

301.162

# ACTA TECHNICA

22.

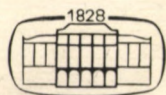
ACADEMIAE SCIENTIARUM HUNGARICAE

107  
1995/1996

EDITOR-IN-CHIEF: P. MICHELBERGER

VOLUME 107  
NUMBERS 1-2

CIVIL ENGINEERING — C/8



AKADÉMIAI KIADÓ, BUDAPEST 1995-96

ACTA TECHN. HUNG. ATSHA8

# ACTA TECHNICA

A JOURNAL OF THE HUNGARIAN ACADEMY OF SCIENCES

---

## CENTRAL EDITORIAL BOARD

T. CZIBERE, K. GÉHER, L. KOLLÁR, P. MICHELBERGER (EDITOR-IN-CHIEF),  
A. LÉVAI, J. PROHÁSZKA, K. REMÉNYI, J. SZABÓ,  
GY. CZEGLÉDI (MANAGING EDITOR)

## EDITORIAL COMMITTEE FOR MECHANICAL ENGINEERING (SERIES C)

A. BÉNYEI, J. FARKAS, ZS. GÁSPÁR, L. KOLLÁR (CHAIRMAN),  
L. SOMLYÓDY

---

*Acta Technica* publishes original papers, preliminary reports and reviews in English, which contribute to the advancement of engineering sciences.

*Acta Technica* is published by

AKADÉMIAI KIADÓ  
H-1117 Budapest, Prielle K. u. 19—35

### *Subscription information*

Orders should be addressed to

AKADÉMIAI KIADÓ  
H-1519 Budapest, P.O. Box 245

Subscription price for Volume 107 (1995–96) in 4 issues US\$ 98.00, including normal postage, airmail delivery US\$ 20.00.

---

*Acta Technica* is abstracted/indexed in Applied Mechanics Reviews, Current Contents–Engineering, Technology and Applied Sciences, GeoRef Information System, Science Abstracts.

---



## CONTENTS

<u>Bagi, K.</u> : Geometrical modelling of granular assemblies .....	1
<u>Bojtár, I.</u> : Numerical estimation of the elastic properties of cemented granular materials with microstructure .....	17
<u>Galaskó, Gy. - Gáspár, Zs. - Nouri-Baranger, T. - Leon, J.C. - Trompette, P. - Veron, P.</u> : Comparison of tent structures calculation in Hungary and France .....	27
<u>Gáspár, Zs. - Károlyi, Gy. - Popper, Gy.</u> : Application of Gröbner basis theory to find global equilibrium paths of the simple arch .....	37
<u>Györgyi, J.</u> : Use of finite element method for testing structures for dynamical effects caused by moving vehicles .....	49
<u>Kurutz, M.</u> : Imperfection sensitivity of the classical stable-symmetric bifurcation problem modified by bilinear material .....	63
<u>Lovas, A.</u> : Mechanical models of the human skull-brain system .....	83
<u>Nédli, P.</u> : Elastic-plastic analysis of frames in case of geometrical nonlinearity .....	99
<u>Popper, Gy. - Károlyi, Gy.</u> : Gröbner basis calculations with applications to mechanics ....	109
<u>Szabó, J. - Tarnai, T.</u> : On the quasi-inverse of the matrix of a system of linear equations ....	115
<u>Tarnai, T. - Gáspár, Zs.</u> : Packing of equal circles in a square .....	123
<u>Vásárhelyi, A. - Lógó, J.</u> : Dynamic elasto-plastic analysis of structures by mathematical programming .....	137







This volume is offered to the honour of

**SÁNDOR KALISZKY**

on the occasion of his 70<sup>th</sup> birthday

Sándor Kaliszky, the well-known professor of Mechanics at the Technical University of Budapest, member of the Hungarian Academy of Sciences becomes 70 in the middle of June 1997. The same time, after more than 45 years of teaching at the Department of Civil Engineering Mechanics, he will be retired. The authors of this volume, namely, his former students, colleagues, co-authors or friends thought that the review of their newest scientific results would be the best way to express their thanks, respect and great honour to their master, quasi to illustrate and prove his role in painting of the panorama of the Structural Mechanics in the recent 40 years.

