lauseen perusteella nämä voimat ovat saman suuruisia mutta vastakkaissuuntaisia voimia [Desloge, 1982, s. 60]. Voimat, jotka eivät ole sisäisiä, ovat ulkoisia voimia. Ulkoisilla voimilla ei ole vastavoimaa järjestelmän sisällä.

M4. **Liikemäärän momentin määritelmä:** Hiukkasen $\mathscr{H}$ liikemäärän momentti mielivaltaisen pisteen A suhteen määritellään

$$\mathbf{l} = \mathbf{r} \times \mathbf{p},$$

missä **r** on hiukkasen $\mathscr{H}$ etäisyys pisteestä A ja **p** on hiukkasen liikemäärä pisteeseen A nähden.

M5. **Momentin määritelmä:** Hiukkasen $\mathscr{H}$ momentti mielivaltaisen pisteen A suhteen määritellään

$$\mathbf{m} = \mathbf{r} \times \mathbf{f},$$

missä **r** on hiukkasen $\mathscr{H}$ etäisyys pisteestä A ja **f** on hiukkaseen $\mathscr{H}$ vaikuttava voima.

M6. **Staattisen tasapainon määritelmä:** Hiukkasjärjestelmän sanotaan olevan staattisessa tasapainossa, jos on mahdollista löytää inertiaalikoordinaatisto, jossa kaikki hiukkaset ovat levossa.

Peruslait, oletusta (**O4**) lukuun ottamatta, on lainattu suurimmalta osin kirjasta [Desloge, 1982]. Perusoletuksella (**O4**) asetetaan massan olemassaolo ja tietylle hiukkaselle ominainen massan suuruus. Massan häviämättömyyden oletus tarvitaan, sillä muutoinhan ei olisi takeita massan säilymiselle edes eristetyssä järjestelmässä. Perusoletusten luetteloa voidaan vielä täydentää energian häviämättömyyden oletuksella eli termodynamiikan ensimmäisellä pääsäännöllä, mutta tämä ei kuulu enää Newtonin mekaniikan piiriin.

Kaikki edellä olevat suureet ovat matemaattisia eikä niillä ole fysikaalista ominaisuutta, mutta niillä on fysikaalinen tulkinta fysikaalisen ja matemaattisen Newtonin mekaniikan peruslakien mukaisesti. Matemaattinen (kinemaattinen) asema ja nopeus määritellään kinematiikassa vapaiksi vektorisuureiksi, joten peruslakien mukaisesti matemaattinen liikemäärä ja matemaattinen voima ovat myös vapaita vektorisuureita, toisin sanoen asema- ja nopeusvektorit eivät ole kiinnitettyjä vaikutuspisteisiinsä. Kuitenkin vapaalla vektorisuureella on eräänlainen polttomerkki, joka ilmaisee mihin olioon kyseinen vektori vaikuttaa tai kuuluu.

Perusmääritelmän (**M2**) mukaisesti annetaan voimalle arvosisältö, joka vastaa liikemäärän muutosnopeutta, joten hiukkaseen vaikuttava voima on yhtä suuri, ei sama asia, kuin hiukkasen kokema liikemäärän muutosnopeus. Mainittakoon, että hitauden lause (Newtonin I laki), voiman ja vastavoiman lause (Newtonin III), liikemäärän säilyminen lause eristetyssä hiukkasjärjestelmässä ja voiman siirtolause jäykälle kappaleelle voidaan johtaa käyttäen perusoletuksia ja –määritelmiä, [Desloge, 1982, s. 12, 60, 57, 256]. Voiman ja vastavoiman lauseen perusteella vuo-

rovaikuttavien voimien vaikutussuunnat eivät välttämättä kohtaa vaikutuspisteiden yhdyssuoran kanssa, joten kuudes perusoletus (**O6**) ei ole turha.

SIDOSTEN DIFFERENTIAALIGEOMETRIAA

Tavallisesti variaatio määritellään Gateaux-variaationa, siis eräänlaisena suuntimana, [Reddy, 1987, s. 154; Oden & Reddy, 1976, s.143; Sagan, 1969, s. 26]. Gateaux-variaatiota funktionaalista $\mathrm{I}(\mathbf{u})$ voidaan merkitä $\delta_g \mathrm{I}(\mathbf{u};\Delta\mathbf{u})$, joka on homogeeninen muutoksen $\Delta\mathbf{u}$ suhteen $(\delta_g \mathrm{I}(\mathbf{u};\lambda\Delta\mathbf{u}) = \lambda\delta_g \mathrm{I}(\mathbf{u};\Delta\mathbf{u}))$ mutta ei ole välttämättä additiivinen $(\delta_g \mathrm{I}(\mathbf{u};\Delta\mathbf{u}_1 + \Delta\mathbf{u}_2) \neq \delta_g \mathrm{I}(\mathbf{u};\Delta\mathbf{u}_1) + \delta_g \mathrm{I}(\mathbf{u};\Delta\mathbf{u}_2))$. Gateaux-variaatio ei siis välttämättä linearisoi käsiteltävää funktionaalia.

Toisaalta tapauksissa, jossa Gateaux-variaatio on lineaarinen ja jatkuva muutoksen $\Delta\mathbf{u}$ suhteen, niin lineaarinen Gateaux-derivaatta on olemassa pisteessä $\mathbf{u}$ mutta ei ole välttämättä olemassa tämän pisteen ympäristössä. Varsinkin iteratiivisissa ratkaisumenetelmissä on järkevää vaatia, että Gateaux-derivaatta on lineaarinen pisteessä $\mathbf{u}$ ja on olemassa myös pisteen ympäristössä, tällöin Gateaux- ja Fréchet-derivaatat ovat olemassa ja saavat saman arvon, [Oden & Reddy, 1976, s. 24].

On siis varsin perusteltua määritellä variaatio Fréchet-variaationa, jolloin variaatioon liittyvä Fréchet-derivaatta on pisteessä $\mathbf{u}$ lineaarinen ja jatkuva, ja derivaatta on olemassa myös tämän pisteen ympäristössä.

Seuraavassa esitellään lyhyesti mekaanisen järjestelmän sidoksiin liittyvää differentiaaligeometriaa, mikä on pitkälti poimittu kirjoista [Abraham ja muut, 1983; Marsden & Hughes, 1983; Rheinbold, 1986; Arnold, 1978]

Differentioituvuuden määritelmä: Olkoon $\boldsymbol{E}$ ja $\boldsymbol{F}$ euklidisia avaruuksia. Kuvaus $\mathbf{h}:\mathscr{U} \subset \boldsymbol{E} \to \boldsymbol{F}$ on differentioituva pisteessä $\mathbf{u}_0 \in \mathscr{U}$, jos on olemassa lineaarinen operaattori $\mathrm{D}\mathbf{h}(\mathbf{u}_0) \in \mathscr{L}(\boldsymbol{E},\boldsymbol{F})$ siten, että

$$\mathbf{h}(\mathbf{u}_0 + \Delta\mathbf{u}) = \mathbf{h}(\mathbf{u}_0) + \mathrm{D}\mathbf{h}(\mathbf{u}_0)\cdot\Delta\mathbf{u} + \omega(\mathbf{u}_0,\Delta\mathbf{u}), \tag{1}$$

missä $\lim_{\Delta\mathbf{u}\to\mathbf{0}} \dfrac{\|\omega(\mathbf{u},\Delta\mathbf{u})\|_F}{\|\Delta\mathbf{u}\|_E} = 0$ ja $\mathbf{u}_0 + \Delta\mathbf{u} \in \mathscr{U}$.

Euklidisella avaruudella tarkoitetaan reaalista, äärellisulotteista sisätuloavaruutta, jossa normina on euklidinen normi.

Jos kuvaus on differentioituva pisteessä $\mathbf{u}_0 \in \mathscr{U}$, niin kuvaus on myös jatkuva tässä pisteessä. Lineaarista operaattoria $\mathrm{D}\mathbf{h}(\mathbf{u}_0) \in \mathscr{L}(\boldsymbol{E},\boldsymbol{F})$ sanotaan Fréchet-derivaataksi, joka on lineaarinen äärellisen muutoksen $\Delta\mathbf{u}$ suhteen ja on olemassa myös pisteen $\mathbf{u}_0$ ympäristössä. Operaattoria $\mathrm{D}\mathbf{h}(\mathbf{u}_0)\cdot\mathbf{du}$, missä $\mathbf{du}$ on differentiaalinen muutos, sanotaan Fréchet-differentiaalioperaattoriksi, joka tavallisesti merkitään $\mathbf{dh}(\mathbf{u};\mathbf{du})$ tai yksinkertaisemmin $\mathbf{dh}$.

Mekaniikassa on järkevää eritellä muuttujat asettamalla $\mathbf{u} = (\mathbf{x}, \mathbf{v}, t) \in \boldsymbol{X} \times \boldsymbol{V} \times \mathbb{R}$, missä $\mathbf{x} \in \boldsymbol{X}$ edustaa asemamuuttujaa, $\mathbf{v} \in \boldsymbol{V}$ nopeusmuuttujaa ja $t \in \mathbb{R}$ aikaa. Tällöin saadaan kuvaukseksi

$$\mathrm{D}\mathbf{h}(\mathbf{u}) \cdot \Delta\mathbf{u} = \mathrm{D}_{\mathbf{x}}\mathbf{h}(\mathbf{x}, \mathbf{v}, t) \cdot \Delta\mathbf{x} + \mathrm{D}_{\mathbf{v}}\mathbf{h}(\mathbf{x}, \mathbf{v}, t) \cdot \Delta\mathbf{v} + \mathrm{D}_{t}\mathbf{h}(\mathbf{x}, \mathbf{v}, t) \cdot \Delta t, \tag{2}$$

missä merkinnät $\mathrm{D}_{\mathbf{x}}, \mathrm{D}_{\mathbf{v}}, \mathrm{D}_{t}$ edustavat Fréchet-osittaisderivaattoja.

Määritelmä: Mekaanisen järjestelmän sidosta sanotaan holonomiseksi, jos sidosehto $\mathbf{g}: \mathscr{U} \subset \boldsymbol{X} \times \mathbb{R} \to \boldsymbol{G}_1 \subset \boldsymbol{X}$ voidaan esittää (äärellisessä) muodossa

$$\mathbf{g}(\mathbf{x}, t) = \mathbf{0}. \tag{3}$$

Joukko $\mathscr{U}$ on euklidisen avaruuden $\boldsymbol{X}$ avoin joukko, missä sidosehto on määritelty. Vaaditaan lisäksi, että kuvaus $\mathbf{x} \times t \mapsto \mathbf{g}(\mathbf{x}, t)$ on jatkuva ja sen derivaatta on jatkuva, ts. $\mathbf{g} \in \mathscr{C}^1$. Vastaavasti, jos sidosta ei voida esittää yllä olevassa muodossa, sidosehtoa sanotaan epäholonomiseksi. Eräs epäholonominen sidostyyppi voidaan esittää differentiaalisessa eli Pfaffin muodossa (1-muodossa) seuraavasti

$$\mathbf{B}(\mathbf{x}, t) \cdot \mathrm{d}\mathbf{x} + \mathbf{b}(\mathbf{x}, t) \cdot \mathrm{d}t = \mathbf{0} \tag{4}$$

missä operaattoreilta $\mathbf{B}: \mathscr{U} \subset \boldsymbol{X} \times \mathbb{R} \to \mathscr{L}(\boldsymbol{X}, \boldsymbol{G}_2)$ ja $\mathbf{b}: \mathscr{U} \subset \boldsymbol{X} \times \mathbb{R} \to \mathscr{L}(\mathbb{R}, \boldsymbol{G}_2)$ vaaditaan jatkuvuus muuttujien $\mathbf{x}$ ja t suhteen. Tämä vaatimus rajaa impulsiiviset sidosvoimat tarkastelun ulkopuolelle. Myöskin kaikki holonomiset sidokset voidaan esittää Pfaffin muodossa, sillä sidosehdon (3) differentiaali on

$$\mathrm{d}\mathbf{g}(\mathbf{x}, t) = \mathrm{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x}, t) \cdot \mathrm{d}\mathbf{x} + \mathrm{D}_{t}\,\mathbf{g}(\mathbf{x}, t) \cdot \mathrm{d}t = \mathbf{0}, \tag{5}$$

missä $\mathrm{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x}, t) \in \mathscr{L}(\boldsymbol{X}, \boldsymbol{G}_2)$ ja $\mathrm{D}_{t}\,\mathbf{g}(\mathbf{x}, t) \in \mathscr{L}(\mathbb{R}, \boldsymbol{G}_2)$. Käänteinen ei välttämättä päde, koska kaikkia Pfaffin muotoisia sidoksia ei voida lausua äärellisessä, holonomisessa muodossa (3), [Rosenberg, 1980, s. 45-48].

Jatkossa oletetaan, että lineaarinen operaattori $\mathbf{D}: \mathscr{U} \subset \boldsymbol{X} \times \mathbb{R} \to \mathscr{L}(\boldsymbol{X}, \boldsymbol{G}_1 \times \boldsymbol{G}_2)$

$$\mathbf{D}(\mathbf{x}, t_0) = \begin{pmatrix} \mathrm{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x}, t_0) \\ \mathbf{B}(\mathbf{x}, t) \end{pmatrix} \tag{6}$$

on surjektiivinen, jolloin operaattorin matriisiesityksen rangi on täysi.

Holonomiset sidokset muodostavat euklidiseen asema-avaruuteen $\boldsymbol{X}$ Riemannin moniston (*Riemannian manifold*) $\mathscr{M}$, joka voidaan määritellä

$$\mathscr{M} = \left\{ \mathbf{x} \times t \in \mathscr{U} \subset \boldsymbol{X} \times \mathbb{R} \,\middle|\, \mathbf{g}(\mathbf{x}, t) = \mathbf{0} \right\}, \tag{7}$$

missä aika t voidaan ajatella olevan moniston $\mathscr{M}$ parametri. Kiinnittämällä aika ja asettamalla $t = t_0 \in \mathbb{R}$ voidaan lausua holonomisten sidosehtojen muodostama monisto $\mathscr{M}_{t_0}$ ajan hetkellä t_0 seuraavasti

$$\mathscr{M}_{t_0} = \left\{ \mathbf{x} \in \mathscr{U} \subset \boldsymbol{X} \,\middle|\, \mathbf{g}(\mathbf{x}, t_0) = \mathbf{0} \right\} \tag{8}$$

Variaation määritelmä: Kuvauksen $\mathbf{h}:\mathscr{X} \subset X \times \mathscr{V} \subset V \to F$ (Fréchet) variaatio $\delta\mathbf{h}$ pisteessä $(\mathbf{x},\mathbf{v}) \in \mathscr{X} \times \mathscr{V}$ voidaan määritellä kaavan (2) erikoistapauksena

$$\delta\mathbf{h}(\mathbf{x},\mathbf{v},t_0) = \mathrm{D}_{\mathbf{x}}\mathbf{h}(\mathbf{x},\mathbf{v},t_0)\cdot\delta\mathbf{x} + \mathrm{D}_{\mathbf{v}}\mathbf{h}(\mathbf{x},\mathbf{v},t_0)\cdot\delta\mathbf{v}, \tag{9}$$

missä aseman $\mathbf{x}$ ja (äärellisen) virtuaalisen siirtymän $\delta\mathbf{x}$ muodostama pari $(\mathbf{x},\delta\mathbf{x})$ sidosehdoilla (3) ja (4) kuuluu vektorikimppuun $\mathscr{W}$ (*vector bundle*), joka määritellään

$$\mathscr{W} = \left\{ \mathbf{x},\delta\mathbf{x} \in X \,\middle|\, \mathbf{x} \in \mathscr{M}_{t_0}, \begin{pmatrix} \mathrm{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x},t_0) \\ \mathbf{B}(\mathbf{x},t_0) \end{pmatrix} \cdot \delta\mathbf{x} = \mathbf{0} \right\}. \tag{10}$$

Virtuaalista siirtymää voidaan kutsua myös lumesiirtymäksi. Vektorikimppu on säikeinen monisto, jonka jokaiseen pisteeseen voidaan liittää vektoriavaruus. Variaatiossa aikaa pidetään kiinnitettynä asettamalla $t = t_0$. Fréchet-derivaatan lineaarisuudesta johtuen myös variaatio-operaattori $\delta: F \to F$ on lineaarinen, ts. $\delta \in \mathscr{L}(F,F)$.

Äärelliselle virtuaalinen nopeudelle $\delta\mathbf{v}$ ei yhtälöiden (3) ja (4) kaltaisilla sidoksilla aseteta ehtoja, mutta virtuaalisen nopeuden on kuitenkin toteutettava differentiaaliyhtälö

$$\delta\left(\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}\right) = \delta\mathbf{v}. \tag{11}$$

Merkitsemällä $\mathbf{h}(\mathbf{x},\mathbf{v},t_0) = \mathbf{x}(t_0)$ nähdään, että paikan $\mathbf{x}$ variaatio on yhtä suuri kuin virtuaalinen siirtymä, ts. $\delta\mathbf{x} = \delta\mathbf{x}$. Samoin nopeuden variaatio on yhtä suuri kuin virtuaalinen nopeus, ts. $\delta\mathbf{v} = \delta\mathbf{v}$. Variaatio ja aikaderivaatta eivät yleisesti ottaen ole vaihdannaisia, sillä esimerkiksi olkoon $\dot{\mathbf{x}} = t\dot{\mathbf{y}}$, jonka variaatio on $\delta\dot{\mathbf{x}} = t\delta\dot{\mathbf{y}}$. Toisaalta virtuaalisen siirtymän $\delta\mathbf{x} = t\delta\mathbf{y} \Leftrightarrow \delta\mathbf{x} = t\delta\mathbf{y}$ aikaderivaatta on

$$\frac{\mathrm{d}(\delta\mathbf{x})}{\mathrm{d}t} = t\frac{\mathrm{d}(\delta\mathbf{y})}{\mathrm{d}t} + \delta\mathbf{y}, \tag{12}$$

joka on selvästi eri suuri kuin $\delta\dot{\mathbf{x}} = t\delta\dot{\mathbf{y}}$. Sidosehtohan $\dot{\mathbf{x}} = t\dot{\mathbf{y}}$ on epäholonominen, joka ei muodosta differentoituvaa monistoa. Tämä hieman yllättäväkin tulos on huomattu ainakin kirjassa [Burge, 1996, s. 314]. Holonomisilla sidoksilla, jotka siis muodostavat differentoituvan moniston, variaatio ja aikaderivaatta ovat kuitenkin vaihdannaisia, ainakin muodollisesti.

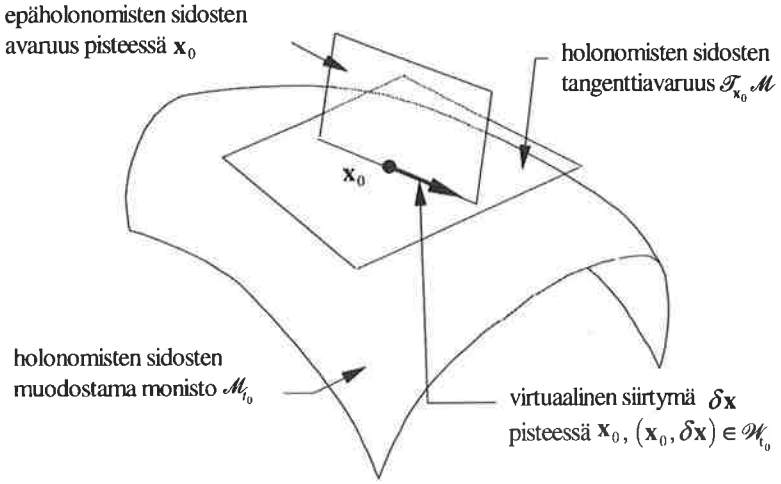
Mikäli järjestelmässä on vain holonomisia sidoksia (3), niin säikeinen monisto $\mathscr{W}$ yksinkertaistuu tangenttikimpuksi (*tangent bundle*) $\mathscr{TM}$

$$\mathscr{TM} = \left\{ \mathbf{x},\delta\mathbf{x} \in X \,\middle|\, \mathbf{x} \in \mathscr{M}_{t_0}, \mathrm{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x},t_0)\cdot\delta\mathbf{x} = \mathbf{0} \right\} \tag{13}$$

Tangenttikimppu on säikeinen monisto, jonka jokaiseen pisteeseen voidaan liittää tangenttiavaruus. Edellä on oletettu, että ajan hetkeä pidetään kiinnitettynä asettamalla $t = t_0$. Samoin tietyn pisteen asemaa $\mathbf{x}_0 = \mathbf{x}(t_0)$ voidaan pitää kiinnitettynä, jolloin tangenttikimppu kutistuu tangenttiavaruudeksi

$$\mathscr{T}_{\mathbf{x}_0}\mathscr{M} = \left\{ \delta\mathbf{x} \in X \,\middle|\, \mathrm{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x} = \mathbf{x}_0, t_0)\cdot\delta\mathbf{x} = \mathbf{0} \right\} \tag{14}$$

Säikeisen moniston nimitys tangenttiavaruudeksi on perusteltua, sillä tangenttiavaruus $\mathscr{T}_{\mathbf{x}_0}\mathscr{M}$ missä tahansa moniston $\mathscr{M}_{t_0}$ pisteessä $\mathbf{x}_0 \in \mathscr{M}_{t_0}$ on yhtä suuri kuin derivaatan $\mathrm{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x},t_0)$ nolla-avaruus, ts. $\ker \mathrm{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x},t_0)$ tässä pisteessä. Kuvassa 1 on esitetty geometrinen tulkinta virtuaaliselle siirtymälle.



Kuva 1. Virtuaalisen siirtymän geometrinen tulkinta

MATEMAATTISEN LAGRANGEN MEKANIIKAN PERUSLAIT

Matemaattisen Lagrangen mekaniikan perusoletukset voidaan asettaa täsmälleen samoiksi kuin matemaattisen Newtonin mekaniikan kuusi perusoletusta (**O1-O6**) vain perusmääritelmissä on hieman eroja. Peruslähtökohta Lagrangen mekaniikassa on eritellä voimat sidosvoimiin $\mathbf{F}^{\mathrm{s}}$ ja voimiin, jotka eivät ole sidosvoimia, joita kutsutaan annetuksi voimiksi $\mathbf{F}^{\mathrm{a}}$. Eri voimatyypit voidaan tunnistaa virtuaalisen työn eli lumetyön avulla.

Virtuaalisen työn määritelmä: Mekaanisen järjestelmän virtuaalisella työllä tarkoitetaan lineaarista funktionaalia δW, jonka määrittelyalue on vektorikimpun $\mathscr{W}$ toimintapisteen $\mathbf{x}_0 = \mathbf{x}(t_0) \in \mathscr{M}_{t_0}$ vektoriavaruus $\mathscr{W}_{\mathbf{x}_0}$

$$\mathscr{W}_{\mathbf{x}_0} = \left\{ \delta\mathbf{x} \in X \middle| \begin{pmatrix} \mathrm{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x}=\mathbf{x}_0, t_0) \\ \mathbf{B}(\mathbf{x}_0, t_0) \end{pmatrix} \cdot \delta\mathbf{x} = \mathbf{0} \right\}. \tag{15}$$

Virtuaalinen työ on siis lineaarinen kuvaus $\delta W: \mathscr{W}_{\mathbf{x}_0} \to \mathbb{R}$. Lineaaristen funktionaalien eli kuvausten $\mathscr{L}(\mathscr{W}_{\mathbf{x}_0}, \mathbb{R})$ muodostamaa avaruutta merkitään tavallisesti $\mathscr{W}^*_{\mathbf{x}_0}$. Koska siirtymät ja voimat muodostavat työkonjugaatin, niin virtuaalinen työ voidaan esittää duaalisena parina

$$\delta W = \mathbf{F}^{\mathrm{a}} \cdot \delta\mathbf{x}, \tag{16}$$

missä virtuaalisen siirtymän $\delta\mathbf{x} \in \mathscr{W}_{\mathbf{x}_0}$ duaalinen pari on voima $\mathbf{F}^{\mathrm{a}} \in \mathscr{W}^*_{\mathbf{x}_0}$, jota kutsutaan annetuksi voimaksi.

Sidosvoimien määritelmä: Sidosvoimat $\mathbf{F}^s$ ovat voimia, joiden virtuaalinen työ on nolla kaikilla virtuaalisilla siirtymillä $\delta\mathbf{x} \in \mathscr{W}_{\mathbf{x}_0}$, ts.

$$\delta W = \mathbf{F}^s \cdot \delta\mathbf{x} = 0, \qquad \forall \delta\mathbf{x} \in \mathscr{W}_{\mathbf{x}_0} \tag{17}$$

tällöin $\mathbf{F}^s \in \mathscr{W}_{\mathbf{x}_0}^{*\perp}$. Vaatimusta voidaan myös pitää ortogonaalisuusehtona.