## Curriculum vitae of Sándor Kaliszky

**Born:** Diósgyőr, Hungary, 15 June 1927. **Academic degrees:** Diploma in Civil Engineering, Technical University of Budapest (1950), Candidate of Technical Science, Hungarian Academy of Sciences (1961), Technical Doctor, Technical University of Budapest (1962), Doctor of Technical Science, Hungarian Academy of Sciences (1967), Corresponding member of the Hungarian Academy of Sciences (1990), Member of the Hungarian Academy of Sciences (1995). **Academic career:** assistant, Department of Civil Engineering Mechanics, Technical University of Budapest (1950-53), research fellow, Institute of Building Sciences, Budapest (1953-55), assistant professor (1955-63), associate professor (1963-69), full professor (1969-), head of department (1971-93), Department of Civil Engineering Mechanics, Technical University of Budapest, deputy dean (1967-73), Faculty of Civil Engineering, Technical University of Budapest, head of Research Group for Technical Mechanics of the Hungarian Academy of Sciences (1993-), resident rector of the International Centre for Mechanical Sciences, CISM, Udine, Italy (1986-), visiting research fellow, Civil Engineering Department, University of Southampton, Great Britain (1964-65), visiting professor, Civil Engineering Department, Monash University, Melbourne, Australia (1973), visiting professor, Civil Engineering Department, University of Wisconsin-Milwaukee, Milwaukee, USA (1980-81, 1984-85). **University Teaching experience:** Technical University of Budapest, Department of Civil Engineering Mechanics, undergraduate and graduate courses: Mechanics (Statics, Strength of Materials, Dynamics), Elasticity, postgraduate courses: Plasticity, Limit Analysis, Optimal Plastic Design; Monash University, Melbourne, Australia: Limit Analysis of RC Slabs; University of Wisconsin-Milwaukee, USA: Strength of Materials, Plasticity, Continuum Mechanics. **General Areas of Research:** Plasticity, Plastic design of RC slabs, Limit and shakedown analysis, Optimal plastic design, Dynamic analysis of plastic structures, Nonlinear analysis of elastic structures, Soil-structure interaction, Analysis of large panel buildings. **Principal investigator of researches supported by the National Science and Research Foundation:** Stress and strain analysis of the contact between the femur and the prosthesis (1986-1991), Dynamic and stability analysis of structures (1991-1994), Mechanical analysis of skull (1990-1995), Analysis of some nonlinear problems of mechanics (1995-). **Memberships in Scientific Committees and Institutions:** Association for Applied Mathematics and Mechanics (GAMM) (1975-), Governing Council of GAMM (1975-84), Hungarian National Committee of the GAMM (1970-), chairman (1979-90), Committee of European Mechanics Colloquia (EUROMECH) (1973-78, 1984-86), European Mechanics Society (1994-), Advisory Board of the European Mechanics Society (1994-), International Union of Theoretical and Applied Mechanics (IUTAM), General Assembly (1987-), Congress Committee (1987-), Executive Committee of Congress Committee (1988-), Hungarian National Committee of the IUTAM (1969-), secretary (1974-90), chairman (1991-), International Centre for Mechanical Sciences (CISM, Udine), resident rector (1986-), International Association for Bridge and Structural Engineering (IABSE) (1970-), Hungarian National Committee of the IABSE (1970-), Polish Society of Theoretical and Applied Mechanics, foreign member (1983-), Society of Building Science (Hungary) (1960-), Committee for Theoretical and Applied Mechanics of the Hungarian Academy of Sciences (1968-), Council of Natural Sciences of the National Scientific and Research Foundation (NSRF) (1993-), Board of Architecture, Civil and Traffic Engineering of the NSRF, chairman (1993-), General Council of International Association for Computational Mechanics (1994-), Hungarian National Committee of the International Council of Scientific Unions (ICSU), (1991-). **Memberships in Editorial boards of Scientific**