Annettujen voimien määritelmä: Voimat, jotka kuuluvat sidosvoima-avaruuden ortogonaalikomplementtiin $\mathscr{W}_{\mathbf{x}_0}^*$, ovat annettuja voimia $\mathbf{F}^a$. Annettu voima $\mathbf{F}^a$ on duaalinen pari virtuaaliselle siirtymälle $\delta\mathbf{x} \in \mathscr{W}_{\mathbf{x}_0}$. Vektoriavaruudet $\mathscr{W}_{\mathbf{x}_0}^{*\perp}$ ja $\mathscr{W}_{\mathbf{x}_0}^*$ muodostavat jaon voima-avaruudelle, jota voidaan merkitä $\boldsymbol{F} = \mathscr{W}_{\mathbf{x}_0}^* \oplus \mathscr{W}_{\mathbf{x}_0}^{*\perp}$ ja $\mathbf{F}^a + \mathbf{F}^s \in \boldsymbol{F}$.

Matemaattisen Lagrangen mekaniikan perusoletukset voidaan asettaa täsmälleen samoiksi kuin matemaattisen Newtonin mekaniikan perusoletukset (**O1-O6**) ja perusmääritelmät eroavat vain määritelmän (**M3**) osalta:

M3a. Virtuaalisen työn määritelmä
M3b. Sidosvoimien määritelmä
M3c. Annettujen voimien määritelmä

Lagrangen mekaniikassa on huomattavasti enemmän matemaattista rakennetta kuin Newtonin mekaniikassa, sillä Lagrangen mekaniikkaan on sisällytettävä jonkin verran differentiaaligeometriaa ja analyysia.

Tarkastellaan seuraavassa matemaattisen Newtonin mekaniikan hiukkasjärjestelmän ongelmaa

$$\mathbf{F} - \mathbf{M} \cdot \mathbf{a} = \mathbf{0}, \tag{18}$$

missä $\mathbf{F} \in \mathbb{F}^n \equiv \left(\mathbb{E}^n\right)^*$ on hiukkasjärjestelmään vaikuttava voimavektori, $\mathbf{M} \in \mathscr{L}(\mathbb{E}^n, \mathbb{F}^n)$ on hitausoperaattori (massamatriisi), $\mathbf{a} \in \mathbb{E}^n$ on kiihtyvyysvektori ($\mathbf{a} = \ddot{\mathbf{x}}$) ja $\mathbb{E}^n$ on n-ulotteinen euklidinen paikka-avaruus sekä $\mathbb{F}^n$ vastaava voima-avaruus. Lisäksi vaaditaan, että siirtymät toteuttavat holonomiset sidosehdot $\mathbf{g}: \mathbb{E}^n \to \mathbb{E}^m$

$$\mathbf{g}(\mathbf{x}) = \mathbf{0} \tag{19}$$

missä $m < n$. Yhtälöt (18) ja (19) muodostavat ylimäärätyn yhtälöryhmän, jossa tuntemattomia on n kappaletta ja yhtälöitä on $m+n$ kappaletta.
Lagrangen mekaniikassa voima $\mathbf{F}$ voidaan jakaa sidosvoimaan ja annettuun voimaa, ts. $\mathbf{F} = \mathbf{F}^s + \mathbf{F}^a$. Laskemalla yhtälön (18) virtuaalinen työ ja huomioimalla, että sidosvoimien virtuaalinen työ on nolla, saadaan tulos

$$\delta W = \left(\mathbf{F}^a - \mathbf{M} \cdot \mathbf{a}\right) \cdot \delta\mathbf{x} = 0 \tag{20}$$

missä annetulla sidosehdolla (19) virtuaalinen siirtymä kuuluu sidosmoniston $\mathscr{M}$ tangenttiavaruuteen $\delta\mathbf{x} \in \mathscr{T}_{\mathbf{x}_0}\mathscr{M}$. Sidosmonisto on tässä tapauksessa yksinkertaisesti

$$\mathscr{M} = \left\{\mathbf{x} \in \mathbb{E}^n \middle| \, \mathbf{g}(\mathbf{x}) = \mathbf{0}\right\}, \tag{21}$$

jonka tangenttiavaruus toimintapisteessä $\mathbf{x}_0 = \mathbf{x}(t_0)$ on vastaavasti

$$\mathscr{T}_{\mathbf{x}_0}\mathscr{M} = \left\{\delta\mathbf{x} \in \mathbb{E}^n \middle|\, \mathrm{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x}=\mathbf{x}_0)\cdot\delta\mathbf{x} = \mathbf{0}\right\} \tag{22}$$

Tavallisesti yhtälöt (19-22) esitetään muodossa:

$$\begin{aligned} &\left(\mathbf{F}^{\mathrm{a}} - \mathbf{M}\cdot\mathbf{a}\right)\cdot\delta\mathbf{x} = 0, \qquad \text{missä virtuaalinen siirtymä } \delta\mathbf{x} \text{ toteuttaa ehdon} \\ &\mathrm{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x}=\mathbf{x}_0)\cdot\delta\mathbf{x} = 0 \qquad \text{ja asema } \mathbf{x} \\ &\mathbf{g}(\mathbf{x}) = \mathbf{0} \end{aligned} \tag{23}$$

Lagrangen kertoimen menettelyllä yhtälöt (19-22) ja vastaavasti (23) ovat samanarvoisia (ekvivalentteja) yhtälöiden [Abraham ja muut, 1983; Rosenberg, 1980]

$$\begin{aligned} &\left(\mathbf{F}^{\mathrm{a}} - \mathbf{M}\cdot\mathbf{a} - \boldsymbol{\lambda}\circ\mathrm{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x}=\mathbf{x}_0)\right)\cdot\delta\mathbf{x} = 0 \qquad \forall\delta\mathbf{x}\in\mathbb{E}^n \\ &\mathbf{g}(\mathbf{x}) = \mathbf{0} \end{aligned} \tag{24}$$

kanssa, missä Lagrangen kerroinvektori $\boldsymbol{\lambda} \in \mathscr{L}(\mathbb{E}^m, \mathbb{R}) = \mathbb{F}^m$ ja $\mathrm{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x}=\mathbf{x}_0) \in \mathscr{L}(\mathbb{E}^n, \mathbb{E}^m)$. Tällöin yhdistetty kuvaus $\boldsymbol{\lambda}\circ\mathrm{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x}=\mathbf{x}_0) \in \mathscr{L}(\mathbb{E}^n, \mathbb{R}) = \mathbb{F}^n$. Yhtälöä (24) sanotaan d'Alembertin periaatteen Lagrangen muodoksi [Rosenberg, 1980]. D'Alembertin periaate on keskeisessä osassa analyyttisessä Lagrangen mekaniikassa, jossa tätä periaatetta käyttäen voidaan johtaa muut Lagrangen mekaniikan periaatteet. Periaate-sanan käyttö tässä yhteydessä on varsin kyseenalaista, sillä d'Alembertin periaate johdettiin puhtaasti määrittelypohjalta käyttäen virtuaalisen työn ja sidosvoimien määritelmiä.

Koska yhtälössä (24) virtuaalinen siirtymä $\delta\mathbf{x}$ on täysin vapaa, niin yhtälö (24) voidaan kirjoittaa myös muodossa

$$\begin{aligned} &\mathbf{F}^{\mathrm{a}} - \mathbf{M}\cdot\mathbf{a} - \left(\mathrm{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x}=\mathbf{x}_0)\right)^{\mathrm{T}}\cdot\boldsymbol{\lambda} = \mathbf{0} \\ &\mathbf{g}(\mathbf{x}) = \mathbf{0} \end{aligned} \tag{25}$$

missä transpoosi $\mathrm{D}_{\mathbf{x}}\mathbf{g}(\mathbf{x}=\mathbf{x}_0)^{\mathrm{T}} \in \mathscr{L}(\mathbb{F}^m, \mathbb{F}^n)$. Vertaamalla yhtälöitä (18) ja (25) nähdään, että sidosvoima $\mathbf{F}^{\mathrm{s}} = -\left(\mathrm{D}_{\mathbf{x}}\mathbf{g}\right)^{\mathrm{T}}\cdot\boldsymbol{\lambda} \in \mathscr{T}^*_{\mathbf{x}_0}\mathscr{M}$, ts. sidosvoimat voidaan lausua sidosehdon derivaatan vaakarivien lineaariyhdistelmänä. Analyyttisen mekaniikan kirjallisuudessa [Rosenberg, 1980; Greenwood, 1977] ns. yleistetyillä koordinaateilla on keskeinen osa. Yleistetty koordinaatisto on itse asiassa sidosmoniston parametrisointi, jonka paikallisen olemassaolon implisiittifunktiolause takaa [Abraham ja muut, 1983, s. 107].

Yleistetty koordinaatisto sidosehdon (19) tapauksessa on kuvaus $\boldsymbol{\varphi}(\mathbf{q}):\mathbb{E}^{n-m} \to \mathscr{M}$, joten sidosehto $\mathbf{g}(\boldsymbol{\varphi}(\mathbf{q})) = \mathbf{0}$ toteutuu automaattisesti. Suorittamalla muuttujan vaihto $\mathbf{x} = \boldsymbol{\varphi}(\mathbf{q})$ voidaan yhtälöt (18) ja (19) muuttaa ns. tilayhtälömuodoksi, jossa ei ole lainkaan sidosehtoja.

YHTEENVETO

Tässä työssä on käsitelty ja koottu yhteen fysikaalisen ja matemaattisen Newtonin mekaniikan oppirakenteen peruslait sekä matemaattisen Lagrangen mekaniikan oppirakenteen peruslait. Matemaattisessa Newtonin ja Lagrangen mekaniikan esittelyssä on eritelty peruslait perusoletuksiin sekä perusmääritelmiin. Keskeisin perusmääritelmä on voiman määritelmä eli Newtonin II laki. Matemaattiset Newtonin ja Lagrangen mekaniikat eroavat vain perusmääritelmiensä suhteen, lisäksi Lagrangen mekaniikkaan sisällytetään jonkin verran differentiaaligeometriaa.

LÄHTEET

Abraham, R., Marsden, J. E., Ratiu, T., (1983), *Manifolds, Tensor Analysis and Applications*, Addison-Wesley, 582 s..
Arnold, V. I., (1978), *Mathematical Methods of Classical Mechanics*, Springer, 462 s..
Beer, F.P., Johnston, E.R., (1976), *Mechanics for Engineers, Statics*, McGraw-Hill, 3. p., 410 s..
Burke, W.L., (1996), *Applied Differential Geometry*, Cambridge University Press, 414 s..
Desloge, E. A., (1982), *Classical Mechanics, volume 1*, John Wiley & Sons, 507 s..
Goldstein, H., (1990), *Classical Mechanics*, Addison-Wesley, 2. toim., 672 s..
Greenwood, D. T., (1977), *Classical Dynamics*, Prentice-Hall, 337 s..
Keskinen, R., (1993), *Analyyttinen mekaniikka*, Limes ry, 121 s..
Kurki-Suonio, K., Kurki-Suonio, R., (1990), *Vuorovaikuttavat kappaleet -- mekaniikan perusteet*, Limes ry, 450 p..
Marsden, J. E., Hughes ,T. J. R., (1983), *Mathematical Foundations of Elasticity*, Prentice-Hall, 556 s..
Oden, J.T., Reddy, J.N., (1976), *Variational Methods in Theoretical Mechanics*, Springer, 302 s..
Pars, L. A., (1965), *A Treatise on Analytical Dynamics*, Heinemann, 641 s..
Reddy, J.N., (1987), *Applied Functional Analysis and Variational Methods in Engineering*, McGraw-Hill, 2. p., 546 s..
Rheinboldt, W. C., (1986), *Numerical Analysis of Parametrized Nonlinear Equations*, John Wiley & Sons, 2. p., 299 s..
Rosenberg, R. M., (1980), *Analytical Dynamics of Discrete Systems*, Plenum, 2. p., 424 s..
Sagan , H., (1969), *Introduction to the Calculus of Variations*, McGraw-Hill, 449 s..
Salmi, T., (1996), *Dynamiikka 1, Kinematiikka*, Pressus, 4. p., 226 s..
Salmi, T., (1997), *Dynamiikka 2, Kinetiikka*, Pressus, 2. p., 312 s..
Salmi, T., (1998), *Statiikka*, Pressus, 388 s..
Salonen, E.-M., (1989), *Dynamiikka*, Otakustantamo, 3 p., 458 s..
Santilli, R. M., (1978), *Foundations of Theoretical Mechanics I: The Inverse Problem in Newtonian Mechanics*, Springer, 266 s..
Spiegell, M. R., (1980), *Schaum's outline of theory and problems of theoretical mechanics with an introduction to Lagrange's equation and Hamiltonian theory*, McGraw-Hill, SI toim., 368 s..
Truesdell, C., (1977), *A First Cource in Rational Continuum Mechanics*, Academic Press, 280 s..
Wells, D A., (1967), *Schaum's outline of theory and problems of dynamics : with a treatment of Euler's equations of motion, Hamilton's equations and Hamilton's principle*, McGraw-Hill, 353 s..
Whittaker, E. T., (1989), *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*, Cambridge, 4. toim. 2. p., 456 s..

Numerical Aspects of Modular Modeling using Lagrangian DAEs

Robert Piché and Mikko Palmroth
Tampere University of Technology
Department of Mathematics

April 19, 2000

Abstract

A modular approach to the Lagrangian DAE method for unified modeling of physical systems is described. Numerical solution aspects are discussed. One example is given: a controlled mechanical system.

1 Background

1.1 Lagrangian DAE modeling

The Lagrangian DAE [1] (LDAE) is a systematic and unified method for mathematical modeling of lumped parameter engineering systems that involve coupled interactions of electrical, mechanical, fluid, and thermal domains. In a Lagrangian DAE model there are $2n + m_1 + m_2 + m_3 + m_4$ configuration variables, as follows:

- displacement variables $q_{1\ldots n}$
- flow variables $f_{1\ldots n}$, $\dot{q} = f$
- displacement constraint Lagrange multipliers $\kappa_{1\ldots m_1}$
- flow constraint Lagrange multipliers $\mu_{1\ldots m_2}$
- implicit efforts ${e^\gamma}_{1\ldots m_3}$
- dynamic variables $s_{1\ldots m_4}$

Displacement and flow variables are called kinematic variables. Kinematic variables and efforts have different interpretations in various physical domains (Table 1).

A Lagrangian system model consists of the following functions:

- potential energy $V(q,t)$

Domain	Displacement q	Flow $f = \dot{q}$	Effort e
mechanical translation	position	velocity	force
mechanical rotation	angle	angular velocity	torque
electrical	charge	current	voltage
incompressible fluid	fluid volume	volumetric flow rate	pressure

Table 1: Unified notation for variables in various physical domains

- kinetic coenergy $T^*(f, q, t)$
- content (Rayleigh potential) $D(f, q, t)$
- applied efforts $[Q_{1\ldots n}] = Q(e^\gamma, t)$
- displacement constraint functions $[\phi_{1\ldots m_1}] = \Phi(q, t)$
- flow constraint functions $[\psi_{1\ldots m_2}] = \Psi(f, q, t)$ that are linear in f
- implicit effort constraint functions $[\gamma_{1\ldots m_3}] = \Gamma(e^\gamma, s, f, q, t)$
- dynamic variable derivative functions $[\lambda_{1\ldots m_4}] = \Lambda(e^\gamma, s, f, q, t)$

The Lagrangian DAE is generated from the system model according to the following formula:

$$\begin{aligned} \dot{q} &= f \\ M\dot{f} + \Phi_q^T \kappa + \Psi_f^T \mu &= \Upsilon \\ \Phi &= 0 \\ \Psi &= 0 \\ \Gamma &= 0 \\ \dot{s} &= \Lambda \end{aligned}$$

Here

$$\begin{aligned} \Upsilon &= Q - (\nabla_f T^*)_q f - (\nabla_f T^*)_t \\ &\quad + \nabla_q T^* - \nabla_q V - \nabla_f D \\ M &= \nabla_f^2 T^* \end{aligned}$$

This set of equations can be summarised as

$$\mathcal{M}(q, f) = 0$$

2 Modular Modeling

In a modeling method, one should be able to be able to build up models from simple reusable components. This has many advantages: libraries of standard components can be developed by experts and used by others; complex systems can be assembled in a hierarchical fashion, etc. It is our intention to describe how this can be achieved using Lagrangian DAEs.

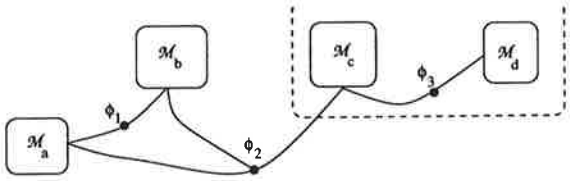


Figure 1: Submodels coupled using kinematic constraints. The dashed box is a submodel that is itself a coupling of submodels.

2.1 Submodels and Connections

To date, the Lagrangian DAE method has not been used in a modular way: problems have always been modelled from scratch. This is not, however, because of innate restrictions: the method already contains the structures to allow a modular approach to modeling. The idea is simple: couple submodels using kinematic constraints.

To elaborate, consider a set of separate system components, each with its own Lagrangian DAE:

$$\begin{aligned} \mathcal{M}_1(q_1, f_1) &= 0 \\ \mathcal{M}_2(q_2, f_2) &= 0 \\ \vdots &\quad \vdots \end{aligned}$$

We assume that the coupling of these submodels can be represented by additional kinematic constraints,

$$\begin{aligned} \Phi(q_1, q_2, \ldots) &= 0 \\ \Psi(q_1, q_2, \ldots, f_1, f_2, \ldots) &= 0 \end{aligned}$$

These constraints contribute additional terms of the form $\Phi_{q_i}^T \kappa + \Psi_{f_i}^T \mu$ to the equations of motion of submodel i. The augmented submodel equations, together with the kinematic constraints, represent the coupled system. This idea can of course be applied recursively, with submodels made up of simpler submodels (Figure 1).

To further illustrate how the coupling changes the submodel's equations, consider two simple Lagrangian components with no dynamic variables or implicit efforts and only flow constraints. If we couple these together with the flow constraint $\Psi_c(f_1, f_2, q_1, q_2, t) = 0$ the resulting equations look like

$$\begin{aligned} \dot{q}_1 &= f_1 \\ M_1 \dot{f}_1 + \Psi_{1f_1}^T \mu_1 + \Psi_{cf_1}^T \mu_c &= \Upsilon_1 \\ \Psi_1 &= 0 \\ \dot{q}_2 &= f_2 \\ M_2 \dot{f}_2 + \Psi_{2f_2}^T \mu_2 + \Psi_{cf_2}^T \mu_c &= \Upsilon_2 \\ \Psi_2 &= 0 \\ \Psi_c &= 0 \end{aligned}$$

In a large class of problems, the coupling of submodels is described entirely by flow constraints of the form $\sum \pm f_i = 0$. The Kirchhoff current law for circuits takes this form, and in mechanics this constraint is the time derivative of the displacement compatibility condition. For systems where this is the only constraint type, the modular modeling method presented here is closely related to the multipole method [2]. However, especially in mechanics, more complicated kinematic constraints are often needed.

It is not always the case that connections between different physical components can be represented by kinematic constraints. Other kind of coupling is also possible. For example it is possible that the potential energy of one component depends on the coordinate of another component, $V_1 = V_1(q_1, q_2)$. This kind of coupling can also be fitted inside our approach by adding an extra variable q_{12} to the first system and an additional kinematic constraint $q_{12} = q_2$. Now the potential energy $V_1 = V_1(q_1, q_{12})$ is only a function of one subsystem's variables. This is a bit cumbersome and the most convenient approach usually is to fit this kind of coupling inside one subsystem, but it does demonstrate that connections with kinematic constraints can handle a very wide range of cases.

2.2 Building a Simulator

The assembly of the Lagrangian DAEs from the submodel and constraint equations is usually done by hand, and is straightforward enough that it can be done automatically using a symbolic computing package, as was done by Layton [1]. With the modular approach suggested here, however, it should be possible to assemble models more efficiently.

The idea is to build a library of reusable software components for submodels and kinematic constraints. The submodels can themselves be built from simpler components that are connected with kinematic constraints. These components would contain all the information needed by a DAE-solver, such as the necessary symbolic derivatives, parameter values, initial values, variable indices, etc. With this information the assembly of the Lagrangian DAE for the entire system can be automated.

In order to use the software library the user would only have to choose suitable components from the library, connect them with kinematic constraints and provide parameter values and initial conditions. No algebraic manipulation is needed, and the user need not worry about the inner workings of the components and constraints. The modeller does not need to learn a new detailed uniform graphical modeling language to use modular Lagrangian DAEs. It is sufficient if diagrams label the constraints and the submodels, and this can be done using the kind of diagram that is commonly used in the appropriate engineering discipline.

The conceptual diagram shown in Figure 1 has an important difference from a standard block diagram: there are no arrows. That is, the links between submodels are not directed, and there are no 'inputs' and 'outputs'. Such concepts are rather artificial in physics, where coupling effects are usually mutual, not one-way. In a block diagram such as Simulink the physical coupling effects have to be represented by additional feedback loops. The resulting block diagram is cluttered and its topology is quite different from that of a free body diagram or of an electric circuit diagram. If one wishes to use block diagrams, e.g. for feedback controllers, this can be accommodated in the Lagrangian DAE using implicit efforts and dynamic variables.

2.3 Numerical Solution of the LDAE

The Lagrangian DAE closely resembles the Euler-Lagrange equations for constrained mechanical systems, but there is one crucial difference. Usually in constrained mechanics all flows are kinetic (associated with kinetic energy) and the matrix $M = \nabla_f^2 T^*$ is positive definite. This is not true with the LDAE, especially if it is built form components with the modular approach suggested here. Our approach has the drawback that it introduces a number of new variables, most of which are not kinetic (see section 3 for an example). In LDAEs matrix M is only positive semidefinite and usually has relatively low rank. This has the result that some algebraic variables (i.e. variables whose derivatives do not appear in the equations) now appear in the algebraic equations and the LDAE can no longer be easily transformed to the well understood Hessenberg form.

The Lagrangian DAE has the linearly implicit form

$$A(y,t)\dot{y} + f(y,t) = 0$$

The matrix M is usually diagonal or block-diagonal with small blocks, and the non-zero blocks are nonsingular (again, see the example). We can therefore easily solve for the derivatives of all kinematic flows and arrive at the semi-explicit form

$$\begin{aligned} \dot{y}_1 &= f_1(y_1, y_2, t) \\ 0 &= f_2(y_1, y_2, t) \end{aligned}$$

The Lagrangian DAEs encountered in practical problems seem to have differential index 1–3. In most holonomic pure mechanical systems all variables are kinematic and the differential index is always 3. Remember, the differential index was defined as the minimum number of times the DAE or parts of it have to be differentiated in order to algebraically solve for the derivatives of all variables.

However, if we allow non-physical (or ideal) components, Lagrangian DAEs with arbitrarily high index can be constructed. For example, ideal operational amplifiers, resistors and capasitors can be used to construct differentiators. By connecting these differentiators in series we can reach as high an index as we want. This example is rather contrived, and it is usually safe to think that LDAEs have at most index 3. The exact index of LDAEs (under suitable assumptions that rule out unphysical components) has yet to be proven.

For most index-3 problems the index can be reduced to 2 by simply replacing all displacement constraints $\Phi(q,t) = 0$ by the corresponding differentiated constraints $\Phi_q f + \Phi_t = 0$. This has the disadvantage that the numerical solution of the index-2 problem may no longer satisfy the original constraints. A better alternative is to include both the original and the differentiated constraints in the equations. This can be done by introducing extra "Lagrange"

multipliers λ:

$$\begin{aligned} \dot{q} &= f - \Phi_q^T \lambda \\ M\dot{f} + \Phi_q^T \kappa + \Psi_f^T \mu &= \Upsilon \\ \Phi &= 0 \\ \Phi_q f + \Phi_t &= 0 \\ \Psi &= 0 \\ \Gamma &= 0 \\ \dot{s} &= \Lambda \end{aligned} \tag{1}$$

This method was originally suggested by Gear, Gupta and Leimkuhler [3] for solving mechanical systems. It is easy to see that this has the same solutions as the original DAE, if the constraints Φ are independent. The modified DAE also has the same index as the DAE obtained from the original by differentiating the displacement constraints.

We are thus faced with solving a linearly implicit or semi-explicit DAE of at most index-2. Before we can apply standard numerical methods such as Runge-Kutta or BDF we still have to find consistent initial conditions. For index-one problems this is relatively easy. For the index-2 case things are more complicated, because the subset of the algebraic equations we have to differentiate to find the hidden constraints varies with the particular problem structure. So far we have used the algorithm for consistent initializing proposed by Schwarz and Lamour [4]. Based on the tractability index concept and some rather general assumptions on the DAE they describe a method for choosing a subset of the algebraic equations for differentiation, and obtaining full rank equations for consistent initial values.

3 Example

The remaining sections of this paper illustrate the modular Lagrangian DAE technique using an example from mechanics. The example shows how to work with a block diagram submodel.

3.1 The Solvers

For the solution of this example and other problems we have used three different solvers: MATLAB, DYNAST[5] and GODESS[6]. The Matlab DAE-solver `ode15s` is only able to handle index-1 problems and we have had to reduce the index of the high-index problems by differentiation and algebraic manipulation. Dynast is much more easy to use. It is able to handle index-2 problems and once the user is familiar with the syntax the DAEs can be solved with very little coding effort. Matlab and Dynast codes for this example and other simulations can be found at

`ftp://ftp.cc.tut.fi/pub/math/piche/ldae/orlando2000/`

These codes only solve the equations of the examples and no not exhibit the modularity concepts any further.

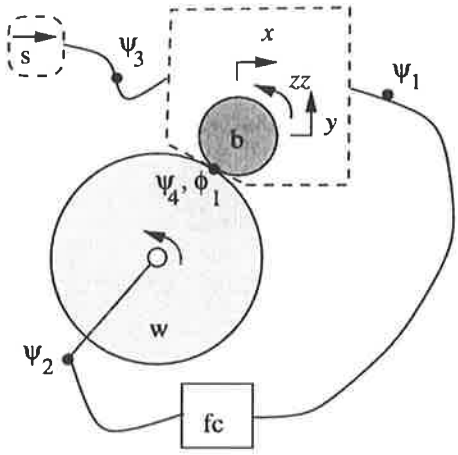


Figure 2: Ball on wheel

Godess is an object oriented DAE-solver written in C++ by Hans Olsson. Using C++ advanced data structures and classes we have implemented the modular models described here. In C++ the component and constraint models can be implemented as base-classes. From these classes it is easy to derive several types of components, such as ideal kinetic and potential stores, dissipators, effort sources, Kirchhoff-type constraints, flow sources etc. More complex structures can be built from these simple components and/or derived from the base classes. If the class-hierarchy is built carefully, adding new components to the library will not require too much effort.The example presented here has been solved with Godess. Index-reduction was done with the GGL-formulation [3] and consistent initial values were computed as in [4]. Naturally the different solvers give identical results.

3.2 A Controlled Mechanical System: Ball on Wheel

This example illustrates how a mechanical system can be coupled to an electro-mechanical feedback controller. The example also shows how Lagrangian submodels can be constructed from smaller components. The original example, without controller, is from [7].

3.2.1 Description

A solid homogeneous spherical ball b rolls without slipping on top of a cylindrical wheel w. The wheel rotates about its longitudinal axis. The ball is acted upon by gravity. A feedback controller fc applies a torque to the wheel. The feedback torque is controlled by a signal that is linearly proportional to the horizontal position and velocity of the ball (Figure 2). An impulsive horizontal force disturbance is applied to the ball.

The submodels in the system are the wheel w, ball b, feedback controller fc and the effort source s (which represents the disturbance force). These submodels are defined in the following sections.

3.2.2 Submodel: Wheel (w)

The wheel's parameters are the radius r_{w} and the moment of rotational inertia I_{w}. The flow variable is f_{w}, the rotation angle (positive counterclockwise) about the axle. The wheel is a kinetic store, with kinetic coenergy function

$$T^* = \frac{1}{2} I_{\mathrm{w}} f_{\mathrm{w}}^2$$

3.2.3 Submodel: Ball (b)

The ball's parameters are

$$\begin{aligned} r_{\mathrm{b}} &= \text{radius} \\ m_{\mathrm{b}} &= \text{mass} \\ I_{\mathrm{b}} &= \text{moment of rotational inertia} = \frac{2}{5} m_{\mathrm{b}} r_{\mathrm{b}}^2 \\ g &= \text{gravitational acceleration} \end{aligned}$$

There are three displacement variables,

$$\begin{aligned} q_{\mathrm{b},x} &= x\text{-position of centre} \\ q_{\mathrm{b},y} &= y\text{-position of centre} \\ q_{\mathrm{b},zz} &= \text{rotation angle about the centre, positive clockwise} \end{aligned}$$

The ball has both kinetic coenergy and potential energy. The kinetic coenergy is the sum of translational and rotational coenergies,

$$T^* = \frac{1}{2} m_{\mathrm{b}} (f_{\mathrm{b},x}^2 + f_{\mathrm{b},y}^2) + \frac{1}{2} I_{\mathrm{b}} f_{\mathrm{b},zz}^2$$

The potential energy due to the gravitational force, which acts in the negative y direction, is

$$V = m_{\mathrm{b}} g q_{\mathrm{b},y}$$

3.2.4 Submodel: Feedback Controller (fc)

The feedback controller, figure 3, is built from five smaller components that are connected with kinematic constraints.

The **sensor** has two flow variables, the input velocity $f_{\mathrm{se,in}}$ and the output current $f_{\mathrm{se,out}}$. It produces a voltage that is proportional to the input velocity:

$$e_{\mathrm{se,out}}^{\gamma} = a f_{\mathrm{se,in}}$$

This gives the effort constraint function

$$\gamma_{\mathrm{se,out}} := e_{\mathrm{se,out}}^{\gamma} - a f_{\mathrm{se,in}} = 0$$

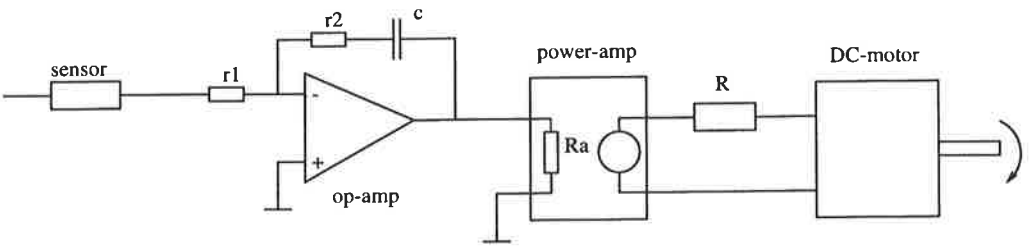


Figure 3: The Controller

The output voltage is an applied effort that apprears in the sensor equations

$$Q_{\text{se,out}} = e^{\gamma}_{\text{se,out}}$$

The only sensor parameter is the constant a.

The ideal **operational amplifier** configuration (that includes the two resistors r_1, r_2 and the capacitor c) produces an output voltage that is linearly proportional to the input voltage and it's integral. This component has two flow variables, the input and output currents $f_{\text{op,in}}$, $f_{\text{op,out}}$. The op-amp can be modeled with two effort-constraints that give the input and output voltages

$$\gamma_{\text{op,in}} := e^{\gamma}_{\text{op,in}} - r_1 f_{\text{op,in}} = 0$$
$$\gamma_{\text{op,out}} := e^{\gamma}_{\text{op,out}} + r_2 f_{\text{op,in}} + \frac{q_{\text{op,in}}}{c} = 0$$

The voltages contribute to the vector of applied efforts

$$Q_{\text{op,in}} = e^{\gamma}_{\text{op,in}}$$
$$Q_{\text{op,out}} = e^{\gamma}_{\text{op,out}}$$

The **power amplifier** takes the signal from the op-amp and runs the DC-motor. The amplifier also has two flow variables, $f_{\text{a,in}}$ and $f_{\text{a,out}}$. The input resistance is R_a and gain $-A$. The amplifier can also be modeled with two effort constraints and applied efforts

$$\gamma_{\text{a,in}} := e^{\gamma}_{\text{a,in}} - R_a f_{\text{a,in}} = 0$$
$$\gamma_{\text{a,out}} := e^{\gamma}_{\text{a,out}} + A e^{\gamma}_{\text{a,in}} = 0$$
$$Q_{\text{a,in}} = e^{\gamma}_{\text{a,in}}$$
$$Q_{\text{a,out}} = e^{\gamma}_{\text{a,out}}$$

The **load resistor** R is an ideal dissipator, with only one flow variable f_r and content

$$D = \frac{1}{2} R f_r^2$$

The **DC motor** is assumed to produce a torque that is proportional to the input current $f_{\text{dc,in}}$ and back-emf that is proportional to the angular velocity $f_{\text{dc,out}}$. This gives the effort

constraints and applied efforts

$$\gamma_{\mathrm{dc,in}} := e^{\gamma}_{\mathrm{dc,in}} - K f_{\mathrm{dc,out}} = 0$$
$$\gamma_{\mathrm{dc,out}} := e^{\gamma}_{\mathrm{dc,out}} - K f_{\mathrm{dc,in}} = 0$$
$$Q_{\mathrm{dc,in}} = e^{\gamma}_{\mathrm{dc,in}}$$
$$Q_{\mathrm{dc,out}} = e^{\gamma}_{\mathrm{dc,out}}$$

These five components inside the controller can be connected together with the flow constraints

$$\psi^c_1 := f_{\mathrm{op,in}} - f_{\mathrm{se,out}} = 0$$
$$\psi^c_2 := f_{\mathrm{a,in}} - f_{\mathrm{op,out}} = 0$$
$$\psi^c_3 := f_r - f_{\mathrm{a,out}} = 0$$
$$\psi^c_4 := f_{\mathrm{dc,in}} - f_r = 0$$

3.2.5 Submodel: Effort Source (s)

The effort function source $e_{\mathrm{s}}(t)$ is a generalised effort,

$$Q_{\mathrm{s}} = e_{\mathrm{s}}(t)$$

3.2.6 Coupling the Submodels using Kinematic Constraints

Kinematic variables are oriented as shown in Figure 2. The coordinate system origin is located at the wheel axle. There are five kinematic constraints, defined as follows.

The connection of the feedback control is modelled by the two flow constraints

$$\psi_1 := f_{\mathrm{se,in}} - f_{\mathrm{b},x} = 0$$
$$\psi_2 := f_{\mathrm{dc,out}} - f_{\mathrm{w}} = 0$$

The effort source is connected to the ball by the flow constraint

$$\psi_3 := f_{\mathrm{s}} - f_{\mathrm{b},x} = 0$$

The no-slip condition is given by the equation

$$r_{\mathrm{b}} f_{\mathrm{b},zz} + r_{\mathrm{w}} f_{\mathrm{w}} = (r_{\mathrm{b}} + r_{\mathrm{w}})\dot{\theta} \tag{2}$$

where θ is the angular position of the centre of the ball in polar coordinates. Differentiating the polar coordinate identities

$$q_{\mathrm{b},x} = (r_{\mathrm{b}} + r_{\mathrm{w}})\cos\theta,\ q_{\mathrm{b},y} = (r_{\mathrm{b}} + r_{\mathrm{w}})\sin\theta$$

gives

$$f_{\mathrm{b},x} + \dot{\theta} q_{\mathrm{b},y} = 0$$

which substituted into (2) gives the flow constraint

$$\psi_4 := r_{\mathrm{b}} f_{\mathrm{b},zz} q_{\mathrm{b},y} + r_{\mathrm{w}} f_{\mathrm{w}} q_{\mathrm{b},y} + (r_{\mathrm{b}} + r_{\mathrm{w}}) f_{\mathrm{b},x} = 0$$

The ball remains at a constant distance from the centre of the wheel. This is the displacement constraint

$$\phi_1 := q^2_{\mathrm{b},x} + q^2_{\mathrm{b},y} - (r_{\mathrm{b}} + r_{\mathrm{w}})^2 = 0$$

3.2.7 Simulation

We compute the response when an impulsive force is applied to the ball, which is initially at rest on top of the wheel ($q_{b,x} = 0$ and $q_{b,y} = 1.2$). The chosen numerical values are

$$
\begin{aligned}
g &= 9.81 \text{ m s}^{-2} & r_w &= 1 \text{ m} \\
r_b &= 0.2 \text{ m} & m_b &= 1 \text{ kg} \\
I_w &= 0.1 \text{ kg m}^2 & e_s(t) &= \begin{cases} 1 \text{ kN} & 0 < t \leq 0.01 \text{ s} \\ 0 & 0.01 \text{ s} < t \end{cases} \\
a &= 10 \text{ V/(m/s)} & r_1 &= 10 \text{ k}\Omega \\
r_2 &= 216 \text{ k}\Omega & c &= 741 \text{ nF} \\
A &= 1000 & R_a &= 1 \text{ M}\Omega \\
R &= 0.5 \text{ M}\Omega & K &= 50 \text{ Nm/A}
\end{aligned}
$$

The controller parameters have been chosen to simplify the control problem. A standard control design technique based on linearized equations has been employed to approximate the feedback control gains.

The numerical solution gives the results shown in Figure 4.

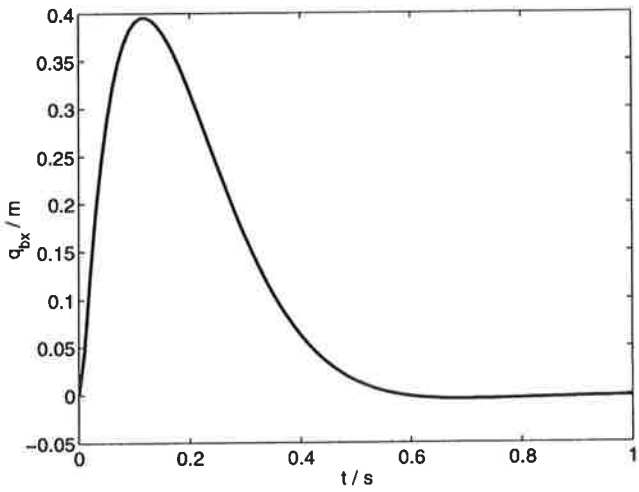


Figure 4: Simulation results for ball on wheel

4 Closing Remarks

We have described how Lagrangian DAE modeling can be extended to allow modularity. The modularity concepts presented are general enough that they could form a basis for software

for modelling and simulation of physical systems. The example illustrates how complex systems can be constructed from simple Lagrangian componenents with very little effort. The procedure does not require us to manipulate the equations for different components, we only need to specify kinematic constraints and the computer generates the appropriate Lagrange multipliers

Although most engineering systems can be decomposed into submodels connected by kinematic constraints, this approach is not always the most convenient way, and we do not try to impose it everywhere. In many systems the best alternative is still to use conventional ODE-modeling, or obtain a DAE-model using special properties of the particular system.

Nevertheless, we feel that the modular Lagrangian DAE technique is a viable alternative for modeling multidisciplinary systems. The technique will become more attractive in the future as numerical methods for DAEs develop and our understanding of numerical properties of Lagrangian DAEs increases.

References

[1] Richard A. Layton, *Principles of Analytical System Dynamics*, Springer, 1998.

[2] Herman Mann, Modular approach to Lagrange formalism for mechatronic systems. 3rd Int. Heinz Nixdorf Symp. on Mechatronics and Advanced Motion Control, pp.67-80, Paderborn 1999

[3] Gear, Gupta & Leimkuhler *Automatic integration of Euler-Lagrange equations with constraints*, J. Comp. Appl. Math. , vol. 12&13, pp. 77–90, 1985

[4] D.E. Schwarz, R. Lamour *The Computation of Consistent Initial Values for Nonlinear Index-2 Differential-Algebraic Equations*, 1999, Preprint, available at `http://taylor.mathematik.hu-berlin.de/publ/pre/1999/p-list-99.html`

[5] Herman Mann, Dynast User's Manual, 1998, `http://icosym.cvut.cz/cacsd/msa/onlinetools.shtml`

[6] Hans Olsson, Lund University *Runge-Kutta Solution of Initial Value Problems, Methods, Algorithms and Implementation*, thesis, 1998

[7] Werner Haas, Kurt Schlacher, and Reinhard Gahleitner, *Modeling of Electromechanical Systems*, 1999, `http://regpro.mechatronik.uni-linz.ac.at/eu-projekt/Theory/cnew.html`

LAGRANGE'S EQUATIONS OF MOTION AND COULOMB FRICTION

E.-M. SALONEN
Laboratory of Theoretical and Applied Mechanics
Helsinki University of Technology
P.O. Box 1100
FIN-02015 HUT, FINLAND

ABSTRACT

The article is concerned with educational aspects of Lagrangian mechanics. In standard mechanics textbooks, when dealing with Lagrange's equations of motion, nothing is mentioned about how Coulomb friction should be taken into account. It is found in the article that even for a holonomic system it is useful to apply a non-minimal-coordinate formalism to be able to treat Coulomb friction in a straightforward way. First a double pendulum is used as an example case to explain in general terms how friction can be introduced into the formulation. Then an extremely simple case — particle motion on an inclined plane — is considered in more detail. It seems that emphasis in engineering mechanics education should be put ever more on non-minimal-coordinate formalisms to follow the trends found useful in practice.

INTRODUCTION

Textbooks on elementary mechanics both in statics and dynamics usually contain a number of applications where Coulomb friction is involved. When the treatment advances further towards the derivation and application of Lagrange's equations of motion, it seems almost as if the existence of friction is from some mutual agreement tactfully forgotten. If the word "friction" is mentioned, it, however, just means viscous damping, that is, forces proportional to velocity; e.g., Reference [1]. These viscous forces are then usually given some attention, e.g., References [1] and [2]. A person entering his studies of analytical mechanics through a path of basic mechanics can hardly be without wondering about this sudden defect in modeling of phenomena. The only reference dealing with Coulomb friction in connection with Lagrange's equations, the writer is aware of, is [3]. This presentation is, however, directed mainly on the input data generation for the ADAMS software, e.g., Reference [4], and not on textbook type theory considerations. In the following for simplicity just holonomic constraints are considered although standard non-holonomic constraints can be dealt with similarly. A rather obvious explanation for the negligence of Coulomb friction is given. Two example cases are treated to show how the effect of friction can be included.

FIRST EXAMPLE

To make the discussion concrete, let us consider the example case shown in Figure 1. The system is a double pendulum of rigid homogeneous bars with equal bar lengths l in plane motion under gravity. No friction is assumed initially at the joints.

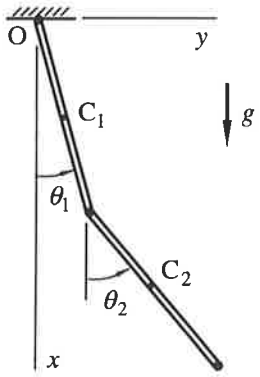


Figure 1 Double pendulum.

If this is considered in a standard textbook, the normal selection would rather probably consist of taking the angles

$$q_1 = \theta_1, \quad q_2 = \theta_2 \tag{1}$$

as the generalized coordinates. The two corresponding Lagrange's equations of motion are then, e.g., [1]:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} - Q_j^{\mathrm{a}} = 0, \quad j = 1,2 \tag{2}$$

The superscript a in the generalized force symbol refers here to the fact the the force is generated only by the to so-called "applied forces" (here gravity). This means that the forces acting on a system are divided into "constraint forces" which maintain the constraints and the rest of the forces are called applied forces. It is well-known that the constraint forces and their generalized counterparts do not appear in Lagrange's equations of motion when the constraints have been taken into account in advance by a suitable selection of the generalized coordinates. In the terminology of structural mechanics we have now a so-called "displacement formulation" as the basic unknowns $\theta_1(t)$ and $\theta_2(t)$ are displacement type quantities. Following Reference [4], we can call this classical formulation also a "minimal-coordinate formalism".

But if we have to take into account possible Coulomb friction at the joints, we are at loss as the normal contact forces being constraint forces acting at the joints, are not directly available and we thus cannot express the friction forces which are of the form: friction coefficient times the normal contact force. To approach this problem we must abandon the classical minimal-coordinate formalism. A rather natural and systematic approach would

now be to take as generalized coordinates the x- and y-coordinates of the mass centers C_1 and C_2 of the bars and the bar orientation angles:

$$q_1 = x_1, \quad q_2 = y_2, \quad q_3 = \theta_1, \quad q_4 = x_2, \quad q_5 = y_2, \quad q_6 = \theta_2 \tag{3}$$

These coordinates are, however, no more independent, and we have to record the obvious constraint equations

$$\begin{aligned}
\Phi_1 &\equiv q_1 - \frac{l}{2}\cos q_3 = 0 \\
\Phi_2 &\equiv q_2 - \frac{l}{2}\sin q_3 = 0 \\
\Phi_3 &\equiv q_1 + \frac{l}{2}\cos q_3 - q_4 + \frac{l}{2}\cos q_6 = 0 \\
\Phi_4 &\equiv q_2 + \frac{l}{2}\sin q_3 - q_5 + \frac{l}{2}\sin q_6 = 0
\end{aligned} \tag{4}$$

Lagrange's equations of motion are now

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} - Q_j^{\mathrm{a}} - \sum_{k=1}^{4} \lambda_k \frac{\partial \Phi_k}{\partial q_j} = 0, \quad j = 1,\cdots,6 \tag{5}$$

Instead of the classical formulation with only two unknown functions, we have now six displacement type unknowns — the generalized coordinates — and four force type unknowns — the Lagrange multipliers — to be determined from the ten equations (4) and (5). Using again the terminology of structural mechanics, this is a "mixed formulation" as both displacements and forces appear as basic unknowns. Also the name "non-minimal-coordinate formalism" can be used. This kind of formulation is not very tempting in introductory presentations of Lagrangian mechanics because of the large number and of the rather involved structure of the equations. First, the equations are no more pure differential equations but differential-algebraic equations. Second, standard numerical solution packages such as the Mathematica program [5], cannot cope with such systems and specially tailored solution software, e.g., [6], must be used. However, these kind of formulations are probably more and more important in future teaching of engineering mechanics as they offer better modeling possibilities and the great number of unknowns becomes less important with the increasing capacity of computers.

We proceed still without friction and notice, e.g., [1], that the generalized forces due to the constraints are given by

$$Q_j^{\mathrm{c}} = \sum_{k=1}^{4} \lambda_k \frac{\partial \Phi_k}{\partial q_j} \tag{6}$$

Subscript c refers to "constraint". Further, in any application we know which physical constraint a constraint equation is associated with. As an example, we consider here the revolute joint O at the origin. The two first constraint equations (4) are associated with it. Thus, the generalized constraint forces due to joint O are

$$(Q_j^c)_O = \lambda_1 \frac{\partial \Phi_1}{\partial q_j} + \lambda_2 \frac{\partial \Phi_2}{\partial q_j} \tag{7}$$

or in detail

$$\begin{aligned}
(Q_1^c)_O &= \lambda_1 \frac{\partial \Phi_1}{\partial q_1} + \lambda_2 \frac{\partial \Phi_2}{\partial q_1} = \lambda_1 \\
(Q_2^c)_O &= \lambda_1 \frac{\partial \Phi_1}{\partial q_2} + \lambda_2 \frac{\partial \Phi_2}{\partial q_2} = \lambda_2 \\
(Q_3^c)_O &= \lambda_1 \frac{\partial \Phi_1}{\partial q_3} + \lambda_2 \frac{\partial \Phi_2}{\partial q_3} = \lambda_1 \frac{l}{2} \sin q_3 - \lambda_2 \frac{l}{2} \cos q_3 \\
(Q_4^c)_O &= \lambda_1 \frac{\partial \Phi_1}{\partial q_4} + \lambda_2 \frac{\partial \Phi_2}{\partial q_4} = 0 \\
(Q_5^c)_O &= \lambda_1 \frac{\partial \Phi_1}{\partial q_5} + \lambda_2 \frac{\partial \Phi_2}{\partial q_5} = 0 \\
(Q_6^c)_O &= \lambda_1 \frac{\partial \Phi_1}{\partial q_6} + \lambda_2 \frac{\partial \Phi_2}{\partial q_6} = 0
\end{aligned} \tag{8}$$

To proceed further, we must consider the situation at the joint. Figure 2 (a) shows schematically the traction distribution exterted on the bar surface from the pin at O.

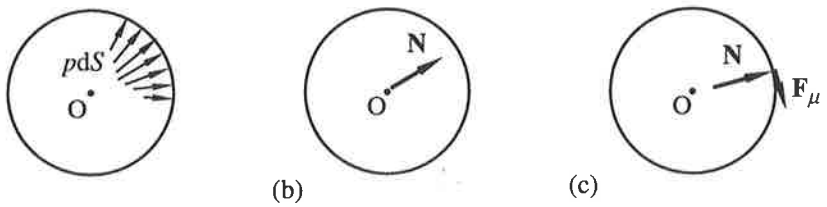


Figure 2 (a) Force system acting on the bar at the joint (no friction). (b) Force system reduced at the origin (no friction). (c) Force system (with friction) reduced at the point of contact on the bar.