**Journals:** Journal of Mechanics of Structures and Machines (1972-), International Journal of Plasticity (1986-89), Structural optimization (1988-), Advances in Mechanics (1978-93), Periodica Polytechnica (1983-88), Construction and Building Science (in Hungarian) (1981-), Archives of Computational Methods in Engineering (1994-), Applied Mechanics Reviews, reviewer (1968-86), associate editor (1987-94), Computer Assisted Mechanics and Engineering Sciences (1995-). **Awards:** Academic Prize of the Hungarian Academy of Sciences (1966), Copernicus Award of the Polish Academy of Sciences (1984), Memorial Medal of the Technical University of Budapest (1993), Golden Medal of the Cracow University of Technology (1995), Medal of the University of Liège (1995), Albert Szent-Györgyi Award of the Hungarian Ministry of Education (1997). **Publications:** 6 books (one with co-authors, one in English, two in German), 4 chapters in engineering handbooks, 12 lecture notes in Mechanics, 2 dissertations, 85 papers in Hungarian and in international journals, 95 lectures at Hungarian and international conferences, congresses.





## GEOMETRICAL MODELLING OF GRANULAR ASSEMBLIES

BAGI, K.\*

(Received 20 October 1996)

The aim of this paper is to establish the geometrical background of the correct definition of continuum-mechanical state variables for granular assemblies. Two complementary cell systems are introduced here; they provide the suitable equivalent continua to replace the granular assembly in continuum-mechanical analysis.

### 1. Introduction

This paper concentrates on the geometrical analysis of *granular assemblies*. Its aim is to provide the background for the application of continuum-mechanical state variables in case of this non-continuous, 'discrete' material.

The term 'granular assembly' is used here for a material consisting of randomly packed grains not fixed to each other (in contrast to the other type of granular materials referred to as 'cemented'). The grains in it are able to slip or roll along each other, contacts can be deleted or new contacts created etc., so even the whole internal structure can be rearranged under the external mechanical loading effects. This phenomena makes the behaviour of granular assemblies so interesting both from theoretical and from practical engineering point of view.

The final goal of the mechanics of granular materials is to provide relationships between the external loads acting on the material and the resulting displacements. Traditionally, the effect of external loads is expressed by the continuum-mechanical state variable *stress* (relation between loads and the stress field is given by the equilibrium equations of continuum mechanics, for example the Cauchy-equations in the simplest case); deformations are reflected by the other continuum-mechanical state variable *strain* (geometrical equations set the link between displacements and the strain field). Stress and strain are related to each other through the *constitutive equations* (which are expected to contain all the necessary information about the mechanical characteristics of the material). The geometrical and equilibrium equations are clear in continuum-mechanics; but to find the proper constitutive equations for granular assemblies is not as simple at all: for many years, a large number of theoretical and experimental studies have been concerned with the problem, and the results seem to be rather limited.

---

\* Bagi, Katalin; Technical University of Budapest, Dept. of Civil Engineering Mechanics, Műegyetem rkp. 3, K. mf. 35. H-1521 Budapest

Recently there are two approaches that most of the researchers follow in order to solve this problem. Let us call the first one the *continuum-mechanical*, and the second one the *microstructural* approach.

The idea of continuum-mechanical approach is to consider the assembly as a continuous domain, accept the concept of an infinitesimally small representative volume element, and apply stress and strain as the fundamental variables that uniquely determine the state of the material in any point. Constitutive relations are searched for in such a way that they would not violate the fundamental laws of physics; and the parameters in the equations, expressing the specific properties of the material, are measured experimentally. This empirical method is the one used today for practical engineering problems.

The problem with this is the limited validity of the results. Experimentally determined relations easily become unreliable if the circumstances for which we want to apply them differ even slightly from those that existed during the experiments. (To improve the constitutive equations, either the mathematical form of the equations must be made more complicated by increasing the number of parameters; or additional state variables are introduced beside the traditional stress and strain.)

The microstructural approach is a relatively new method, and - in the long run - it may be an advantageous alternative to the previous one. The aim of the microstructural approach is to find macro-level state variables that are based on micro-variables such as contact forces, grain displacements, and local geometrical characteristics. Since it would reflect those characteristics of the material that are most significant in determining the macro-behaviour, and the relationships between its state variables would be strongly connected to the phenomena taking place in the microstructure, a microstructural theory is expected to be far more reliable and general than the existing continuum-mechanical models.