As we assume here still no friction, the line of action of each force differential $p\mathrm{d}S$ goes through the axis of the pin and the force system from the tractions reduced at O gives a resultant (Figure 2 (b))

$$\mathbf{N} = N_x \mathbf{i} + N_y \mathbf{j} \tag{9}$$

without a moment. A short calculation gives the virtual work expression

$$\delta W = N_x \delta q_1 + N_y \delta q_2 + (N_x \frac{l}{2} \sin q_3 - N_y \frac{l}{2} \cos q_3) \delta q_3 \tag{10}$$

for this force system. The corresponding generalized forces are thus

$$
\begin{aligned}
(Q_1^c)_O &= N_x \\
(Q_2^c)_O &= N_y \\
(Q_3^c)_O &= N_x \frac{l}{2}\sin q_3 - N_y \frac{l}{2}\cos q_3
\end{aligned}
\tag{11}
$$

Comparison of these with expressions (8) gives the equations

$$
\begin{aligned}
N_x &= \lambda_1 \\
N_y &= \lambda_2 \\
N_x \frac{l}{2}\sin q_3 - N_y \frac{l}{2}\cos q_3 &= \lambda_1 \frac{l}{2}\sin q_3 - \lambda_2 \frac{l}{2}\cos q_3
\end{aligned}
\tag{12}
$$

This is an overdetermined but consistent system of equations giving the simple results

$$
N_x = \lambda_1, \qquad N_y = \lambda_2 \tag{13}
$$

for the constraint force components.

We can now start to include the effect of friction. To deal with Coulomb friction we would in principle need to know the minute details of the contacting surfaces as the pressure distribution depends on it and we have in fact a complicated problem of contact mechanics. We proceed here in the standard way by assuming some looseness between the surfaces so that the contact is over a small area, in fact a point contact, and we have the simplified force distribution of Figure 2 (c) where $\mathbf{F}_\mu$ is the friction force. The diagram has been drawn assuming a positive $\dot{q}_3$. It must be realized that when there is relatove motion, the Coulomb friction force is an applied force although it depends on the constraint force value. Thus the constraint force expressions (13) are seen still to be valid. The constraint force values will naturally in general depend also on friction but the formulas remain unaltered. The friction force expression is thus

$$
\begin{aligned}
\mathbf{F}_\mu &= \mu N \operatorname{sign}(\dot{q}_3)\frac{N_y\mathbf{i} - N_x\mathbf{j}}{N} = \mu \operatorname{sign}(\dot{q}_3)(N_y\mathbf{i} - N_x\mathbf{j}) \\
&= \mu \operatorname{sign}(\dot{q}_3)(\lambda_2\mathbf{i} - \lambda_1\mathbf{j})
\end{aligned}
\tag{14}
$$

We do not consider here the special case $\dot{q}_3 = 0$. The generalized forces corresponding to (14) can be determined and added as contributions to to the left-hand sides of (5). A similar study can be performed with respect to the second joint. Now Coulomb friction effects are in principle included in the formulation. We, however, do not perform the rather involved computations here but consider a second much simpler example case in more detail.

SECOND EXAMPLE

An extremely simple demonstration case shown in Figure 3 (a) is considered: motion of a particle under gravity on an inclined plane. We take the cartesian coordinates x and y of the particle as the generalized coordinates of the problem and use also these symbols in the equations to follow.

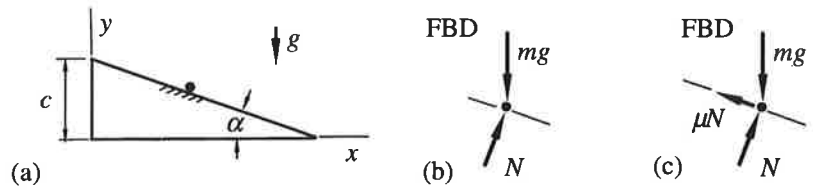


Figure 3 (a) Particle in motion on an inclined plane. (b) Free-body diagram of the particle without friction and (c) with friction.

We have now one constraint equation:

$$\Phi \equiv y + x\tan\alpha - c = 0 \tag{15}$$

The kinetic energy expression is

$$T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) \tag{16}$$

We consider the problem again first without friction. The generalized forces from the applied forces (gravity) are found simply to be

$$Q_x^{\mathrm{a}} = 0, \qquad Q_y^{\mathrm{a}} = -mg \tag{17}$$

and Lagrange's equations of motion are

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial T}{\partial \dot{x}} - \frac{\partial T}{\partial x} - Q_x^{\mathrm{a}} - \lambda\frac{\partial \Phi}{\partial x} = 0 \\ \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial T}{\partial \dot{y}} - \frac{\partial T}{\partial y} - Q_y^{\mathrm{a}} - \lambda\frac{\partial \Phi}{\partial y} = 0 \end{aligned} \tag{18}$$

A short calculation gives the detailed forms

$$\begin{aligned} m\ddot{x} - \lambda\tan\alpha = 0 \\ m\ddot{y} + mg - \lambda = 0 \end{aligned} \tag{19}$$

The three governing equations are thus (19) and (15) for the three unknowns $x(t)$, $y(t)$, $\lambda(t)$.

The generalized constraint force expressions are

$$Q_x^{\mathrm{c}} = \lambda\tan\alpha, \qquad Q_y^{\mathrm{c}} = \lambda \tag{20}$$

The next problem is again to determine the constraint force expression in terms of the Lagrange multiplier. From the free-body diagram of Figure 3 (b) we can readily deduce the alternative generalized constraint force expressions. The virtual work of the constraint force is

$$\delta W = N\sin\alpha\,\delta x + N\cos\alpha\,\delta y \tag{21}$$

and the corresponding generalized constraint forces are thus

$$Q_x^c = N\sin\alpha, \qquad Q_y^c = N\cos\alpha \tag{22}$$

Comparison of these with (20) gives the equations

$$\begin{aligned} N\sin\alpha &= \lambda\tan\alpha \\ N\cos\alpha &= \lambda \end{aligned} \tag{23}$$

This is an overdetermined but consistent system of equations for N and we obtain the solution

$$N = \frac{\lambda}{\cos\alpha} \tag{24}$$

Now we include friction. Again, the constraint force expression (24) is not changed by the effect of friction. Using the notation of Figure 3 (c), we obtain the virtual work

$$\begin{aligned} \delta W &= -\operatorname{sign}(\dot{x})\mu N\cos\alpha\,\delta x + \operatorname{sign}(\dot{x})\mu N\sin\alpha\,\delta y \\ &= -\operatorname{sign}(\dot{x})\mu\frac{\lambda}{\cos\alpha}\cos\alpha\,\delta x + \operatorname{sign}(\dot{x})\mu\frac{\lambda}{\cos\alpha}\sin\alpha\,\delta y \end{aligned} \tag{25}$$

due to friction (we do not consider the case $\dot{x} = 0$ here) and the additional generalized forces are thus

$$Q_x^{af} = -\operatorname{sign}(\dot{x})\mu\lambda, \qquad Q_y^{af} = \operatorname{sign}(\dot{x})\tan\alpha\cdot\mu\lambda \tag{26}$$

Equations of motion (19) are thus modified with the inclusion of friction into the forms

$$\begin{aligned} m\ddot{x} - \lambda\tan\alpha + \operatorname{sign}(\dot{x})\mu\lambda &= 0 \\ m\ddot{y} + mg - \lambda - \operatorname{sign}(\dot{x})\tan\alpha\cdot\mu\lambda &= 0 \end{aligned} \tag{27}$$

As a check we can differentiate (15) twice to obtain

$$\ddot{y} + \tan\alpha\cdot\ddot{x} = 0 \tag{28}$$

and then solve for $\ddot{x}$, $\ddot{y}$, λ from (27) and (28). There is obtained

$$\begin{aligned} \ddot{x} &= g\cos\alpha\cdot\sin\alpha - \operatorname{sign}(\dot{x})\mu g\cos^2\alpha \\ \ddot{y} &= -g\sin^2\alpha + \operatorname{sign}(\dot{x})\mu g\cos\alpha\cdot\sin\alpha \\ \lambda &= mg\cos^2\alpha. \end{aligned} \tag{29}$$

It is rather easily found that the first two equations (29) are indeed the same obtained as conventional component equations of motion from the free-body diagram of Figure 3 (c) combined with the differentiated constraint equation (28).

CONCLUDING REMARKS

The ever increasing importance of computers is changing inevitably in many ways the emphasis in the contents of engineering mechanics curricula. The mechanism simulation packages used in practice — such as the ADAMS software — employing Lagrangian mechanics formulation, seem to favor non-minimal-coordinate formalisms to achieve

systematic and flexible formulations usable in various applications. The price is a large number of unknowns but this is acceptable because of the growing computational power. This trend must be taken into account also in teaching. In this article it is shown how Coulomb friction can be included with a non-minimal-coordinate formalism. It seems that the reason for the textbooks avoiding discussion of Coulomb friction in connection with Lagrange's equations of motion has its basis in the desire to apply minimal-coordinate formalism which is understandable when hand calculations are used to generate the equations. Then, however, direct information about the constraint forces disappears.

ACKNOWLEDGMENT

The writer wants to thank Mr. Mauri Vesimäki for bringing Reference [3] to his attention.

REFERENCES

1. Goldstein, H. *Classical Mechanics*, Addison-Wesley, Reading, Massachuchetts, 1971.
2. Meirovitch, L. *Methods of Analytical Dynamics*, McGraw-Hill, New York, 1970.
3. Krueger, M. Modeling Coulomb Friction In Joints, *1992 International ADAMS Users' Conference*, Ypsilanti, Michigan, May 13, 1992.
4. Ryan, R.R. ADAMS Multibody System Analysis Software. *Mechanical Dynamics Inc.* 3055 Plymoth Road, Ann Arbor, MI 48105-3203, 1989.
5. Wolfram, S. *The Mathematica Book*, 4th ed., Wolfram Media/Cambridge University Press, 1999
6. Gear, C. W. The Simultaneous Numerical Solution of Differential Algebraic Systems, *IEEE Transactions on Circuit Theory*, Vol.1, 1971.

YHDISTETTY SÄHKÖMAGNEETTIS–MEKAANINEN TEHTÄVÄ

S. SUURINIEMI & L. KETTUNEN
TTKK, Sähkömagnetiikan laitos
PL 692, 33101 Tampere
email: *Saku.Suuriniemi@tut.fi*

TIIVISTELMÄ

Tarkastelemme johtavien ja magnetoituvien jäykkien kappaleiden liikkeen laskemista dynaamisessa magneettikentässä. Liike ja magneettikentän lähteiden muutokset ajan suhteen aiheuttavat johtaviin kappaleisiin induktiovirtoja, joihin magnetoitumien ohella ulkoinen magneettikenttä kohdistaa voimia ja vääntömomentteja, muuttaen kappaleiden liiketilaa.

Esitämme laskentamenetelmän, joka perustuu heikosti kytkettyyn mekaanisen ja sähkömagneettisen osatehtävän aikaintegrointiin. Parametrisoimme rotaatiot Eulerin parametreilla, laskemme voimat magneettisella ekvivalenttilähdemenetelmällä ja ratkaisemme magneettikenttätehtävän niin kutsutulla H-tyyppisellä hybridielementtimenetelmällä. Esitämme myös muutamia menetelmän laskentatuloksia.

1 JOHDANTO

Sähkömagneettis–mekaaniset tehtävät ovat usein kiinnostavia suunniteltaessa sähkötekniisiä laitteita. Moottorien, releiden, generaattoreiden ja levitaatiolaitteiden numeerisessa mallintamisessa joudutaan ratkaisemaan sähkömagneettinen kenttätehtävä yhtäaikaisesti jonkintyyppisen mekaanisen tehtävän kanssa.

Ratkaistaessa sähkömagneettis–mekaanisia tehtäviä on pystyttävä ratkaisemaan kolme osatehtävää: Mekaaninen malli (tässä jäykän kappalen dynamiikka), sähkömagneettisten voimien laskenta, sekä sähkömagneettinen kenttätehtävä. Osatehtävistä koottua dynaamista tehtävää voidaan aikaintegroida joko heikosti kytkettynä, jolloin osatehtäviä integroidaan vuorotellen, tai vahvasti kytkettynä, jolloin koko fysikaalinen malli

kuvataan yhdellä differentiaalialgebrallisella (DAE) yhtälöryhmällä, jonka kaikki yhtälöt ratkaistaan samanaikaisesti.

Sähkömagneettis–mekaanisten tehtävien ratkaiseminen kolmessa dimensiossa vaatii huomattavia laskentaresursseja, ja siksi siihen ollaankin vakavasti ryhdytty vasta melko taannoin. Tehtävää on tyypillisesti lähdetty ratkaisemaan elementtimenetelmällä, jota on laajennettu siten, että voidaan laskea jonkin geometrialtaan yksinkertaisen kappaleen verrattain yksinkertainen liike, kuten sähkökoneen sylinterimäisen roottorin tasorotaatio. Liikkeiden on tavallisesti sovitettu verkkoon jollain erityisellä tekniikalla, kuten liukuverkolla. Koska tähtäämme mielivaltaisen muotoisten jäykkien kappaleiden vapaan liikkeen laskemiseen, joutuisimme perinteistä elementtimenetelmää käyttäessämme verkottamaan uudelleen kappaleiden välisiä – usein suuria ja hankalan muotoisia – alueita joka aika-askeleella. Välttääksemme tämän hitaan ja vaivalloisen työn, käytämme apunamme integraalioperaattoreita asettaessamme magneettista kenttätehtävää.

Käytämme liikkuvissa kappaleissa sekä mekaanisten että sähkömagneettisten ilmiöiden tarkastelussa *Lagrangen tarkastelua*, jokaisessa kappaleessa omassa lokaalissa koordinaatistossaan. Lokaalissa koordinaatistossa ilmaistuja suureita merkitsemme pilkulla, kuten kulmanopeutta ω'. Rajoitamme tarkastelun kvasistaattiseen tapaukseen, jolloin sähkömagneettista aaltoilmiötä ei esiinny ja magneettivuon tiheysvektori B sekä magneettikentän voimakkuusvektori H pysyvät samoina havaitsijan liiketilasta riippumatta [1].

2 OSATEHTÄVÄT

2.1 Mekaaninen malli

Chaslesin teoreeman nojalla jäykän kappaleen yleinen liike on jaettavissa translaatioon ja rotaatioon, ja näitä liikkeitä kuvaavat vastaavasti Newtonin ja Eulerin yhtälöryhmät. Mekaaniseksi malliksi kullekin jäykälle kappaleelle yhdistämme yhtälöt Newton–Euler-yhtälöryhmäksi

$$F = ma \tag{1}$$
$$\tau' = \mathbf{J}'\alpha' + \omega' \times \mathbf{J}'\omega'. \tag{2}$$

Äärellisen rotaation parametrisointiin on olemassa useita eri tapoja, joista yksinkertaisimpia ovat kolmiparametriesitykset, kuten Eulerin ja Bryantin kulmat. Nämä pa-

rametrisoinnit ovat kuitenkin singulaarisia. Siten ne johtavat helposti DAE-ryhmiin, joiden derivointi-indeksi [2] saattaa olla hyvinkin korkea, ja mikä pahinta, muuttua rotaatiosta riippuen. Valitsemme siis Eulerin parametrit [3], jotka ovat neliparametrinen esitystapa. Eulerin parametreilla e ja kulmanopeudella (ω', lokaalissa koordinaatistossa) on verrattain yksinkertainen yhteys,

$$\omega' = 2\mathbf{L}\dot{e} = 2\begin{bmatrix} -e_1 & e_0 & e_3 & -e_2 \\ -e_2 & -e_3 & e_0 & e_1 \\ -e_3 & e_2 & -e_1 & e_0 \end{bmatrix}\dot{e}. \tag{3}$$

Muotoilemme [3], [4] yhtälöryhmän (2) uudelleen Lagrangen kertojaa käyttäen:

$$\mathbf{J}'\dot{\omega}' = \tau' - \omega' \times \mathbf{J}'\omega', \tag{4}$$

$$\dot{e} = \frac{1}{2}\mathbf{L}^T\omega' + e\lambda, \tag{5}$$

$$0 = ||e|| - 1. \tag{6}$$

Tehtävän derivointi-indeksi on aina 2 [4], ja useimmat DAE-ratkaisijat kykenevät ratkaisemaan tämäntyyppisiä tehtäviä.

2.2 Magneettikenttätehtävä

Niin kutsuttua pyörrevirtatehtävää kuvataan *kvasistaattisilla Maxwellin yhtälöillä*,

$$\mathrm{rot}H = J, \tag{7}$$

$$\mathrm{rot}E = -\frac{\partial B}{\partial t}, \tag{8}$$

$$\mathrm{div}B = 0. \tag{9}$$

Lisäksi tarvitsemme väliaineyhtälöt

$$B = \mu H, \tag{10}$$

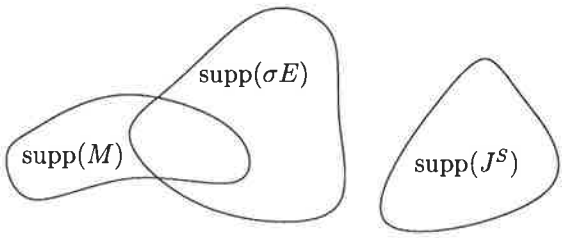
$$D = \epsilon E, \tag{11}$$

$$J = \sigma E + J^S, \tag{12}$$

jossa σE merkitsee induktiovirtoja ja J^S lähdevirtoja, jotka tunnetaan ennalta ja joiden kantajia kutsutaan tässä käämeiksi.

Mikäli tunnemme lähdevirrat J^S, induktiovirrat σE ja magnetoituman M, saamme laskettua kaiken magneettikentän informaation muodostamalla näiden kantajien[1] yli tietyt *integraalioperaattorit*. Nämä ovat i) Biot–Savartin integraalioperaattori, jolla voimme laskea virtojen ympärilleen luoman kentän, sekä ii) integraalioperaattori, jolla laskemme magnetoitumien aiheuttaman magneettikentän. Nämä voidaan puolestaan ilmaista H:n avulla, koska $\sigma E + J^S = \mathrm{rot}H$, ja $M = \chi(H)H$.

[1] support, $\mathrm{supp}(f) = \{x \in \mathrm{Dom}(f) | f(x) \neq 0\}$



Kuva 1: Virtojen ja magnetoitumien kantajat.

Omaksumme merkinnän $\text{supp}(M) \cup \text{supp}(\sigma E) = \Omega$. Edellisessä kappaleessa esitetystä seuraa, että mikäli tunnemme H:n Ω:ssa sekä J^S:n omassa kantajassaan, voimme ilmaista H:n integraalioperaattorein alueen Ω reunalla $\partial\Omega$:ssa (kuten missä tahansa muuallakin). Asettamalla reunaehto tällä tavoin *implisiittisenä* $\partial\Omega$:lla, reuna–arvotehtävä on hyvin asetettu kiintopisteteorian [5] nojalla.

Reuna–arvotehtävälle löydetään approksimaatio määrittelyalueessa Ω diskretoimalla magneettikenttä H samoin kuin tavallisessa elementtimenetelmässä. Diskreeteissä avaruuksissa operaattoreita vastaavat matriisit, ja lopputuloksena saadaan hyvin asetettu yhtälöryhmä, jolla on siis yksikäsitteinen ratkaisu. Tässä yhtälöryhmässä määrittelyalueessa Ω toteutuvat yhtälöt (7)–(9) muodostavat diskretoituina indeksin 2 DAE:n [6]. Lopulta magneettikenttä voidaan integroida diskreetistä H:n ratkaisusta haluttuihin Ω:n ulkopuolisiin pisteisiin. Tätä tarvitaan esimerkiksi voimien laskennassa.

2.3 Voimien laskenta

Sähkökenttä (E) ja magneettikenttä (B) aiheuttavat varaustiheyksiin (ρ) ja virrantiheyksiin (J) *Lorentz-voiman*,

$$F = \int_V \rho E + J \times B dV. \tag{13}$$

Tässä tapauksessa oletamme johdekappaleiden nettovarauksen nollaksi, ja kvasistaattisessa tapauksessa virrantiheyden lähteettömäksi (7). Virran jatkuvuusyhtälön

$$\text{div} J + \frac{\partial \rho}{\partial t} = 0 \tag{14}$$

nojalla $\rho \equiv 0$ ja voimme jättää huomioimatta (13):n ensimmäisen termin pyörrevirtatehtävässä.

Liikkuviin johtaviin kappaleisiin indusoituu pyörrevirtoja (σE), ja lisäksi magnetoituviin kappaleisiin muodostuu magnetoitumia (M) materiaalin magneettisten dipolien

järjestymistaipumuksen seurauksena. Ulkoinen magneettikenttä kohdistaa voiman sekä pyörrevirtoihin että magnetoitumiin.

Voimme korvata magnetoituneen materiaalin *ekvivalenttisella virrantiheydellä*

$$J^M = \mathrm{rot} M \tag{15}$$

ja saada materiaalin ulkopuolelle tarkalleen saman kentän kuin korvatusta magnetoitumasta aiheutuisi [7]. Kahden materiaalin rajapinnoille (jotka ovat magnetoituman epäjatkuvuuskohtia) voimme laskea *ekvivalenttisen pintavirrantiheyden*

$$j^M = [M_1 - M_2] \times n. \tag{16}$$

Tässä n on materiaalin 1 ulkonormaali. Toisaalta voimme myös laskea magnetoitumaan kohdistuvan voiman käyttäen ekvivalenttisia virtoja ja *ulkoisten lähteiden aiheuttamaa magnettikenttää* B^* [8].

Summaamalla virrantiheyskomponentit saamme lausekkeen kappaleeseen 1 kohdistuvalle kokonaisvoimalle:

$$F = \int_V (\sigma E + \mathrm{rot} M_1) \times B^* dV + \oint_{\partial V} ([M_1 - M_2] \times n) \times B^* da. \tag{17}$$

Diskretoidussa mallissa käytämme lauseketta elementeittäin. Vääntömomenttien laskennassa noudatetaan aivan samaa menetelmää.

3 AIKAINTEGROINTI

Aikaintegrointimenetelmää valittaessa on pohjimmiltaan valittava *heikosti* ja *vahvasti kytkettyjen* menetelmien välillä [9]: Vahvasti kytketyissä menetelmissä osatehtävät kootaan yhdeksi yhtälöryhmäksi, jonka yhtälöt ratkaistaan samanaikaisesti yhdellä aikaintegraattorilla. Heikosti kytketyissä menetelmissä osatehtävien integrointi suoritetaan vuorotellen, mahdollisesti eri integraattoreilla.

Heikosti kytketyt menetelmät soveltuvat sähkömagneettis–mekaanisten tehtävien ratkaisemiseen, koska tavallisesti sähkömagneettisten ja mekaanisten ilmiöiden muutosnopeudet ovat aivan eri kertaluokkaa. Tällöin voidaan käyttää erillisiä integraattoreita, erilaisin askelpituusin. Heikosti kytkettyjen menetelmien ongelmaksi nousee usein ehdollinen stabiilisuus, joka rajoittaa käytettyä askelpituutta.

Ryhmän (4)–(6) integrointiin tarvitaan sopiva DAE–integraattori. Tapauksissa, jossa liikkeeseen liittyy rotaatio, on integrointiin käytetty Radau5–integraattoria [2], joka kykenee rakaisemaan epälineaarisia, mahdollisesti jäykkiä DAE–tehtäviä, joiden indeksi on korkeintaan 3. Mikäli liike on ollut puhdasta translaatiota – esimerkiksi symmetrian vuoksi – on käytetty ODE–ryhmän ratkaisemiseen RKSUITE90–paketin eksplisiittisiä Runge–Kutta–integraattoreita [10].

4 RATKAISUMENETELMÄ

Tehtävässä käytetty H-hybridiformulaatio on implementoitu BARITONE–ohjelmiston aikatransienttiratkaisijaan [11]. Sähkömagneettisten yhtälöiden ratkaiseminen suoritetaan BARITONEn DAE–integraattorilla [6]. Tämä tekee heikosti kytketystä integrointimenetelmästä varsin luonnollisen valinnan. Voimanlaskennassa käytetty menetelmä valittiin aikaisemman kokemuksemme perusteella [8].

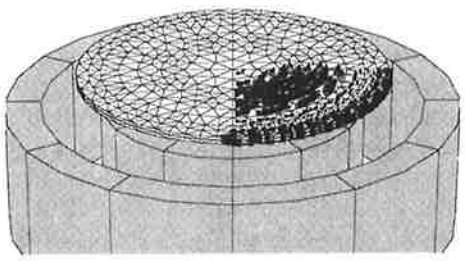
Edellisiin valintoihin perustuva heikosti kytketty menetelmä on:

1. Ratkaise H.
2. Muuta magnetoitumat evivalenttisiksi virroiksi.
3. Laske kappaleisiin kohdistuvat voimat ja vääntömomentit.
4. Integroi mekaanisen järjestelmän aika-askel.
5. Päivitä geometria ja aloita askeleesta 1.

5 ESIMERKKEJÄ

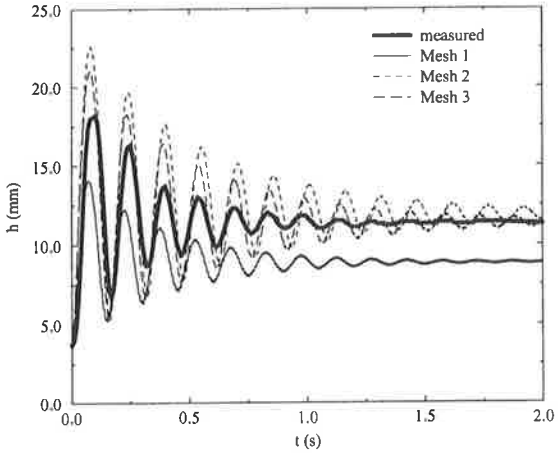
5.1 Standarditestitehtävä

Erityyppisiä standarditestitehtäviä, joilla voidaan tarkistaa numeeristen ratkaisumenetelmien tarkkuutta ja luotettavuutta, on koottu TEAM–sarjaan (Testing Electromagnetic Analysis Methods). Tehtävänkuvauksen mukaiset koejärjestelmät on rakennettu ja niiden vasteet on mitattu, jonka jälkeen mittaustulokset on julkaistu vertailutuloksiksi.



Kuva 2: Tehtävä TEAM 28: Alumiinilevy leijuu käämien sinimuotoisen virran aiheuttamassa magneettikentässä. Virrantiheys levyssä esitetty hetkellä t=2ms.

TEAM–tehtävä numero 28:ssa [12] (kuva 2) asetetaan pyöreä alumiinilevy kahden samankeskisen toroidikäämin päälle. Alumiinilevy sijoitetaan tarkalleen keskelle, jolloin tehtävä on symmetrinen, eikä rotaatioita esiinny. Kun käämeihin aletaan syöttää sinimuotoista virtaa (vaiheet vastakkaisia), alumiinilevyyn indusoituu pyörrevirtoja. Näihin pyörrevirtoihin kohdistuu puolestaan repulsiivinen voima, jolla on tasapainopiste painovoiman suhteen 11.3 mm:n korkeudessa. Tehtävänä on laskea levyn leijuntakorkeus ajan funktiona.

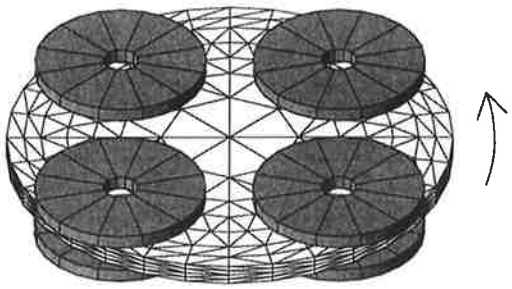


Kuva 3: Levyn leijuntakorkeus (Verkko 1: 344, Verkko 2: 1492 and Verkko 3: 3735 tetraedriä).

Tehtävä laskettiin edellä kuvatulla menetelmällä, useita verkkoja käyttäen. Tulokset on esitetty kuvassa 3. Huomaamme, että alun transientti on tarkemissa laskelmissa liian voimakas: Virtaa pidettiin mallissa pakotettuna, kun taas todellisuudessa pyörrevirrat pyrkivät vastustamaan käämeihin johdettua virtaa. Laskentatuloksen vaihevaste on melko hyvin kohdallaan, mutta värähtelyn vaimeneminen on hieman liian hidasta. Arvelemme tämän johtuvan voimien laskentamenetelmästä.

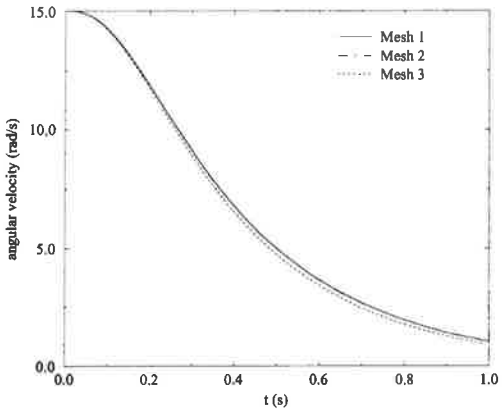
5.2 Pyörrevirtajarru

Toinen ratkaistu testitehtävä on pyörrevirtajarrun malli (kuva 4). Pyöreälle alumiinilevylle annetaan alkunopeudeksi 15 rad/s. Levyn läheisyyteen asetetaan kahdeksan käämiä, joiden magneettivuot lävistävät alumiinilevyn. Käämeihin syötetään kasvava virta, $I = I_0(1 - e^{-t/0.1s})$, jotta pyörivään levyyn saataisiin indusoitumaan pyörrevirtoja. Pyörrevirrat aiheuttavat levyä hidastavan vääntömomenetin.



Kuva 4: Pyörrevirtajarru, 3861 elementtiä yläpuoliskossa. Käämit varjostettu.

Levystä rakennettiin symmetriaa hyväksi käyttäen kolme mallia, joissa levyn yläpuolikossa on verkossa 1 3861, verkossa 2 4764, ja verkossa 3 7587 elementtiä. Kuvassa 5 näkyvät laskentatulokset, jotka sopivat hyvin yhteen odotusten kanssa: Kun käämien virta hetkellä $t = 0.1$ s on saavuttanut huomattavan osan (63%) lopullisesta arvostaan, kulmanopeus alkaa hidastua eksponentiaalisesti.



Kuva 5: Alumiinilevyn kulmanopeus ω pyörrevirtajarrussa.

6 YHTEENVETO

Tässä paperissa on esitetty H-hybridiformuloituun elementtimenetelmään perustuva numeerinen ratkaisumenetelmä haastavalle yhdistetylle tehtävätyypille, magneettikenttätehtävään yhdistetylle jäykän kappaleen dynamiikalle. Esitämme kaksi esimerkkitehtävää, joista lähinnä ensimmäisen nojalla voidaan eksakteihin mittauksiin perustuen väittää menetelmää onnistuneeksi.

Menetelmän jatkotutkimus tulee keskittymään eri voimanlaskentamenetelmiin, deformoituviin kappaleisiin, sekä mahdollisesti monikappaledynamiikkaan.

LÄHTEET

[1] S. Kurz, J. Fetzer, G. Lehner, W.M. Rucker: "A novel formulation for 3D eddy current problems with moving bodies using Lagrangian description and BEM-FEM coupling", *IEEE Trans. Magn.* vol. 34, pp. 3068-3073, September 1998.

[2] E. Hairer, G. Wanner, *Solving ordinary differential equations II, stiff and differential–algebraic equations*, Springer series in computational mathematics 14, Berlin: Springer 1991.

[3] E.J. Haug, *Computer aided kinematics and dynamics of mechanical systems*, vol.1, Boston, London, Sydney, Toronto: Allyn and Bacon 1989.

[4] K.E. Brenan, S.L. Campbell, L.R. Petzold, *Numerical solution of initial–value problems in differential–algebraic equations*, SIAM series Classics in Applied Mathematics, Philadelphia: SIAM 1996.

[5] A. Bossavit: *Computational electromagnetism*, London: Academic Press, 352 pp. 1998.

[6] F. Cameron, R. Piché, K. Forsman: "Variable Step Size Time Integration Methods for Transient Eddy Current Problems", *IEEE Trans. Magn.* vol. 34, pp. 3319-3322, September 1998.

[7] J.A. Stratton, *Electromagnetic theory*, New York, London: McGraw–Hill, 1941.

[8] A. Koski, K. Forsman, T. Tarhasaari, J. Kangas, L. Kettunen: "Force and torque computations with hybrid methods", *IEEE Trans. Magn.* 1999, painossa.

[9] W.L. Wood: *Practical Time–Stepping Schemes*, Oxford: Clarendon Press, 373 pp. 1990.

[10] R.W. Brankin, I. Gladwell: "A Fortran 90 Version of RKSUITE: An ODE Initial Value Solver", *Annals of Numer. Math.*, vol. 1 1994, pp. 363-375.

[11] L. Kettunen, K. Forsman, A. Koski, T. Tarhasaari: "Integral operators in computational electromagnetics", *4th International Workshop on Electric and Magnetic fields*, (Marseille, Ranska) 1998.

[12] H. Karl, J. Fetzer, S. Kurz, G. Lehner, W.M. Rucker: "Description of TEAM workshop problem 28: An electrodynamic levitation device", *Proc. TEAM workshop, 6th round*, (Rio de Janeiro, Brasilia) 1997.

ON SOME NEW DEVELOPMENTS IN CONJUGATION BASED FACTORIZED APPROXIMATE INVERSE PRECONDITIONING

Michele BENZI[§], Reijo KOUHIA[¶] and Miroslav TŮMA[‡]

[§] *Scientific Computing Group CIC-19, Los Alamos National Laboratory, Los Alamos, NM 87545, USA*
[¶] *Laboratory of Structural Mechanics, Helsinki University of Technology, P.O. Box 2100, 02015 HUT, Finland*
[‡] *Institute of Computer Science, Czech Academy of Sciences, 182 07 Prague 8, Czech Republic*

ABSTRACT

Some new developments of conjugation-based sparse approximate inverse preconditioners are considered in application to problems of solid and structural mechanics. A block version of the AINV algorithm and its stabilized modification are presented. The stabilized AINV algorithm, which in exact arithmetic is breakdown-free for any SPD matrix, is found to result in robust preconditioners even in difficult thin shell applications.

INTRODUCTION

Preconditioned Krylov subspace methods have proved to be efficient in solving large, sparse linear systems in many areas of scientific computing. However, lack of robustness is the main weakness of iterative solvers in comparison to direct ones which are still favoured in some areas of industrial computations. It is generally agreed that the preconditioning is the key factor to the success or failure of the iterative scheme.

In the last few years there has been considerable interest in explicit preconditioning techniques based on directly approximating the inverse of the coefficient matrix with a sparse matrix. Sparse approximate inverses have been shown to result in good rates of convergence of the preconditioned iteration (comparable to those obtained with incomplete factorization methods) while being well-suited for implementation on vector and parallel architectures. Because the application of the preconditioner reduces to matrix–vector products, parallelization is straightforward.

Although the main motivation for the development of sparse approximate inverse techniques comes from parallel processing, an additional important aspect is their robustness and reliability.

Aproximate inverse preconditioning methods can be grouped into two categories: (a) methods based on Frobenius norm minimization (SPAI), (b) factorized sparse approximate inverses. Frobenius norm based techniques can be rather effective, but

they suffer from some disadvantages (e.g. the preconditioner can be singular) which can be avoided by considering factorized sparse approximate inverses. Factorized approximate inverse techniques can be subdivided into four groups: (i) the FSAI approach, (ii) incomplete (bi)conjugation, (iii) bordering methods and (iv) inverse ILU methods. Methods in groups (i) to (iii) belong to the class of methods which do not require any information about the trianglar factors of $\boldsymbol{A}$.

In this paper, some new developments of conjugation based sparse approximate inverse preconditioners (AINV) are considered in application to problems of solid and structural mechanics where the coefficient matrix is a symmetric and positive definite (SPD) matrix. For comprehensive review of approximate inverse preconditioners, see ref. [9]. These new developments include the stabilization approach and blocking. The basic AINV approach may not be well-defined for an arbitrary SPD matrix, resulting in breakdowns of the conjugation process. Stabilized AINV is a modification of the basic approach which is inherently stable and well defined for an arbitrary SPD matrix. It is mathematically equivalent to the standard AINV algorithm when no dropping is applied. Even though the pointwise stabilized AINV is breakdown free the PCG iterations may fail due to slow convergence. It is shown that blocking can improve the convergence behaviour, especially in the analysis of shells. The effect of various ordering strategies is also studied.

THE AINV ALGORITHM

The AINV algorithm [6, 7] builds a factorized sparse approximate inverse of the form

$$\boldsymbol{M} = \boldsymbol{Z}\boldsymbol{D}^{-1}\boldsymbol{Z}^T \approx \boldsymbol{A}^{-1}$$

where $\boldsymbol{Z}$ is a unit upper triangular matrix and $\boldsymbol{D}$ is diagonal. The approximate inverse factor $\boldsymbol{Z}$ is a sparse approximation of the inverse of the $\boldsymbol{L}^T$ factor in the $\boldsymbol{L}\boldsymbol{D}\boldsymbol{L}^T$ decomposition of $\boldsymbol{A}$. When the AINV process is performed exactly, the diagonal matrix $\boldsymbol{D}$ is the same in the two decompositions, and contains the *pivots* down the main diagonal.

The AINV algorithm computes $\boldsymbol{Z}$ and $\boldsymbol{D}$ directly from $\boldsymbol{A}$ by means of an incomplete conjugation ($\boldsymbol{A}$-orthogonalization) process applied to the unit basis vectors. In this process, small elements are dropped to preserve sparsity. The underlying assumption is that most entries in $\boldsymbol{L}^{-1}$ are small in magnitude. This is true for many problems of practical interest, particularly for discretizations of partial differential equations of elliptic type. Sparsity can also be achieved by combining dropping of small entries with suitable sparse matrix orderings, such as Nested Dissection and Minimum Degree. These orderings are beneficial in that they result in smaller time and space requirements for forming and storing the preconditioner, while at the same time improving the quality of the preconditioner in a significant number of cases; see [8, 10, 11].

The basic incomplete $\boldsymbol{A}$-conjugation procedure is well-defined if $\boldsymbol{A}$ is an H-matrix (in exact arithmetic). In ref. [6], a dynamic strategy to shift negative or numerically small pivots away from zero was proposed. A better strategy is to use

a priori diagonal shift technique. However, the optimal value of the shift is difficult to estimate, and in most cases the shifting strategy produces preconditioners of poor quality.

In refs. [4, 13] a reformulation of the AINV algorithm that is breakdown free for any SPD matrix was proposed. The right-looking stabilized AINV algorithm can be written as follows.

Algorithm 1 Stabilized $\boldsymbol{A}$-conjugation algorithm

(1) *Let* $\boldsymbol{z}_i^{(0)} = \boldsymbol{e}_i \quad (1 \le i \le n)$
(2) *For* $i = 1, 2, \ldots, n$ *do*
(3) $\quad \boldsymbol{v}_i := \boldsymbol{A}\boldsymbol{z}_i^{(i-1)}$
(4) $\quad$ *For* $j = i, i+1, \ldots, n$ *do*
(5) $\quad\quad p_j^{(i-1)} := \boldsymbol{v}_i^T \boldsymbol{z}_j^{(i-1)}$
(6) $\quad$ *End do*
(7) $\quad$ *if* $i = n$ *go to (12)*
(8) $\quad$ *For* $j = i+1, \ldots, n$ *do*
(9) $\quad\quad \boldsymbol{z}_j^{(i)} := \boldsymbol{z}_j^{(i-1)} - \left(\dfrac{p_j^{(i-1)}}{p_i^{(i-1)}}\right) \boldsymbol{z}_i^{(i-1)}$
(10) $\quad$ *End do*
(11) *End do*
(12) *Let* $\boldsymbol{z}_i := \boldsymbol{z}_i^{(i-1)}$ *and* $p_i := p_i^{(i-1)}$, *for* $1 \le i \le n$. *Return* $\boldsymbol{Z} = [\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_n]$ *and* $\boldsymbol{D} = \mathrm{diag}(p_1, p_2, \ldots, p_n)$.

BLOCK ALGORITHMS

In the finite element discretization of solids and shells in particular, the corresponding matrix has a block form. It will be shown that the block approach has a considerable stabilizing effect on the preconditioner construction in difficult thin shell applications. The block form also enhances the efficiency of the code through the use of BLAS3 arithmetic if the block dimensions are large enough. However, in present applications, especially shells, most of the blocks has dimension six, therefore the program can be coded with explicit loop unrolling to get the best performance.

In the finite element context where there are several unknowns in a node, the construction of the block structure is trival. However, the block structure, or approximate block structure, can be constructed based on the graph compression procedure [3].

Consider the coefficient matrix partitioned in the block form as follows:

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \boldsymbol{A}_{1N} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \cdots & \boldsymbol{A}_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{A}_{N1} & \boldsymbol{A}_{N2} & \cdots & \boldsymbol{A}_{NN} \end{pmatrix}.$$

Here $\boldsymbol{A}_{ij}$ has order $n_i \times n_j$, where $1 \le n_i \le n$. N is the block dimension of the matrix, n is its dimension. Denote $\sum_{j<i} n_j$ by p_i (it is the offset of the i-th block).

Denote also the block rows of $\boldsymbol{A}$ by $\boldsymbol{A}_i, i = 1, \ldots, N$. Note that the diagonal matrix blocks are square symmetric positive definite matrices.

The block AINV algorithm computes block partitioned $\boldsymbol{Z}$ and $\boldsymbol{D}$ directly from $\boldsymbol{A}$ by means of a block incomplete conjugation process applied to the block unit basis vectors. In this process, blocks with small norms are dropped to preserve sparsity. In our case we used the infinity norm for the matrix blocks.

The basic left–looking block $\boldsymbol{A}$-conjugation procedure can be written as follows.

Algorithm 2 Block $\boldsymbol{A}$-conjugation algorithm:

(1) *Let* $\boldsymbol{Z}_i^{(0)} = (\boldsymbol{0}_{p_i},\ \boldsymbol{I}_{n_i},\ \boldsymbol{0}_{n-p_{i+1}})^T \quad (1 \le i \le N)$
(2) *For* $i = 1, 2, \ldots, N$ *do*
(3) *For* $j = 1, \ldots, i-1$ *do*
(4) $\boldsymbol{P}_i^{(j-1)} := \boldsymbol{A}_j^T \boldsymbol{Z}_i^{(j-1)}$
(5) $\boldsymbol{Z}_i^{(j)} = \boldsymbol{Z}_i^{(j-1)} - \boldsymbol{Z}_j^{(j-1)} \boldsymbol{P}_j^{(j-1)^{-1}} \boldsymbol{P}_i^{(j-1)}$
(6) *End do*
(6) $\boldsymbol{P}_i^{(i-1)} = \boldsymbol{A}_i^T \boldsymbol{Z}_i^{(i-1)}$
(7) *End do*
(8) *End do*
(9) *Let* $\boldsymbol{Z}_i := \boldsymbol{Z}_i^{(i-1)}$ *and* $\boldsymbol{D}_i := \boldsymbol{P}_i^{(i-1)}$, *for* $1 \le i \le N$. *Return* $\boldsymbol{Z} = [\boldsymbol{Z}_1,\ \boldsymbol{Z}_2, \ldots, \boldsymbol{Z}_N]$ *and* $\boldsymbol{D} = \mathrm{diag}(\boldsymbol{D}_1, \boldsymbol{D}_2, \ldots, \boldsymbol{D}_N)$.

The block stabilized AINV algorithm (see [4]) is obtained from Algorithm 2 by replacing the computation of $\boldsymbol{P}_i^{(j-1)}$ and $\boldsymbol{P}_i^{(i-1)}$ by $\boldsymbol{Z}_i^{(i-1)} \boldsymbol{A} \boldsymbol{Z}_i^{(j-1)}$ and $\boldsymbol{Z}_i^{(i-1)} \boldsymbol{A} \boldsymbol{Z}_i^{(i-1)}$, respectively.

NUMERICAL EXAMPLES

In this section the performance of various AINV-preconditioners is discussed. Rather comprehensive sets of test data for the performance of AINV preconditioning and its stabilized versions can also be found in [4, 6, 7, 8].

All the computations reported here have been performed on a single processor of an SGI Origin 2000 at the Center for Scientific Computing, Espoo, Finland. Convergence of the iterative process is achieved when $\|\boldsymbol{r}_i\| = \|\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}_i\| \le RTOL\|\boldsymbol{b}\| + ATOL$, and the values of $RTOL = 10^{-5}$ and $ATOL = 10^{-9}$ have been used in all test runs reported here.

Cylindrical shells

Matrices arising from the finite element discretizations of thin shells have been found to be generally difficult to solve by preconditioned iterations, especially cases which cover bending-dominated deformations of the shell.

In [5] the standard AINV approach was tested on cylindrical shell examples with modest results. The reason was identified to be pivot modifications needed to avoid

negative or zero pivots. However, it will be demonstrated here that the stabilized AINV and especially the block stabilized AINV preconditioners are successful in solving these problems.

In this paper the performance of new AINV approaches in shell problems is demonstrated by two examples. The shell element is a facet-type 3-node triangular element using the Hughes–Brezzi formulation [12] for drilling rotations. Also, an Allman-type displacement field [2] is added to the in-plane interpolation to improve the coarse mesh accuracy of the element. The plate bending part of the element is based on the stabilized MITC theory [14] and the stabilization parameter has the value 0.4.

The value of the regularizing penalty parameter γ used in the formulation of Hughes and Brezzi affects the condition number of the stiffness matrix and thus the convergence of the PCG iteration. The effect of the parameter γ on the spectral condition number as well as the convergence of the PCG iteration will be demonstrated in an forthcoming paper. In this case the value is equal to $\gamma = 10^{-3}G$, where G is the shear modulus.

The first example case corresponds to the matrix `S3RMT3M3` in the Matrix Market set CYLSHELL.[1] It corresponds to a well-known shell element test where a cylinder is loaded by two normal and equal point loads applied centrally at the opposite sides of the cylindrical surface. The length of the shell is chosen to be equal to its diameter ($L = 2R$) and the Poisson's ratio is 0.3. The shell is thin, the thickness to radius ratio is $t/R = 10^{-3}$. One octant of the shell is discretized in 1666 triangular 3-noded elements resulting in 5357 unknowns ref. [5].

Performance of the pointwise right-looking and the block left-looking AINV precontitioners in the `S3RMT3M3` case are shown in fig. 1 where the effect of different ordering strategies are also studied. The pointwise SAINV performs satisfactorily in this example, although it is consistently worse than the block version. In addition, the block version seems to be less sensitive to nodal orderings than the pointwise versions. However, there is evidence which is in favour of Jacobi or block Jacobi scalings also for the block algorithm in more difficult problems.

In this particular case the iterative part of the CG iteration with the block AINV preconditioner runs in a speed of over 150 Mflops, which is 25 % of the theoretical peak performance of the SGI Origin R12k processor (600 Mflops).

Comparison to a robust incomplete Cholesky preconditioner by Ajiz and Jennings [1] shows that the block AINV results in similar convergence rates. However, incomplete Cholesky preconditioners are difficult to parallelize. It should be noted that the no-fill IC preconditioner fails in this problem.

The second cylindrical shell example is a bent tube. In this test case neither membrane nor bending energy dominates asymptotically. Minimal number of restraints to prevent the rigid body motion of the tube are imposed. The characteristic thickness t/R is 10^{-2}, and the length to radius ratio is $L/R = 8$. Highly refined mesh at the middle part of the tube is used resulting in 7088 elements and 21498 unknowns.

Due to the different deformation characteristics, the convergence behaviour is

[1] `http://math.nist.gov/MatrixMarket/data/misc/cylshell/cylshell.html`

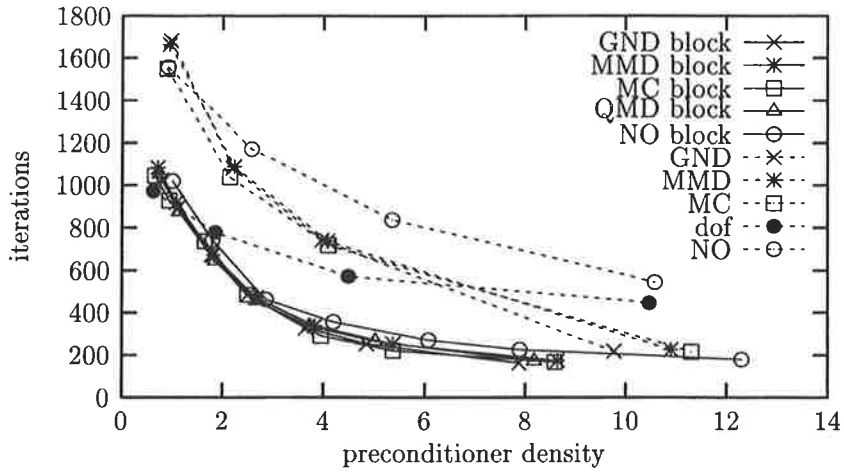


Figure 1: Test case S3RMT3M3, comparison with right-looking poitwise SAINV and the left-looking block SAINV. GND = gereralized nested dissection, MMD = multiple minumum degree, MC = multicolor, QMD = quotient minimum degree, NO = no ordering, dof = specific dofwise ordering. $n = 5357, nz = 208111, N = 938$.

completely different to the previous example. In the S3RMT3M3 problem the residual norm does not decrease until at the very end of the iterative process. In the case of the bent tube, the relative residual decreases rapidly to the level $10^{-2} - 10^{-3}$ and then stagnates and oscillates for a long time, until convergent solution is obtained. The rapid decrease at the beginning of the iteration is most probably due to the convergence of the membrane deformation state and the stagnation/oscillation phase is a sign from the difficulty of the iterative methods to capture the bending deformations.