(It has to be mentioned that there exists a 'microstructural' approach within continuum-mechanics too (see [10] for instance): the representative volume in it has a finite size and can have deformations independently from the macro-level distortions. However, it was found by Füzy et al. [11] that the phenomena remain of boundary nature.)

Researchers generally accept today that - from practical reasons - even in the microstructural approach the continuum-mechanical state variables should be re-interpreted for the discrete material in a suitable and theoretically correct way; and they should be the state variables of the granular assemblies too. (This concept is, in fact, not self-evident at all. There are several doubts about the applicability of stress and strain that were originally defined for a domain without any internal structure, so they might be strongly insufficient for the description of the state-changes of this material whose most important characteristic is just its internal structure. Pioneering efforts can already be found in the literature to apply new methods as graph theory, fractal geometry, percolation theory etc. instead of continuum-mechanics.)

Our own researches [1], [2] in the last few years focused on the microstructural analysis of granular assemblies. The survey of international literature showed that even the fundamental questions of this area were still unsolved: the microstructural definition of stress and strain, and the interpretation of their meaning from the point of view of micro-level variables were still missing. In other words: the transition between the discrete internal structure (micro-level analysis) and the continuum (macro-level analysis) was not established. Though there were good ideas for the definition of stress tensor [3], [4], their theoretical background was insufficient; and the problem of strain tensor was completely unsolved - mainly because of the lack of proper geometrical fundamentals.



The solution of the above problems was strongly inspired by the works of E. Tonti, an Italian mathematician who was searching for the reason of the common experience that physical theories having very different meaning show close analogies in the mathematical built-up of their basic equations [5]. He pointed out that the differential operators used in the mathematical equations correspond to a so-called *coboundary process* executed on two complementary cell systems and this fact leads to conclusions that give the explanation for the existence of the analogies. The two systems can be introduced as follows. To define the primal cell system, consider the analysed continuous region  $\Omega$  of the n-dimensional Euclidean space (for simplicity, this summary will be restricted to  $n=3$  only, though the considerations are valid for smaller or larger  $n$  too). Subdivide  $\Omega$  into small three-dimensional cells whose faces are formed by the co-ordinate surfaces of a co-ordinate system  $x^1, x^2, x^3$ ; these will be called as 3-cells. Every 3-cell is composed of faces, edges and nodes that will be considered as 2-cells, 1-cells and 0-cells. To construct the dual cell system, consider the centres of the 3-cells; they become the nodes (0-cells) in the dual cell system that is built upon the dual nodes in the same way as seen above. Obviously, for every p-cell of the primal system there corresponds a (n-p)-cell of the dual system and vice versa (see Figure 1.). In general, in a physical theory if the geometrical and kinematical variables are referred to the p-cells of the primary system, the corresponding statical and dynamical variables are referred to the p-cells of the dual system.

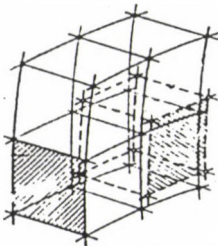
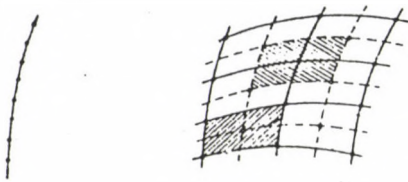
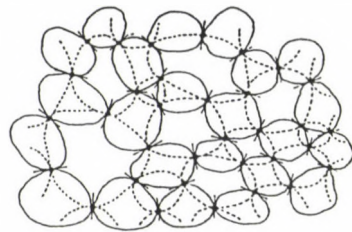
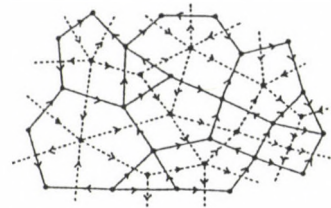


Fig. 1.



(a) A Granular Assembly  
(Dotted line indicates dual particles)



(b) Particle- and Void-Graphs

Fig. 2.