The Gibbs-King profile reduction resulted in smallest RMS-bandwidth (741), thus it was used for the block Ajiz-Jennings type IC preconditioner. In Table 1 some statistics are shown from computations where the preconditioner density is tried to get close to one.

Since the CPU-times in RISC-architecture based machines is not independent of the workload of the computer, the CPU-timings shown are averages from three runs. The post-filtration does not have much effect on block preconditioners, therefore such results are not presented. For pointwise AINV preconditioners it has positive effect on the convergence rate of the accelerator iteration as it can be seen from Table 1.

Engine head

The last example is from thermal stress analysis of an engine head. Unstructured mesh with 151483 linear 4-node tetrahedral elements is used, which results in 143571 unknowns (number of nonzero elements in the stiffness matrix $nz = 4706073$). This

Table 1: Bent tube, $n = 21498, nz = 897056, N = 3584$. ψ, ψ_P, ρ and ρ_P are the drop-tolerance, the post-filtration tolerance, preconditioner density and the preconditioner density after post-filtration, respectively.

preconditioner	ψ	ψ_P	ρ	ρ_P	iterations	P-time	I-time	P+I
SAINV	0.035	0	1.042	1.042	1789	5.8	158	164
SAINV	0.02	0.1	1.743	0.937	1319	11.1	108	119
block SAINV	0.08	0	0.908	0.908	1428	2.0	57	59
block IC (A-J)	0.05	0	1.056	1.056	699	0.3	38	38

problem is an easy one to solve with preconditioned iterative methods. The standard AINV approach as well as IC preconditioners can be constructed without shifting. In Table 2 some results from computations are shown where the load has been a thermal load from one centigrade temperature increase. The preconditioners used in this comparison are: standard right-looking AINV, left-looking ILUT-type AINV, right-looking SAINV, and the standard no-fill IC. MMD reordering of nodes has been used in computations with the AINV algoritmn, and the Gibbs-King profile reduction resulted in smallest RMS-bandwidth (2566), thus it was used for the IC preconditioner.

In Table 2 the column LSize indicates the maximum number of entries in each column for the ILUT-type AINV approach.

In this serial implementation, the basic no-fill IC gives the fastest execution.

Table 2: Engine head, $n = 143571, nz = 4706073$.

preconditioner	ψ	ρ	iterations	P-time	I-time	P+I	LSize
Right-looking AINV preconditioners							
AINV	0.1	0.302	637	10	292	302	-
AINV	0.05	0.631	536	17	264	281	-
AINV	0.025	1.219	414	32	264	296	-
SAINV	0.1	0.291	568	15	226	241	-
SAINV	0.05	0.628	427	26	209	235	-
Left-looking AINV preconditioners							
ILUT-type AINV	0.01	0.792	798	104	506	610	20
ILUT-type AINV	0.01	1.058	733	112	448	560	30
IC(0)	-	1.000	379	2	193	195	-

CONCLUDING REMARKS

The stabilized AINV approach has been shown to result in robust and versatile preconditioners for large-scale structural analysis, even on scalar computers. For thin shell applications the block stabilized AINV gives considerable improvement in convergence rate over the pointwise version of the SAINV approach. The multiple minimum degree, the quotient minimun degree and the generalized nested dissection orderings give with the same AINV-preconditioner density similar convergence rates and the computing times are within 10 % range, however, the MMD seems to be the most robust ordering for AINV type preconditioners. Together with the block preconditioners the use of block Jacobi scaling is recommended. On cache-based architectures, the use of blocking is also beneficial in terms of performance.

REFERENCES

[1] M. A. Ajiz and A. Jennings. A robust incomplete Cholesky conjugate gradient algorithm, *Int. J. Num. Meth. Engng*, 20:949–966, 1984.

[2] D. J. Allman. A compatible triangular element including vertex rotations for plane elasticity analysis, *Computers and Structures*, 19:1–8, 1984.

[3] C. Ashcraft. Compressed graphs and the minimum degree algorithms, *SIAM J. Sci. Comput.*, 16:1404–1411, 1995.

[4] M. Benzi, J. K. Cullum and M. Tůma. Robust approximate inverse preconditioning for the conjugate gradient method, *SIAM J. Sci. Comput.*, to appear.

[5] M. Benzi, R. Kouhia and M. Tůma. An assessment of some preconditioning techniques in shell problems, *Comm. Numer. Meth. Engng*, 14:897-906, 1998.

[6] M. Benzi, C. D. Meyer and M. Tůma. A sparse approximate inverse preconditioner for the conjugate gradient method, *SIAM J. Sci. Comput.*, 17:1135–1149, 1996.

[7] M. Benzi and M. Tůma. A sparse approximate inverse preconditioner for nonsymmetric linear systems, *SIAM J. Sci. Comput.*, 19:968–994, 1998.

[8] M. Benzi and M. Tůma. Orderings for factorized sparse approximate inverse preconditioners, *SIAM J. Sci. Comput.*, to appear.

[9] M. Benzi and M. Tůma. A comparative study of sparse approximate inverse preconditioners, *Applied Numerical Mathematics*, 20:305–340, 1999.

[10] R. Bridson and W.-P. Tang. Ordering, anisotropy and factored sparse approximate inverses, *SIAM J. Sci. Comput.*, 21:867–882, 1999.

[11] M. R. Field. Improving the performance of factorized sparse approximate inverse preconditioner, Technical Report HDL-TR-98-199, Dublin, Ireland, 1998.

[12] T. J. R. Hughes and F. Brezzi. On drilling degrees of freedom, *Comp. Meth. Appl. Mech. Engng*, 72:105–121, 1989.

[13] S. A. Kharchenko, L. Yu. Kolotilina, A. A. Nikishin and A. Yu. Yeremin. A reliable AINV-type preconditioning method for constructing sparse approximate inverse preconditioners in factored form. Preprint, 1999.

[14] M. Lyly, R. Stenberg, and T. Vihinen. A stable bilinear element for the Reissner-Mindlin plate model, *Comp. Meth. Appl. Mech. Engng*, 110:343–357, 1993.

ELASTO-PLASTIC ANALYSIS OF SYMMETRICALLY LOADED THIN SHELLS OF REVOLUTION

SAMI PAJUNEN
Tampere University of Technology
Institute of Applied Mechanics
P.O. Box 589, 33101 Tampere
E-mail: sami@mohr.me.tut.fi

ABSTRACT

The paper presents an extension of a stress resultant based axisymmetric shell element [1] to the elasto-plastic computations. In the elastic range, the element exhibits exact kinematics for large displacements and rotations thus enabling the use of large increments in the global equilibrium path continuation. In order to preserve this desired property also in the elasto-plastic range, the general closest point projection algorithm [2] is utilised for the elasto-plastic stress increment computation. In the case of the used element, consistent linearisation of the stress updating algorithm results in a very simple form so that the element offers an efficient and computationally cheap approach to the large deflection elasto-plastic axisymmetric shell problems.

1. INTRODUCTION

Analysis of shell structures belongs to the category in which elegant mathematical and mechanical theories meet the interest of practical engineering in a fruitful way. From the engineering point of view, shells often provide structural solutions with high-grade strength and stiffness and relatively low weight. From the analysis point of view, shell structures stand for a challenging problem mostly due to the three-dimensional finite rotations. In the literature, there is a large number of papers devoted to the very problem of accurate and efficient shell structure analysis. For review of various shell theories and approaches, see e.g. [3,4] and references therein.

In practical engineering design, the main interest is usually directed to the load-carrying capacity of the designed structure. However, in order to follow the requirements of safe structural design, the designer must know how the structure behaves if the design load level is exceeded e.g. in the case of accident causing unexpected additional loads. This is especially true for thin shells that are known to be highly sensitive to external imperfections,

so that minor additional loads may change the load-carrying capacity and the global load-deflection path dramatically. Hence, it is of primary importance to have computational methods that are capable of analysing the structural response also in far post-critical and elasto-plastic range.

In this study, we concentrate on one particularly important shell type, namely to thin shell of revolution that appears frequently in practical applications as roofs, containers, silos etc. For the shell discretisation we adopt a stress-resultant based axisymmetric shell element that is derived using the Reissner's axisymmetric shell theory [1]. For detailed element stiffness matrix derivations we refer to [5]. For the non-linear path-following we use the arc-length method [6] with orthogonal trajectory constraint [7] and for the elasto-plastic stress updating the general idea of the closest point projection scheme is utilised [2]. As a result, we achieve an accurate and efficient computing routine suitable for handling axisymmetric shell problems with large plastic strains, large displacements and large rotations.

In what follows, in chapter 2 we first discuss some computational aspects of different types of yield functions in the stress resultant space. In chapter 3 we derive the elasto-plastic stress updating algorithm for the used element and for the chosen yield function and linearise it in order to construct the consistent tangential material stiffness matrix [8]. Finally, in chapter 4, the accuracy and efficiency of the proposed approach is numerically verified in examples of a uniformly loaded circular plate and stretched cylindrical shell, respectively.

2. EXACT AND APPROXIMATE YIELD FUNCTIONS IN THE STRESS-RESULTANT SPACE

The stress-resultant approach has certain advantages and disadvantages when applied to plastic flow problems. The approach is computationally cheap because no integration over the cross section is needed in contrast with the so called layer approach, see e.g. [9]. Consequently, the main drawback of the approach is, that partially plastic cross sections can not be handled accurately. Knowing this restriction, the analyst has to be careful not to use the element e.g. in problems exhibiting large displacements with only partially plastic cross sections. However, when restricting to thin shell applications, such problems are quite rare. The second drawback of the approach is the derivation of the yield surface that may prove to be an extremely complicated task if a strict accuracy requirement is set.

Generally, the results of the limit analysis can be used for the construction of stress resultant yield surfaces. Typically, such a yield surface has some regions exhibiting only C^0-continuity. This causes numerical problems e.g. in the case of the associative flow rule, in which normals to the yield surface has to be constructed. Numerical problems in the vicinity of such C^0-continuous 'corners' can be circumvented in many ways. In the context of the so called explicit methods the corners can be numerically rounded as in [10] so that a unique expression for the normal vector will be obtained. However, this approach is not valid when

an implicit, general closest point stress updating algorithm is adopted. When an implicit method is used, the yield surface corners can be handled by utilising a specially designed stress updating procedure [11] or by approximating the yield surface with a C^1-continuous function. In this study the latter approach is chosen because of the computational simplicity and relatively small approximation error.

In [12] a yield surface for axisymmetrically loaded thin shells of revolution is derived. In order to get rid of the abrupt corners in the surface, we approximate the exact yield surface with an inscribed ellipsoid causing conservative error as depicted in Fig 1. Following the results presented in [12], the yield surface is expressed in terms of meridional force N, circumferential force N_θ and meridional moment M_s while the influence of the other stress resultants to the plastic behaviour is negligible. In Fig. 1, the plastic interaction curves between N and M_s and between M_s and N_θ are drawn using solid line and the approximate ellipsoid yield surface is drawn using dashed line. The ellipsoid can be written as

$$N^2 - NN_\theta + N_\theta^2 + \left(\frac{4}{\beta h}\right)^2 M_s^2 - \left(\alpha\sigma_y h\right)^2 = 0 \qquad (1)$$

in which σ_y is the yield stress, $\alpha = 0.8668$ and $\beta = 1.000$ so that the approximation error is 13.3% along each axis N, N_θ and M_s, respectively, while the maximum error is ca. 23%. Fortunately, as seen from Fig.1 this relatively large error occurs in a rather narrow segment in the plane $N = 0$. By varying the values of shape parameters α and β, someway better approximations can be achieved, but simultaneously the desired character of conservativeness will be lost. An experienced analyst can use the shape parameters for tuning the algorithm for different types of problems. For instance, setting $\alpha = 0.960$ and $\beta = 1.042$ and thus allowing non-conservative yield surface in the plane $N_\theta = 0$ as depicted in Fig. 2, the maximum approximation error would drop to 12%.

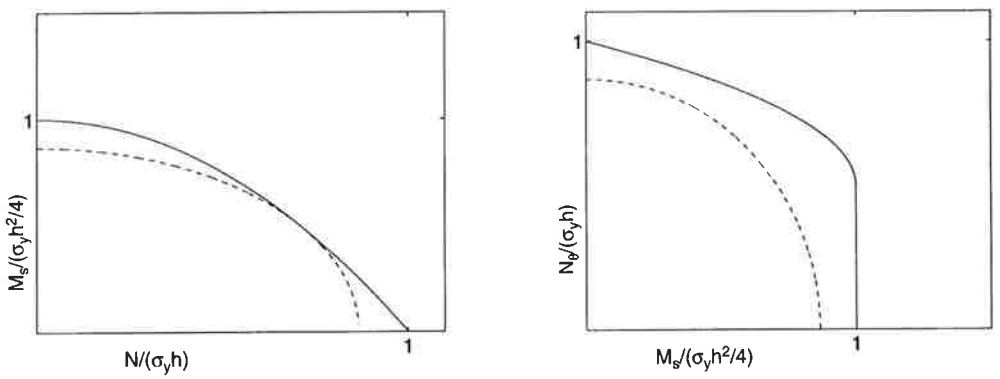


Figure 1. Exact (solid line) and conservatively approximate (dashed line) yield surfaces in $N_\theta = 0$ and $N = 0$ planes ($\alpha = 0.8668$, $\beta = 1.000$).

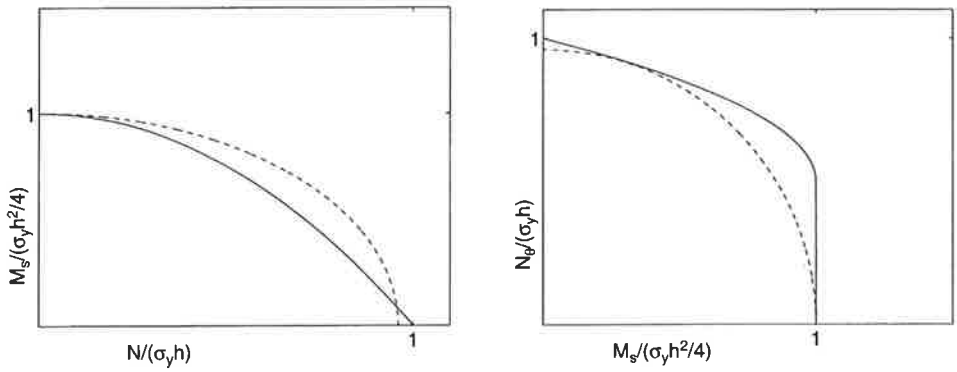


Figure 2. Exact (solid line) and approximate (dashed line) yield surfaces in $N_\theta = 0$ and $N = 0$ planes ($\alpha = 0.960$, $\beta = 1.042$).

3. ELASTO-PLASTIC STRESS UPDATING ALGORITHM

The central problem in the strain-driven elasto-plastic equilibrium path continuation procedure is calculation of the stress resultant increment $\Delta\mathbf{S}$ due to elasto-plastic strain increment $\Delta\boldsymbol{\varepsilon}$. Using the proposed smooth yield function (10), the stress resultant increment $\Delta\mathbf{S}$ can be computed efficiently by adopting the so called general closest point projection algorithm [2]. Following the standard argument, we first write the elasto-plastic constitutive equations as

$$\mathbf{S} = \mathbf{D}(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^p) \tag{2}$$

$$\dot{\boldsymbol{\varepsilon}}^p = \gamma \partial_{\mathbf{S}} f(\mathbf{S}, q) \tag{3}$$

$$\dot{q} = -\gamma E_p h \partial_q f(\mathbf{S}, q) \tag{4}$$

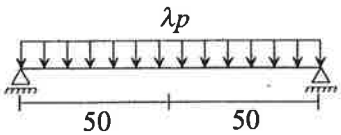
in which the operator $(\dot{})$ denotes the differentiation with respect to time. In (2) $\mathbf{D}$ is the elastic material matrix, $\boldsymbol{\varepsilon}^p$ is plastic part of the total strain vector $\boldsymbol{\varepsilon}$, γ is the so called consistency parameter and E_p is the plastic modulus that has the well known connection with the tangential modulus E_t as $E_p = EE_t/(E-E_t)$. Assuming that the plastic modulus is obtained from a uniaxial tension test exhibiting bilinear behaviour, the internal plastic variable q can be written as

$$q = -\sigma_y h = -(\sigma_0 + E_p e^p)h \tag{5}$$

in which σ_0 is the initial yield stress and e^p is the equivalent plastic strain. The general loading/unloading criteria can be expressed in the following Kuhn-Tucker conditions [13] that has to be satisfied simultaneously with the equations (2 - 4)

$$f \le 0 \tag{6}$$

$$\gamma \ge 0 \tag{7}$$



E=210000
E_t=210
ν=0.3
σ_0=360
h=2
p=2

Figure 3. Plane view of the uniformly loaded circular plate.

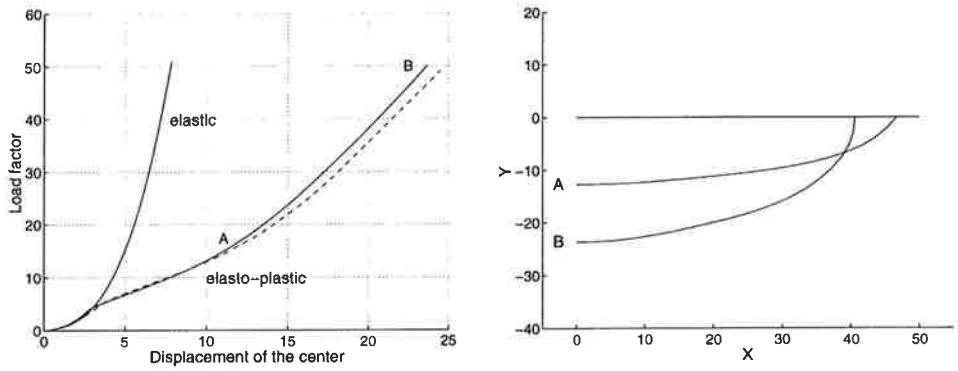


Figure 4. Load-deflection paths for the circular plate (---- = ABAQUS, —— = present approach), and some deformed configurations from the elasto-plastic analysis.

4.2 Stretched cylindrical shell

The second example is used to assess the robustness of the approach. An elastic - ideally plastic cylindrical shell is stretched with uniform end loads p = 1000. The material and geometrical properties of the shell are: E = 210000, ν = 0.3, σ_0 = 360, shell thickness h = 1, radius R = 50 and length L = 500. The computed equilibrium path depicted in Fig. 5 is obtained by using five constant arc-length increments, from which three are in the elasto-plastic range. As seen from Table 1, the convergence rate of the global equilibrium iteration is quadratic inside the convergence cone. In the global equilibrium iteration the convergence is studied using the relation $\left\|\mathbf{u}_{n+1}^{k+1}-\mathbf{u}_{n+1}^{k}\right\| < tol_3\left\|\mathbf{u}_{n+1}^{k+1}-\mathbf{u}_{n+1}^{0}\right\|$ in which $\mathbf{u}$ contains the nodal degrees-of-freedom and tol_3=10^{-10}.

$$\gamma f = 0 \tag{8}$$

The continuum differential equations (2-4) can be solved by adopting either explicit or implicit time integration method. When reaching for a continuation procedure exhibiting quadratic convergence rate with large increments, the latter approach is compulsory. Adopting backward Euler method to (2-4) and (6-8), the resulting discretised equations can be written as

$$\mathbf{S}_{n+1} = \mathbf{D}(\boldsymbol{\varepsilon}_{n+1} - \boldsymbol{\varepsilon}^{p}_{n+1}) \tag{9}$$

$$\boldsymbol{\varepsilon}^{p}_{n+1} = \boldsymbol{\varepsilon}^{p}_{n} + \Delta\gamma \partial_{S} f(\mathbf{S}_{n+1}, q_{n+1}) \tag{10}$$

$$q_{n+1} = q_{n} - \Delta\gamma E_{p} h \partial_{q} f(\mathbf{S}_{n+1}, q_{n+1}) \tag{11}$$

$$f(\mathbf{S}_{n+1}, q_{n+1}) \le 0 \tag{12}$$

$$\Delta\gamma \ge 0 \tag{13}$$

$$\Delta\gamma f(\mathbf{S}_{n+1}, q_{n+1}) = 0 \tag{14}$$

in which the subscript denotes the increment number and $\Delta\gamma$ denotes the value of the consistency parameter cumulated during one increment. Equations (9 - 14) can be solved using e.g. Newton-Raphson iteration, but because of the inequalities (12) and (13), the solution procedure has to applied in two steps. First a fully elastic trial step with $\Delta\gamma = 0$ is taken resulting in $\mathbf{S}^{trial}_{n+1}$. If the yield criterion (12) is satisfied, also the other equations (9 - 14) are fulfilled and the continuation can be proceeded. However, if the yield criterion (12) is violated, the step is elasto-plastic with $\Delta\gamma > 0$ and the resulting vector $\mathbf{S}_{n+1}$ has to be computed iteratively. In order to derive the iterative equations, let us first divide (11) with $(-E_p h)$, and denote the unbalance of (10) and (11) with a residual vector $\mathbf{R}$

$$\mathbf{R} = \begin{Bmatrix} -\boldsymbol{\varepsilon}^{p}_{n+1} + \boldsymbol{\varepsilon}^{p}_{n} \\ (-q_{n+1} + q_{n}) / (-E_{p} h) \end{Bmatrix} + \Delta\gamma \begin{Bmatrix} \partial_{S} f_{n+1} \\ \partial_{q} f_{n+1} \end{Bmatrix} \tag{15}$$

in which

$$f_{n+1} = f(\mathbf{S}_{n+1}, q_{n+1}) \tag{16}$$

so that the linearisation of (15) yields

$$\mathbf{R}^{k}_{n+1} + \begin{Bmatrix} -\Delta\boldsymbol{\varepsilon}^{p\,k}_{n+1} \\ -\Delta e^{p\,k}_{n+1} \end{Bmatrix} + \Delta\gamma \begin{Bmatrix} \partial_{SS} f^{k}_{n+1} \Delta\mathbf{S} \\ \partial_{qq} f^{k}_{n+1} \Delta q \end{Bmatrix} + \Delta^{2}\gamma \begin{Bmatrix} \partial_{S} f^{k}_{n+1} \\ \partial_{q} f^{k}_{n+1} \end{Bmatrix} = \mathbf{0} \tag{17}$$

in which the superscript denotes the iteration cycle number. From (17) it is straightforward to derive the iterative formulas for plastic strain vector $\boldsymbol{\varepsilon}^{p}$ and equivalent plastic strain e^{p} that can be written as

$$\begin{Bmatrix} \Delta\boldsymbol{\varepsilon}^{p\,k}_{n+1} \\ \Delta e^{p\,k}_{n+1} \end{Bmatrix} = \begin{bmatrix} \mathbf{D}^{-1} & \mathbf{0} \\ \mathbf{0} & (E_{p} h)^{-1} \end{bmatrix} \Xi^{k}_{n+1} \left\{ \mathbf{R}^{k}_{n+1} + \Delta^{2}\gamma \begin{Bmatrix} \partial_{S} f^{k}_{n+1} \\ \partial_{q} f^{k}_{n+1} \end{Bmatrix} \right\} \tag{18}$$

in which the matrix Ξ is

$$\Xi = \begin{bmatrix} \Xi_{11} & \mathbf{0} \\ \mathbf{0} & \Xi_{22} \end{bmatrix} := \begin{bmatrix} \mathbf{D}^{-1} + \Delta\gamma\partial_{SS}f & \mathbf{0} \\ \mathbf{0} & \left(E_p h\right)^{-1} + \Delta\gamma\partial_{qq}f \end{bmatrix}^{-1} \tag{19}$$

Similar formula for the consistency parameter γ can be obtained by linearising (16) resulting in

$$f^k_{n+1} + \partial_S f^k_{n+1}\Delta\mathbf{S} + \partial_q f^k_{n+1}\Delta q = 0 \tag{20}$$

Taking into account that $\Delta\mathbf{S} = -\mathbf{D}\Delta\boldsymbol{\varepsilon}^p$ and $\Delta q = -E_p h\Delta e^p$ and substituting (18) into (20) yields, after some manipulation the following result

$$\Delta^2\gamma = \frac{f^k_{n+1} - \mathbf{R}^{Tk}_{n+1}\Xi^k_{n+1}\begin{Bmatrix}\partial_S f^k_{n+1} \\ \partial_q f^k_{n+1}\end{Bmatrix}}{\begin{Bmatrix}\partial_S f^k_{n+1} \\ \partial_q f^k_{n+1}\end{Bmatrix}^T \Xi^k_{n+1}\begin{Bmatrix}\partial_S f^k_{n+1} \\ \partial_q f^k_{n+1}\end{Bmatrix}} \tag{21}$$

Updated values of the plastic strain vector $\boldsymbol{\varepsilon}^p$, equivalent plastic strain e^p, consistency parameter γ and stress resultant vector $\mathbf{S}$ are obtained using (18), (21) and (9) together with the updating procedure that can be expressed as

$$\boldsymbol{\varepsilon}^{p\,k+1}_{n+1} = \boldsymbol{\varepsilon}^{p\,k}_{n+1} + \Delta\boldsymbol{\varepsilon}^{p\,k}_{n+1} \qquad \& \qquad \boldsymbol{\varepsilon}^{p\,0}_{n+1} = \boldsymbol{\varepsilon}^{p}_{n} \tag{22}$$

$$e^{p\,k+1}_{n+1} = e^{p\,k}_{n+1} + \Delta e^{p\,k}_{n+1} \qquad \& \qquad e^{p\,0}_{n+1} = e^{p}_{n} \tag{23}$$

$$\Delta\gamma^{k+1} = \Delta\gamma^k + \Delta^2\gamma^k \qquad \& \qquad \Delta\gamma^0 = 0 \tag{24}$$

The iteration is performed until convergence, i.e. until the yield function and residual vector norm values satisfy the conditions

$$f^k_{n+1} < tol_1 \qquad \& \qquad \|\mathbf{R}\|^k_{n+1} < tol_2 \tag{25}$$

in which tol_1 and tol_2 are user-specified suitably small positive numbers.

After the stress increment iteration has converged, the global equilibrium path continuation can be proceeded. In order to maintain the quadratic rate of convergence, the tangential material stiffness matrix $d\mathbf{S}/d\boldsymbol{\varepsilon}$ has to be constructed so that it is consistent with the stress updating algorithm. As shown in [2], using the closest point projection algorithm in stress updating, a closed form solution for the matrix $d\mathbf{S}/d\boldsymbol{\varepsilon}$ can be obtained. Using the axisymmetric shell element [1], the closed form solution can be derived in similar way than in [2] resulting in

$$\left.\frac{d\mathbf{S}}{d\boldsymbol{\varepsilon}}\right|_{n+1} = \Xi_{11\,n+1}\left(\mathbf{I} - \partial_S f_{n+1}\mathbf{n}^T_{n+1}\right) \tag{26}$$

in which **I** is the identity matrix and the matrix Ξ_{11} and the derivatives of the yield function (1) with respect to stress resultant vector **S** and internal parameter q can be written in closed form as

$$\Xi_{11} = \frac{Eh}{1-\nu^2}\begin{bmatrix} \frac{2\Delta\gamma Eh+1}{c_1} & \frac{\Delta\gamma Eh+\nu}{c_1} & 0 & 0 & 0 & 0 \\ & \frac{2\Delta\gamma Eh+1}{c_1} & 0 & 0 & 0 & 0 \\ & & \frac{1-\nu}{2} & 0 & 0 & 0 \\ & & & \frac{-1}{c_2} & \frac{-\nu}{c_2} & 0 \\ & & & & \frac{\frac{8\Delta\gamma Eh}{3\beta^2}+1}{c_2} & 0 \\ & \textit{symmetric} & & & & \frac{h^2(1-\nu)}{144} \end{bmatrix} \tag{27}$$

$$c_1 = 1+\Delta\gamma\frac{Eh}{1-\nu^2}(3\Delta\gamma Eh-2\nu+4) \tag{28}$$

$$c_2 = -\frac{32\Delta\gamma E}{\beta^2 h(1-\nu^2)}-\frac{12}{h^2} \tag{29}$$

$$\partial_S f^T = \begin{bmatrix} 2N-N_\theta & 2N_\theta-N & 0 & \frac{32}{\beta^2 h^2}M_s & 0 & 0 \end{bmatrix} \tag{30}$$

$$\partial_q f = 2\alpha^2 h\sigma_y \tag{31}$$

4. NUMERICAL VERIFICATION

4.1 Uniformly loaded circular plate

The first example is used to assess the accuracy of the presented approach. The circular plate, depicted in Fig. 3 is loaded uniformly and the associated elasto-plastic equilibrium path is computed until the maximum displacement is ca. one fourth of the whole span length. The plate is discretised with 10 three noded elements with equal length and the material is modelled using the conservative yield surface approximation (1) with $\alpha = 0.8668$ and $\beta = 1.000$. The computed solution, depicted in Fig.4 is compared with the one obtained by using an axisymmetric shell element SAX2 and von Mises yield criterion within the general purpose finite element package ABAQUS. As seen from Fig. 4, excellent agreement between the solutions is found.

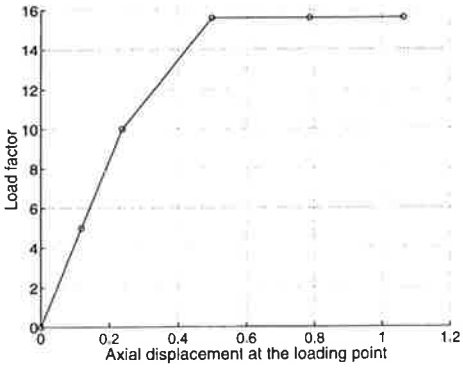


Figure 5. Load deflection path for the elastic - ideally plastic stretched cylindrical shell.

Table 1. Convergence of the equilibrium iteration in an analysis of a stretched cylindrical shell.

Iteration	*Value of the residual norm* $\|\mathbf{u}_{n+1}^{k+1} - \mathbf{u}_{n+1}^{k}\|$ *at increment number*				
	1 (elastic)	*2 (elastic)*	*3*	*4*	*5*
1	3.30E-1	3.30E-1	4.14E-1	5.95E-2	1.10E-1
2	2.00E-4	1.81E-4	3.37E-1	6.55E-2	1.55E-1
3	4.17E-7	3.58E-7	4.84E-2	4.79E-2	2.27E-1
4	1.08E-10	8.75E-11	4.23E-2	7.86E-2	2.39E-2
5	7.32E-14	6.01E-14	8.51E-2	2.75E-3	7.57E-5
6			2.51E-3	2.50E-5	6.70E-7
7			3.38E-6	1.09E-7	5.30E-9
8			5.81E-8	1.08E-9	5.10E-11
9			5.10E-10	1.05E-11	
10			4.78E-12		

5. CONCLUSIONS

In the present investigation numerical treatment of large plastic deformations occurring in thin shells of revolution is considered. The computing routines derived in the paper can be used in large displacement and large rotation problems that appear e.g. when a designed structure is analysed until breakdown. The routines are computationally cheap because for the used element the consistent linearisation of the stress updating algorithm results in a very simple form as seen from equations (26 -31). The major drawback of the approach is the approximate character of the chosen yield criterion (1). However, if the analyst abandons the requirement of conservativeness, the approximation error can be reduced significantly, as shown in Figs. 1 and 2.

ACKNOWLEDGEMENT

The author would like to thank Prof. Markku Tuomala for many helpful discussions.

REFERENCES

1. Reissner E. On finite axi-symmetrical deformations of thin elastic shells of revolution. *Comp. Mech.* 1989; **4**:387-400.
2. Simo JC, Hughes TJR. *Computational Inelasticity*. Springer-Verlag: New York, 1998.
3. Eberlain R, Wriggers P. Finite element concepts for finite elastoplastic strains and isotropic stress response in shells: theoretical and computational analysis. *Comp. Meth. Appl. Mech. Eng.* 1999; **171**:243-279.
4. Simo JC, Fox DD, Rifai MS. On a stress resultant geometrically exact shell model. Part III: Computational aspects of the nonlinear theory. *Comp. Meth. Appl. Mech. Eng.* 1990; **79**:21-70.
5. Pajunen S, Tuomala M. Post-buckling analysis of plates and shells. Proc. of the 6th Finnish Mechanics Days, J. Aalto, T. Salmi (eds.), 1997; 237-251.
6. Riks E. The application of Newton's method to the problem of elastic stability. *J. Appl. Mech.* 1972; **39**:1060-1066.
7. Fried I. Orthogonal trajectory accession to the nonlinear equilibrium curve. *Comp. Meth. Appl. Mech. Eng.* 1984; **47**:283-297.
8. Hughes TJR, Taylor RL. Unconditionally stable algorithms for quasi-static elasto/viscoplastic finite element analysis. *Computers & Structures* 1978; **8**:169-173.
9. Argyris JH, Balmer H, Kleiber M, Hindenlang U. Natural description of large inelastic deformations for shells of arbitrary shape - application of TRUMP element. *Comp. Meth. Appl. Mech. Eng.* 1980; **22**:361-389.
10. Nayak GC, Zienkiewicz OC. Elasto-plastic stress analysis. A generalization for various constitutive relations including strain softening. *Int. J. Num. Meth. Eng.* 1972; **5**:113-135.
11. Simo JC, Kennedy JG, Govindjee S. Non-smooth multisurface plasticity and viscoplasticity. Loading/unloading conditions and numerical algorithms. *Int. J. Num. Meth. Eng.* 1988; **26**:2161-2185.
12. Drucker DC, Shield RT. Limit analysis of symmetrically loaded thin shells of revolution. *J. Appl. Mech.* 1959; **26**:61-68.
13. Simo JC, Ortiz M. A unified approach to finite deformation elastoplasticity based on the use of hyperelastic constitutive equations. *Comp. Meth. Appl. Mech. Eng.* 1985; **49**:221-245.

ANALYSIS OF ELASTO-PLASTIC BODY BY THE FINITE ELEMENT METHOD

M. REIVINEN, J. FREUND
Helsinki University of Technology
Laboratory of Theoretical and Applied Mechanics
P.O.Box 1100, FIN-02015 HUT

ABSTRACT

A finite element scheme for the analysis of elasto-plastic material is presented. In the method, the approximation of displacement is continuous whereas the approximations of the stress and plasticity parameter are discontinuous. As numerical examples, two simple deformation modes of elastic body, and necking of an elasto-plastic strip in tension are studied.

INTRODUCTION

This study has started from the need to develop numerical methods and material models for the analysis of elasto-plastic material in metal forming processes. Some aspects of the forming processes and analysis are discussed for example in [1], [2], [3] and [4]. Often, the displacements, rotations, and strains are large, and consequently, the typical formulation applied in the analysis may be Total Lagrangian using the second Piola-Kirchhoff stress measure and Green-Lagrange strain measure. Alternatively, the Updated Lagrangian with Cauchy stress and logarithmic strain can be used. The use of the most general large strain formulation will always be correct, but the use of a more restrictive formulation may be computationally more effective. Here, we restrict ourselves to the simplified forms of the governing equations of an elasto-plastic material and consider an incremental formulation in an Eulerian frame of reference. For demonstration, some simple cases of elastic and elasto-plastic bodies under forced displacement are analyzed.

The boundary value problem for a body $V \subset \mathbb{R}^n$ ($n = 2$ or $n = 3$) can be written as

$$\sigma_{ij,j} - q_i = 0 \quad x_i \in V\,, \tag{1}$$

$$u_i - \hat{u}_i = 0 \quad x_i \in A_D\,, \tag{2}$$

$$t_i - \hat{t}_i = 0 \quad x_i \in A_N \,, \tag{3}$$

in which σ_{ij} is the Cauchy stress tensor, q_i is the given force per unit mass, u_i is the displacement, $t_i = \sigma_{ij} n_j$ is the traction, n_j is the unit outward vector to the boundary, and the comma notation is used for partial derivatives i.e. $u_{i,j} \equiv \partial u_i / \partial x_j$. The parts of the boundary, on which the Dirichlet (A_D) and Neumann (A_N) conditions are given, satisfy $A_D \cap A_N = \varnothing$ and $\bar{A}_D \cup \bar{A}_N = A$ (the overbar means closure).

INCREMENTAL APPROACH

Due to the nature of the stress-strain relation of elasto-plastic material, an incremental step-by-step approach is natural. For this purpose, the external loading or displacement data is parameterized with $t \in [0,1]$ so that $t = 0$ corresponds to the zero loading and the value $t = 1$ to the true loading. In the incremental scheme, we start from the known solution at $t = 0$, seek the solution at $t + \Delta t$, and repeat the same until $t = 1$. In what follows, overbar ($\bar{x}_i$) means a quantity of the known solution, hat ($\hat{x}_i$) a given quantity, and tilde ($\tilde{x}_i$) an intermediate value during a step. The delta notation is used for the increment of a quantity in a step (Δx_i).

The incremental strain and the rotation tensor expressions are taken to be

$$\Delta\varepsilon_{ij} = \frac{1}{2}(\Delta u_{i,j} + \Delta u_{j,i} + \Delta u_{i,k}\Delta u_{j,k}) \,, \tag{4}$$

$$\Delta R_{ij} = \delta_{ij} + \Delta\omega_{ij} + \frac{1}{2}\Delta\omega_{ik}\Delta\omega_{kj} \tag{5}$$

in which $\Delta u_i(\bar{x}_j) = x_i(\bar{x}_j) - \bar{x}_i$, $\Delta\omega_{ij} = (\Delta u_{i,j} - \Delta u_{j,i})/2$, and the derivatives are with respect to the coordinates with an overbar. Strain has been defined so that the effect of a displacement field representing an incremental rotation vanishes. The rotation tensor can be taken as the truncated Taylor series of the exact one.

The incremental stress-strain and plastic strain-stress relationships are taken to be

$$\sigma_{ij} = \Delta R_{ik}\Delta R_{jl}\bar{\sigma}_{kl} + \lambda\Delta\varepsilon_{kk}\delta_{ij} + 2\mu(\Delta\varepsilon_{ij} - \Delta\varepsilon_{ij}^p) \,, \tag{6}$$

$$\Delta\varepsilon_{ij}^p = \Delta L\frac{\partial f}{\partial\sigma_{ij}} = \Delta L b_{ij} \,, \tag{7}$$

where λ and μ are the material parameters. In (6) we have considered the motion of a material element as result of translation, rotation and deformation, of which only the two last parts affect the stress expression. In the first term of (6), the initial stress field

following the material is rotated to the fixed coordinate system used in the description. In the relationship (7) giving the increment of the plastic strain $\Delta\varepsilon_{ij}^{p}$ as function of stress, ΔL is the plastic parameter increment. The material is assumed to follow the von Mises condition according to which the boundary of the elastic region of the stress-space is given by

$$f(\sigma_{ij}) \equiv g(\sigma_{ij}) - k = \sqrt{\sigma'_{ij}\sigma'_{ij}} - k = 0. \tag{8}$$

Above, σ'_{ij} is the deviatoric part of the stress tensor i.e. $\sigma'_{ij} = \sigma_{ij} - \sigma_{kk}\delta_{ij}/3$ and k the radius of the elastic region in the stress space.

FINITE ELEMENT METHOD

The basic unknowns of the finite element formulation are the coordinates of the body $x_i(\overline{x}_j, \Delta t)$, stress $s_{ij}(\overline{x}_j, \Delta t)$ (symbol σ_{ij} refers to expression (6)) and plastic multiplier $L(\overline{x}_j, \Delta t)$. Approximations of all these quantities are chosen to be the same degree polynomials. For the stress components and the plastic multiplier, the approximations are discontinuous, i.e.

$$U(V) = \{\, u \,|\, \kappa \in P_k(\kappa) \quad \forall \kappa \in K \}, \tag{9}$$

where K denotes the element division of the body at t, κ an element, and $P_k(\kappa)$ the set of polynomials of degree k on element κ. The mesh parameter is defined as $h = \max_K \operatorname{diam}(\kappa)$. The approximations of the displacement field components are continuous, i.e. $W(V) = U(V) \cap C^0(V)$.

The weak form of a typical loading step can be written as: find $(x_i, s_{ij}, L) \in (W)^n \times (U)^{n \cdot n} \times U$ ($n = 2$ or $n = 3$) such that

$$(\delta\Delta u_{i,j}, \sigma_{ij})_V - (\delta\Delta u_i, q_i)_V - \langle \delta\Delta u_i, \hat{t}_i \rangle_{A_N} +$$

$$-((1-\alpha)\delta\Delta L, f)_V + (\alpha\,\delta\Delta L, 2\mu\,\Delta L)_V + (\alpha\,\delta\Delta\varepsilon_{ij}\, b_{ij}, 2\mu\,\Delta L)_V +$$

$$-\langle \delta\Delta u_i, t_i \rangle_{A_D} + \langle \delta t_i, u_i - \hat{u}_i \rangle_{A_D} + \langle C\,2\mu h^{-1}\,\delta\Delta u_i, u_i - \hat{u}_i \rangle_{A_D} +$$

$$+(\delta s_{ij}, s_{ij} - \sigma_{ij})_V = 0. \tag{10}$$

for all $(\delta x_i, \delta s_{ij}, \delta L) \in (W)^n \times (U)^{n \cdot n} \times U$. Above, $(u,v)_V$ and $\langle u,v \rangle_A$ mean the $L_2(V)$ and $L_2(A)$ inner products, and the shorthand notations used are $\Delta u_i = x_i - \overline{x}_i$, $\delta\Delta u_i = \delta x_i$, $\delta\Delta\varepsilon_{ij} = (\delta x_{i,j} + \delta x_{j,i})/2$, $\Delta L = L - \overline{L}$ and $\delta\Delta L = \delta L$. The information about the plastic and elastic regions is included with the aid of the mapping $\alpha : \mathbb{R}^n \to \{0,1\}$

$$\alpha = \begin{cases} 1 & f(\sigma_{ij}) < 0, \\ 0 & f(\sigma_{ij}) \geq 0. \end{cases} \tag{11}$$

The yield criterion is satisfied by the terms on the second row: in the elastic region ($\alpha = 1$) the two last terms imply that $\Delta L = 0$. In the plastic region ($\alpha = 0$), the first term on the second row implies $f = 0$. The terms on the third row act so that the boundary condition on the Dirichlet part of the boundary A_D is satisfied in the weak sense. Constant C determines the precision with which this condition is satisfied.

The discrete problem corresponding to a loading step is solved by using a segregated outer-inner iteration: in the inner iteration, the Newton method is first applied only to variables x_i and L by keeping s_{ij} fixed. After that the stress s_{ij} is calculated by keeping x_i and L, fixed. The inner iteration is repeated until the solution does not change any more. This constitutes the outer loop of the iteration.

NUMERICAL EXAMPLES

We consider first two simple deformation modes (Figure 1) without the additional complexity introduced by plastic behavior. The exact solutions to these problems are known and therefore they serve well for illustrative purposes. In the examples, the mesh consists of 3×3 bilinear elements and the given displacement field is imposed on the boundary.

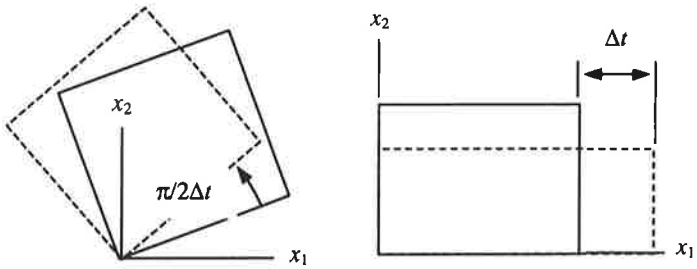


Figure 1. A body under pure rotation (left) and stretch (right).

Rotation

In the pure rotation example (Figure 1), the initial stress is given by $\bar{\sigma}_{11} = 1$ and $\bar{\sigma}_{ij} = 0$ otherwise, and the displacement field increments (enforced on the boundary) are given by

$$\begin{aligned} \Delta u_1 &= \bar{x}_1 \cos(\pi \Delta t / 2) - \bar{x}_2 \sin(\pi \Delta t / 2), \\ \Delta u_2 &= \bar{x}_1 \sin(\pi \Delta t / 2) + \bar{x}_2 \cos(\pi \Delta t / 2). \end{aligned} \tag{12}$$

In this case, the strain as calculated from (4) vanishes and, after rotation $t=1$, the initial stress is transformed to $\sigma_{22}=1$ and $\sigma_{ij}=0$ otherwise. The RMS error of the stress obtained by the finite element method is shown in Table 1 for some selections of Δt. The results indicate that the approximation (5) is accurate enough for the practical purposes.

Stretch

In the pure stretch example (Figure 1), the displacement field increments

$$\Delta u_1 = \bar{x}_1 \Delta t\,, \qquad \Delta u_2 = -\bar{x}_2 \Delta t \tag{13}$$

are such that rotations of the material elements vanish and the areas (volumes) of the material elements are conserved. The incremental strains are given by $\Delta\varepsilon_{11}=\Delta t+\Delta t^2/2$ and $\Delta\varepsilon_{22}=-\Delta t+\Delta t^2/2$ and the non-zero stress components by $\sigma_{11}=2\mu t(1+\Delta t/2)$ and $\sigma_{22}=-2\mu t(1-\Delta t/2)$. Therefore, the relative errors in the stress components are $\Delta t/2$. The calculated stress is compared with the exact stress ($\Delta t\to 0$) in Table 1 as function of Δt.

Table 1. Relative error $|\sigma''_{ij}-\sigma_{ij}|_2 / |\sigma_{ij}|_2$ in the stress as function of the increment Δt of the control parameter at $t=1$.

Δt	Rotation	Stretch
0.2	$1.2\cdot10^{-2}$	$2.3\cdot10^{-1}$
0.1	$1.5\cdot10^{-3}$	$1.1\cdot10^{-1}$
0.05	$1.9\cdot10^{-4}$	$2.3\cdot10^{-2}$

Necking of an elasto-plastic strip

As the final example we consider necking of an elasto-plastic strip. Figure 2 shows a tensile test simulation of metal specimen with a two-dimensional model based on the plane-stress assumption. The right end of the strip is subjected to forced displacement which is increased gradually during the experiment. Yielding starts at points where the exterior boundary has a non-convex corner, and proceeds rapidly to the right end. After that the yielding takes place only in the elements near the right end of the strip.

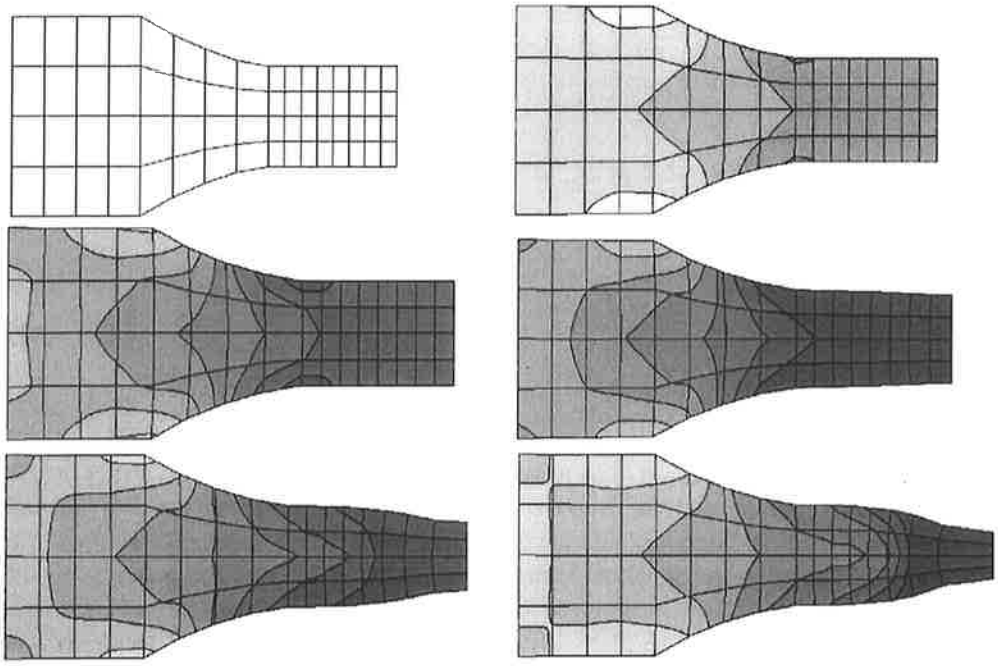


Figure 2. Level curves of the von Mises stress measure during the necking simulation.

CONCLUSIONS

The goal of the numerical method presented is to serve in the future as a tool in the analysis of metal forming processes, such as drawing, strip rolling, extrusion, and forging. To assess the validity of integration procedure of stresses and strains, comparisons to some analytical solutions were performed. The second example simulates tension test, which is commonly used to determine parameters for materials to be used in drawing processes. The results show that the numerical method works also when the plastic behavior has to be accounted for.

REFERENCES

[1] M. Reivinen, J. Freund: *On Finite Element Analysis of Large Deformation Metal Forming Processes*, in print, Espoo (2000).

[2] M.A. Crisfield, *Non-linear Finite Element Analysis of Solids and Structures*, Volume 2: Advanced Topics, 494 pp., Wiley, (1997).