This concept is obviously well-suited for continuum mechanics. Here the cell size can be infinitesimally small, and the two cell systems - in a limit sense - slip onto each other. The granular material, however, does not make this transition possible: first, the grains have a finite size so an infinitesimally small cell size cannot be applied; second, a cell system defined according to the above way (faces and edges parallel to the co-ordinate surfaces and axes) is

completely meaningless in case of the random internal structure of granular assemblies. So the concept of Tonti cannot directly be applied for the granular case.

However, the idea of applying two complementary systems already arose in the micromechanics of granular assemblies, though in a very different way. In the graph-theoretical approach of Satake [6], [7] two complementary graphs (particle- and void-graph) are applied for the topological characterization of 2D random assemblies. As shown in Figure 2., the nodes in the particle graph correspond to loops in the void graph (representing the grains); branches of the two graphs correspond to each other (they represent the contacts); loops in the particle-graph correspond to nodes in the void-graph (voids). The topological structure of the graphs is expressed by two *topological matrices*. The equilibrium and compatibility equations of the system are compiled with the help of these matrices; and it was found by Satake that the topological matrices here have the same role as the Schaefer-operators in generalized continuum-mechanics [8]. (The correspondence between the topological matrices and Schaefer-operators can be understood with the help of Tonti's results. Tonti's concept includes that the differential operators of continuum-mechanics show an analogy with the coboundary processes. Indeed, the topological matrices of Satake could be considered as operators prescribing some kind of 'coboundary processes' in random granular assemblies.)

These preliminaries inspired our theoretical work whose geometrical background will be introduced in this paper.

The aim was to establish the transition between the discrete and continuous analysis: to define suitable *equivalent continua* for the replacement of the assemblies. Before introducing them in Chapters 2. and 3., let us shortly summarize those existing methods that can often be found in the literature for the representation of the internal structure.

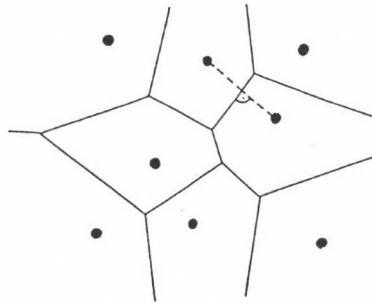


Fig. 3.

The widely-used *Voronoi-tessellation* can be applied in several ways. In the simplest case it is defined for a set of discrete points given in the 2D or 3D Euclidean space. In the 2D problem (see Figure 3.) the plane is subdivided into polygonal domains each of them containing exactly one point. The edges of the domains are the bisecting lines of those straight segments that connect the neighbouring points. (In 3D the system is similar: faces of polyhedral domains are given by the bisecting planes between neighbouring points.)

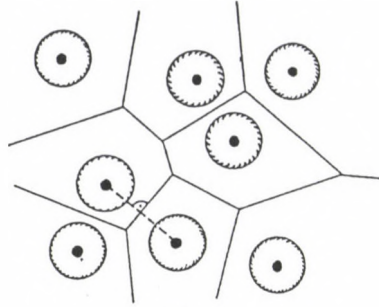


Fig. 4.

Its generalized version can be applied for monosize assemblies of circular or spherical grains. For 2D assemblies of equal circles (Figure 4.) the plane can be divided into polygons whose edges are the bisecting lines of straight segments joining the centres of neighbouring grains. (The same can be done for 3D too.) A cell system ('Voronoi-cells') results in such a way that there is exactly one grain in each cell. This system is especially suitable for the analysis of regular assemblies.

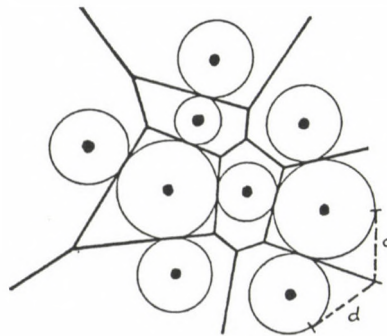


Fig. 5.