[3] J.C. Simo, T.J.R Hughes, *Computational inelasticity*, Springer-Verlag, Berlin, (1998).

[4] R.H. Wagoner, J-L. Chenot, *Fundamentals of metal forming*, Wiley, (1997).

A THEORY OF FINITE ELEMENTS FOR THE STIFFNESS, MIXED AND FLEXIBILITY METHODS

Pentti Tuominen
Department of Civil Engineering, University of Oulu

ABSTRACT

Paper is dealing with generation of finite elements. Introducing concepts of complement vectors and matrices it is possible to present static equation of a two-noded analytical element with one expression, which is valid for the stiffness, mixed and flexibility methods. An other equation is needed for handling inside of the element. A conical shell element is considered as an example.

In addision of analytical elements it's considered how to use equations when presumed shape functions are as base of generating a two-noded element. A difficulty in applying the equations for more-noded elements is discussed briefly.

THEORETICAL CONSIDERATIONS

A two-noded finite element presented in Figure 1 is considered

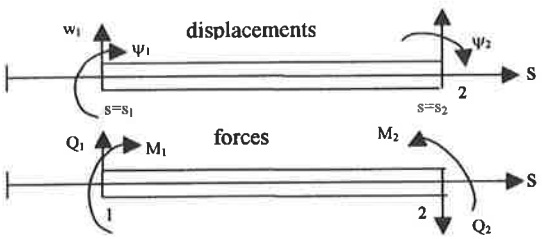


Figure 1. A two-noded structural element with displacements and forces shown as positive.

It is assumed that we know the governing differential equation of the problem and its analytic solution. Thus we have expressions for all quantities of Fig. 1. The homogenous parts of these expressions contain four integration constants $\mathbf{c} = \{c_1\ c_2\ c_3\ c_4\}$. In addition of this eight particular solution terms are expected to been known (some can be zeros).

Four quantities of Figure 1 are components of the state vector **a.** Its complement vector **b** contains the other four quantities. For these vectors can now be written expressions

$$\mathbf{a} = \mathbf{A}\mathbf{c} + \mathbf{a}_p \tag{1}$$

and

$$\mathbf{b} = \mathbf{B}\mathbf{c} + \mathbf{b}_p, \tag{2}$$

where vectors $\mathbf{a}_p$ and $\mathbf{b}_p$ contain the particular solution terms and matrices **A** and **B** coefficients of the integration constants **c**. Also matrices **A** and **B** are said to be each others complements.

Equation 1 gives the expression

$$\mathbf{c} = \mathbf{A}^{-1}\mathbf{a} - \mathbf{A}^{-1}\mathbf{a}_p \tag{3}$$

for integration constants. The substitution of it to Equation 2 will produce the final interaction equation

$$\mathbf{b} = \mathbf{B}\mathbf{A}^{-1}\mathbf{a} + \mathbf{b}_p - \mathbf{B}\mathbf{A}^{-1}\mathbf{a}_p \tag{4}$$

between state vectors **a** and **b**.

In the case of Figure 1 Equation 4 is a statical equation. Its physical meaning depends on the content of the vectors and matrices of it. If vector **b** contains forces Q_i and M_i ($i = 1,2$) the interaction matrix $\mathbf{B}\mathbf{A}^{-1}$ is the stiffness matrix and its inverse $\mathbf{A}\mathbf{B}^{-1}$ the flexibility matrix (if $\mathbf{B}^{-1}$ exists) of the element. If **b** contains quantities at node 2 matrices $\mathbf{B}\mathbf{A}^{-1}$ and $\mathbf{A}\mathbf{B}^{-1}$ are the two transfer matrices between nodes. The part $\mathbf{b}_p - \mathbf{B}\mathbf{A}^{-1}\mathbf{a}_p$ in Equation 4 can be called a loading vector broaderly suitable appellations are a source vector or a forcing vector.

Equations 3 and 4 are valid if matrix **A** is regular (its inverse $\mathbf{A}^{-1}$ exists). If **A** is singular, it has to be reduced to a smaller non-singular one. (Also its complement **B** and state vectors **a** and **b** must then be reduced.) An example of this replacing vectors and matrices with smaller ones is given in [1].

When we are calculating vector $\mathbf{e}(s) = \{w\ \psi\ Q\ M\}$ at a point s inside our element we have available Equation

$$\mathbf{e}(s) = \mathbf{E}(s)\mathbf{c} + \mathbf{e}_p(s) \tag{5}$$

In it $\mathbf{E}(s)$ is the coefficient matrix of integration constants $\mathbf{c}$ at the point $s_1 \leq s \leq s_2$. Vector $\mathbf{e}_p$ contains the particular solution terms at the point. Substitution of $\mathbf{c}$ from Eq. 3 into Eq. 5 will give

$$\mathbf{e}(s) = \mathbf{E}(s)\mathbf{A}^{-1}\mathbf{a} + \mathbf{e}_p(s) - \mathbf{E}(s)\mathbf{A}^{-1}\mathbf{a}_p \tag{6}$$

for computing vector $\mathbf{e}$ at any point s. When using Equation 6 in connection of the transfer matrix it coincides with Equation 4, when the element length $\mathbf{l} = \mathbf{s}_2 - \mathbf{s}_1$, is replaced with $\mathbf{s} - \mathbf{s}_1$.

Equation 4 has also an interchangeable form

$$\mathbf{b} = [\mathbf{B}\mathbf{A}^{-1}]\,\mathbf{\Gamma}^{-1}\mathbf{a} + \mathbf{b}_p - [\mathbf{B}\mathbf{A}^{-1}]\,\mathbf{\Gamma}^{-1}\mathbf{a}_p, \tag{4a}$$

where $\mathbf{I}$ is the unit matrix of the same size with $\mathbf{A}$ and $\mathbf{B}$. As a comment we can say that any interaction matrix $[\mathbf{B}\mathbf{A}^{-1}]$ has a complement, unit matrix $\mathbf{I}$. When substituting this into Equation 3 we will obtain the result

$$\mathbf{c} = \mathbf{\Gamma}^{-1}\mathbf{a} - \mathbf{\Gamma}^{-1}\mathbf{a}_p = \mathbf{a} - \mathbf{a}_p \tag{3a}$$

and can have a further interpretetion: Generation of a finite element can be regarded as procedure, where mathematical integration constants $\mathbf{c}$ are replaced with physical ones given in Equation 3a.

NUMERICAL CONSIDERATIONS

As an example is considered an isotropic conical shell element under its own weight g = 3,75 kN/m^2. Shell is fixed at the inner boundary and free at the outer one. Measures of the shell and its data are presented in Figure 2.

Young's modulus $\mathbf{E}$ = 10000 kN/m^2
Poisson's ratio ν = 0,20
Bending stiffness $\mathbf{B}$ = 2,9296875 kN/m

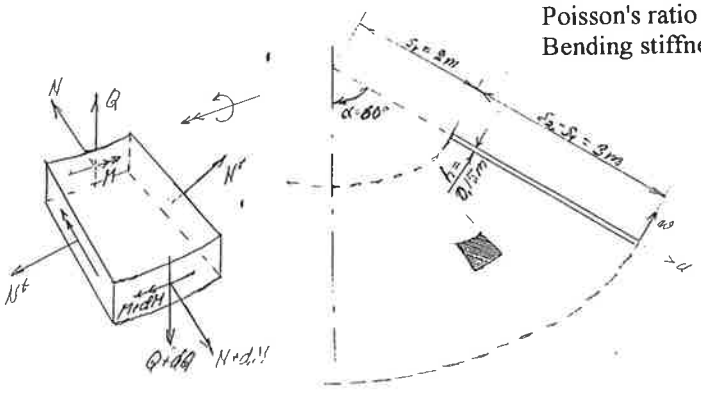


Figure 2. Conical shell considered as example its measures and data. A shell element with stress resultants and couples as positive.

When handling a conical shell six integration constants c_i ($i = 1...6$) are needed. Coefficients of integration constants and particular solution terms for own weight g are given in Table 1. [2, p.11, pp.15-19]

Table 1. Content of the matrix **E** and particular solution for own weight g (see Eq. 6), when handling a conical shell. The argument **y** of Thomson functions (ber etc.) of zero and second kind is $y = 2\eta\sqrt{s}$, where **s** is the distance from the apex and conical parameter $\eta = [12(1 - \gamma]^{1/4}(h\tan\beta)^2$. For coefficients of constants c_3, c_5 and c_6 Thomson functions ber are replaced with bei, ker and kei. Constants $\mathbf{c_1}$ and $\mathbf{c_4}$ have an obvious physical meaning. Terms: $f = 2 + \nu - 1{,}5/\cos^2\alpha$ and $\varphi = (1-2\nu\sin^2\alpha)/\cos\alpha$.

Quantity	$c_1 = u_1$	c_2	$c_4 = N_1/Eh$	particular solution
u	1	$\tan[\mathrm{ber} + (1+\nu)\mathrm{ber}_2]/Eh$	$s_1\ln(s/s_1)$	$-gs^2/Eh/4\cdot\varphi$
w	$-\tan\alpha$	$-\tan\alpha\ (2\mathrm{ber} - y\mathrm{ber})/Eh$	$-r_1\ln(s/s_1)/\cos\alpha$	$+gs^2\cdot\sin\alpha/Eh/2\cdot f$
ψ	0	$\mathrm{ber}_2/B/\eta$	$s_1/s\cdot\tan\alpha$	$-gs\sin\alpha/Eh\cdot f$
N	0	$-\tan\alpha\ \mathrm{ber}_2/s$	$s_1/s\cdot Eh$	$-gs/2/\cos\alpha$
Q	0	ber_2/s	0	0
M	0	$2(y\mathrm{ber}_2' + \nu\mathrm{ber}_2)/y$	$(1-\nu)B\tan\alpha\ s_1/s^2$	$gh^2\sin\alpha/12/(1-\nu)\cdot f$

Equations 4 and 6 were tested using the transfer matrix and stiffness methods.

Comparison results were calculated using a finite difference shooting method developed for computing thin axisymmetric orthotropic and composite structures. This method can be described as a transfer matrix method, where a new matrix is determined at every qrid point for the next step. The results are sharpened using several calculation loops and then Richardson's extrapolation toward zero grid length. First the program determines the transfer matrix and the loading vector for the elements. On the base of this local and global stiffness matrices and nodal load vectors are calculated . Deviation from the symmetry in the stiffness matrix is used as an error parameter. Inside quatities of the element are calculated using transfer matrices at every grid point of the first calculation loop.

In the next will be brief descriptions of computing with Equations 4 and 6 using the transfer matrix and stiffness methods.

The transfer matrix method

For the transfer matrix method vectors **a** and **b** are vector $\mathbf{e}(\mathbf{s}) = \{u\ w\ \psi\ N\ Q\ M\}$ calculated with arguments $s = s_1 = 2m$ and $s = s_2 = 5m$. Equation 4 is now in numerical form as

$$\begin{bmatrix} 1{,}470804 & 0{,}271819 & -0{,}186141 & 0{,}153317\cdot10^{-2} & 0{,}323191\cdot10^{-2} & 0{,}0214573 \\ -38{,}539074 & -21{,}247635 & 16{,}493004 & -0{,}0287824 & -0{,}232134 & -1{,}961246 \\ -29{,}151238 & -16{,}830475 & -15{,}150729 & 0{,}529531\cdot10^{-2} & 1{,}270381 & 3{,}611878 \\ -1587{,}2510 & -416{,}39978 & 313{,}15960 & -0{,}369094 & 11{,}323683 & -11{,}660495 \\ 916{,}39978 & 529{,}08366 & -180{,}80278 & 0{,}0444037 & -6{,}137731 & 6{,}732190 \\ 495{,}00695 & 285{,}79240 & -27{,}873685 & 0{,}176765 & -6{,}908573 & -9{,}157266 \end{bmatrix} \cdot \begin{Bmatrix} u_1 = 0 \\ w_1 = 0 \\ \psi_1 = 0 \\ N_1 \\ Q_1 \\ M_1 \end{Bmatrix} + \begin{Bmatrix} -0{,}00292771 \\ 1{,}13130731 \\ -0{,}05565189 \\ 47{,}64261323 \\ -18{,}41320884 \\ -7{,}92591531 \end{Bmatrix} = \begin{Bmatrix} u_2 \\ w_2 \\ \psi_2 \\ N_2 \\ Q_2 \\ M_2 \end{Bmatrix}$$

Three last lines of Equation 4 are now used for solving unknowns N_1, Q_1 and M_1 at the starting boundary, where the vector $\{u\ w\ \psi\ N\ Q\ M\} = \{0\ \ 0\ \ 0\ \ 32{,}730762\ 3{,}836052\ 1{,}396604\}$ is then obtained. The deviation of the matrix

$$\mathbf{E}(s_1)\mathbf{A}^{-1} = \begin{bmatrix} 1{,}000 & -0{,}125\cdot10^{-17} & -0{,}143\cdot10^{-17} & 0{,}426\cdot10^{-20} & 0{,}146\cdot10^{-19} & -0{,}295\cdot10^{-19} \\ -0{,}2112\cdot10^{-15} & 1{,}000000 & -0{,}231\cdot10^{-16} & -0{,}450\cdot10^{-19} & 0{,}171\cdot10^{-17} & 0{,}151\cdot10^{-17} \\ -0{,}116\cdot10^{-14} & 0{,}114\cdot10^{-14} & 1{,}000000 & 0{,}157\cdot10^{-17} & 0{,}868\cdot10^{-17} & -0{,}392\cdot10^{-16} \\ -0{,}666\cdot10^{-13} & -0{,}500\cdot10^{-13} & 0{,}726\cdot10^{-14} & 1{,}000000 & -0{,}475\cdot10^{-15} & 0{,}138\cdot10^{-14} \\ -0{,}109\cdot10^{-13} & 0{,}889\cdot10^{-13} & -0{,}596\cdot10^{-15} & 0{,}365\cdot10^{-17} & 1{,}000000 & 0{,}674\cdot10^{-15} \\ -0{,}350\cdot10^{-13} & -0{,}265\cdot10^{-14} & 0{,}538\cdot10^{-14} & -0{,}270\cdot10^{-16} & -0{,}501\cdot10^{-16} & 1{,}000000 \end{bmatrix}$$

from the unit matrix will give some idea of the size of round-off error. Results at boundaries and five inner points are presented at Table 2.

Table 2. Some results for the conical shell considered above. Loading is own weight g = 3,75 kN/m^2. Data are in Figure 2. In Table meridional forces are given.

Distance from [m] the apex	u [m]	w [m]	f rad	N [kN/m]	Q [kN/m]	M [kNm/m]
2,0	$-0{,}474\cdot10^{-18}$	$-0{,}173\cdot10^{-16}$	$0{,}170\cdot10^{-15}$	32,730763	3,836052	-1,396605
2,5	0,0095190	-0,0551153	0,104711	26,164395	1,131956	-0,137583
3,0	0,0176448	-0,0869845	0,0963362	19,753732	0,142183	0,106319
3,5	0,0242797	-0,129888	0,0761892	13,779034	-0,068312	0,083771
4,0	0,0293851	-0,164662	0,0644069	8,541408	-0,059992	0,034945
4,5	0,0330843	-0,195466	0,0596134	3,995105	-0,021230	0,008200
5,0	0,0355400	-0,224769	0,0578527	0,000000	0,000000	0,000000

Stiffness method

The numerical form of Equation 4 is now:

$$\begin{bmatrix} -899{,}784 & -67{,}8470 & 14{,}4733 & 795{,}256 & 7{,}49781 & -0{,}65311 \\ 67{,}8470 & 62{,}5518 & -17{,}2417 & 45{,}0155 & 2{,}60934 & 1{,}14944 \\ -14{,}4733 & -17{,}2417 & 10{,}2378 & -16{,}9989 & -0{,}928756 & -0{,}12646 \\ -318{,}103 & 18{,}0061 & -6{,}79954 & 414{,}432 & 37{,}6097 & 14{,}7316 \\ 2{,}89914 & -1{,}04372 & 0{,}371497 & -37{,}6097 & -18{,}9387 & -8{,}19635 \\ -0{,}261233 & -0{,}459768 & 0{,}050585 & -14{,}7316 & -8{,}19635 & -6{,}89841 \end{bmatrix} \begin{bmatrix} u_1 = 0 \\ w_1 = 0 \\ \psi_1 = 0 \\ u_2 \\ w_2 \\ \psi_2 \end{bmatrix} + \begin{Bmatrix} -6{,}190393 \\ -2{,}756201 \\ 0{,}993903 \\ 7{,}127699 \\ 2{,}445997 \\ 0{,}919631 \end{Bmatrix} = \begin{Bmatrix} N_1 \\ Q_1 \\ M_1 \\ N_2 = 0 \\ Q_2 = 0 \\ M_2 = 0 \end{Bmatrix}$$

Three last lines are used to determine the unknown displacements and we will obtain the vector

$\mathbf{a} = \{u_1\ w_1\ \psi_1\ u_2 w_2\ \psi_2\} = \{0\ \ 0\ \ 0\ \ 0{,}035540\ -0{,}224768\ 0{,}057852\}$

In matrix $E(s_1)A^{-1}$ three first tines are now an approximation of the unit matrix and three last lines are similar with the first three lines of the element stiffness matrix (deviations due to round-off error). The final results coincide with the ones given in Table 2 with minimal differences in values of displacements at the inner boundary.

Comparison results were computed using the numerical method described briefly above. Calculations were made with one element in five loops with 12 to 72 meridional segments. Error parameter pointed an error in the nineth digit. Also the results of this program coincided with the ones of Table 2.

Examples of using shape functions

How to use Equation 4, when we have presumed shape functions instead of accurate solution functions of the homogenous differential equation?

We are considering first a two-noded bar element and assume a linearly varying displacement.

$$u(s) = \frac{s}{l} u_2 + \left(1 - \frac{s}{l}\right) u_1, \ (0 \le s \le l) \tag{7}$$

where l is the length of the element.

The strain of the bar is

$$\varepsilon(s) = \frac{du}{ds} = \frac{u_2 - u_1}{l} \tag{8}$$

Further the normal forces at note 1 and 2 are

$$N_i = EA_i \frac{(u_2 - u_1)}{l}, \ (i = 1,2) \tag{9}$$

where EA is the streching stiffness. In fact we know the stiffness matrix but for demonstrating the validity of Equation 4 it has to be noticed that in Equations 7 and 3a u_1 and u_2 have the role of integration constants. When calculating the stiffness matrix, matrix **A** is the unit matrix and vector **b** = $\{N_1 N_2\}$ and matrix **B** from Equation 9 will be

$$\mathbf{B} = \begin{bmatrix} -EA_1/l & EA_1/l \\ -EA_2/l & EA_2/l \end{bmatrix}, \tag{10}$$

which is the final result ($\mathbf{BA}^{-1} = \mathbf{B}$).

When determining the transfer matrix we have vector **a** = $\{u_1 N_1\}$ and **b** = $\{u_2 N_2\}$.
Matrices **A** and **B** are

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ -EA_1/l & EA_1/l \end{bmatrix} \tag{11}$$

and

$$\mathbf{B} = \begin{bmatrix} 0 & 1 \\ -EA_2/l & EA_2/l \end{bmatrix}. \tag{12}$$

The transfer matrix will be

$$\mathbf{T} = \mathbf{BA}^{-1} = \begin{bmatrix} 1 & \frac{l}{EA} \\ 0 & \frac{EA_2}{EA_1} \end{bmatrix}. \tag{13}$$

Results 10 and 13 can be achieved also by assuming the displacement $u(s) = c_1 + c_2 s (0 \le s \le l)$
Obviously results 10 and 13 are correct for a bar with a constant cross section.

A better choice as a shape function will be the constant value N_1 for the normal force. The strain will be then

$$\varepsilon(s) = \frac{N_1}{EA(s)}, \tag{14}$$

which will produce the expression

$$u(s) = u_1 + N_1 \cdot \int_0^s \frac{ds}{EA(s)} \tag{15}$$

where u_1 and N_1 are replacing integration constants $\mathbf{c} = \{c_1\ c_2\}$. Displacement $u_2 = u(l)$ is now

$$u_2 = u_1 + N_1 \cdot \int_0^l \frac{ds}{EA(s)} \equiv u_1 + N_1 S \ . \tag{16}$$

When calculating now the stiffness matrix, we will have $\mathbf{a} = \{u_1 u_2\}$, $\mathbf{b} = \{N_1 N_2\}$ and the operation

$$\mathbf{BA}^{-1} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 1 & S \end{bmatrix}^{-1} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ -1/S & 1/S \end{bmatrix} = \begin{bmatrix} -1/S & 1/S \\ -1/S & 1/S \end{bmatrix} \tag{17}$$

A more general result will be obtained by presenting vector $\{\mathbf{f}_1\ \mathbf{f}_2\}$ with Equation

$$\begin{Bmatrix} f_1 \\ f_2 \end{Bmatrix} = K \begin{Bmatrix} d_1 \\ d_2 \end{Bmatrix} + \begin{Bmatrix} r_1 \\ r_2 \end{Bmatrix} = \begin{bmatrix} C & D \\ F & G \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}^{-1} \begin{Bmatrix} d_1 \\ d_2 \end{Bmatrix} + \begin{Bmatrix} r_1 \\ r_2 \end{Bmatrix} \tag{18}$$

where $\mathbf{f}_i$, $\mathbf{d}_i$ and $\mathbf{r}_i$ are force, displacement and nodal load vectors at node i. Matrices **C, D, E** and **G** are submatrices of the stiffness matrix **K**. As before in Equations 10...13 we can shift over from the stiffness method to the transfer matrix method. For this purpose vectors **a** and **b** are

$$\mathbf{a} = \begin{bmatrix} I & 0 \\ C & D \end{bmatrix} \begin{Bmatrix} d_1 \\ d_2 \end{Bmatrix} + \begin{Bmatrix} 0 \\ r_1 \end{Bmatrix} \tag{19}$$

and

$$\mathbf{b} = \begin{bmatrix} 0 & I \\ F & G \end{bmatrix} \begin{Bmatrix} d_1 \\ d_2 \end{Bmatrix} + \begin{Bmatrix} 0 \\ r_2 \end{Bmatrix} \tag{20}$$

For the transfer matrix we will obtain

$$\mathbf{T} = \mathbf{B}\mathbf{A}^{-1} = \begin{bmatrix} 0 & I \\ F & G \end{bmatrix} \cdot \begin{bmatrix} I & 0 \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} 0 & I \\ F & G \end{bmatrix} \cdot \begin{bmatrix} I & 0 \\ -D^{-1}C & D^{-1} \end{bmatrix} = \begin{bmatrix} D^{-1}C & D^{-1} \\ F - GD^{-1}C & GD^{-1} \end{bmatrix} \tag{21}$$

The loading vector has the expression

$$\begin{Bmatrix} q_1 \\ q_2 \end{Bmatrix} = \begin{Bmatrix} 0 \\ r_2 \end{Bmatrix} - BA^{-1} \begin{Bmatrix} 0 \\ r_1 \end{Bmatrix} = \begin{Bmatrix} D^{-1}r_2 \\ r_2 - GD^{-1}r_2 \end{Bmatrix}, \tag{22}$$

where $\mathbf{q}_i$ are subvectors of the loading vector. Equation 21 is given inter alia in reference [3, p. 105].Results 21 and 22 are also valid when going back from the transfer matrix method to the stiffness one, when in Equation 22 $\mathbf{q}_i$ and $\mathbf{r}_i$ change places with each others.

Then the question of using Equation 4 for generating a more-noded element for the stiffness method. Matrix **A** is regular and $\mathbf{A}^{-1}$ exists if shape functions of displacements fulfill geometrical boundary conditions. Also matrix **B** can be determined and forces at boundaries and inside the element can be calculated. An open question is the additional fiction of the finite element method. The forces have to be concentrated at the nodal points. Reference 4 highlights this problem.

ACKNOWLEDGEMENTS

The author wishes to express his gratitude to professor Eero-Matti Salonen for valuable discussions and cooperation.

REFERENCES

1. Tuominen, P. Static equation of an analytic element (A unified theory of the stiffness, mixed and flexibility methods). Proceedings of the NSCM-11: Nordic Seminar on Computational Mechanics, Anders Ericsson and Costin Pacoste (editors) Royal Institute of Technology, Department of Structural Engineering, pp. 171-174. Stockholm 1998.

2. Outinen, H. Isotrooppisen kartiokuoren rotaatiosymmetrinen statiikka, Tampere University of Technology, Department of Mechanical Engineering, Applied Mechanics. Tampere 1983.

3. Tesár, A. & Fillo, L'u: Transfer Matrix Method. Kluwer Academic Publishers, Dordrecht/Boston/London 1988.

4. Salonen, E-M.: The concept of fictitious loading. International Journal for Numerical Methods in Engineering, Vol. 14, 1653-1663 (1979).