*Dirichlet-tessellation*, the next system we shall introduce here, has the main advantage to the Voronoi-tessellation that the grains do not necessarily have the same size. Consider a set of non-intersecting circular grains in 2D, or non-intersecting spherical grains in 3D. A domain can be assigned to each grain, consisting of those points which have a shorter or equal tangent to that grain than to any other grain (Figure 5.). The common faces of the domains are the power lines (power planes in 3D) of neighbouring grains. (It may be worthwhile to mention that in case of monosize assemblies the Dirichlet- and Voronoi-tesselations are equivalent.) Similar tessellation was suggested by Gellatly and Finney [9] for the characterization of assemblies having circular grains with different sizes.



The Dirichlet-tessellation, in principle, could be generalized for particles with arbitrary smooth convex shape, but as far as we know the problem was solved and construction algorithms were found only for grains and assemblies having very special regular geometry.

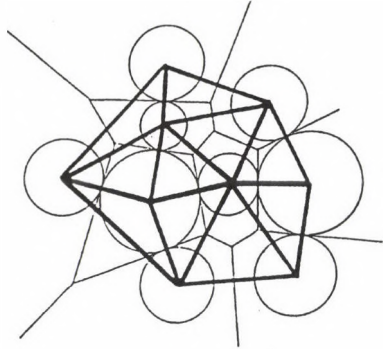


Fig. 6.

The *Delaunay-network* can also be a useful tool in characterizing granular systems. Consider an assembly of circular or spherical grains; if the Dirichlet-cells of two grains have a common side, connect the two grain centres by a straight line. These connecting lines form the Delaunay-network of the assembly (Figure 6.).

The definition can be modified to give a more physical meaning to the network if the centres of grains being in contact are connected (Figure 7.). In this version the branches in the network correspond to the internal supports in the microstructure. However, the duality with the Dirichlet-tessellation does not necessarily holds in this case.

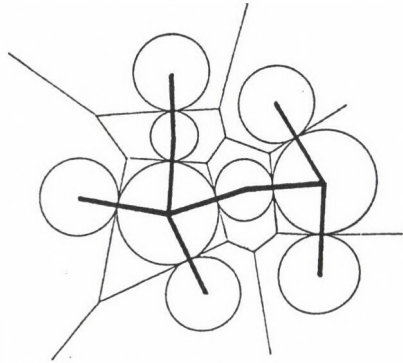


Fig. 7.

In Chapter 2. and 3. of the present paper two complementary geometrical systems will be introduced, as an alternative to the previous descriptions. They will be defined for assemblies of grains having arbitrary convex shape; and there will be a clear duality between them. Our opinion is that these advantages make the suggested cell systems more powerful for the modelling of granular assemblies than the presently applied Voronoi-, Dirichlet- and



Delaunay-systems. The two complementary systems, namely the material cell system and the space cell system, served as equivalent continua for granular assemblies in our previous works [1], [2] in which continuum-mechanical derivations led to the physically and theoretically based microstructural definitions of stress and strain tensors. The derivations will not be repeated in the present paper since its aim is to give the geometrical background to the mechanical analysis; but the results will shortly be summarised in Chapter 4.

## 2. The material cell system

### 2.1. The material cells

Consider an assembly consisting of grains with convex but otherwise arbitrary shape in the 2D or 3D Euclidean space where the distance between two points is understood in the usual sense; and the  $\overline{PG}$  distance between a point P and a grain G is the following:

1. If P is outside G or on its boundary,  $\overline{PG}$  is the distance between P and that point of G which has the smallest distance from P. (This includes that for a Q boundary point of G the distance is zero:  $\overline{QG} = 0$ .)
2. By definition, if P is inside G then  $\overline{PG}$  is negative and its absolute value is the smallest distance between P and the boundary points of G.

Consider now a grain  $G_0$  and collect all those P points whose distance from  $G_0$  is less or equal than from any other grain:

$$\overline{PG_0} \leq \overline{PG_k} \quad (k \neq 0)$$

These P points form a domain around the  $G_0$  grain. The domain has the following important characteristics:

- The internal and boundary points of  $G_0$  all belong to it.
- If  $\overline{PG_0} < \overline{PG_k}$  for all  $k \neq 0$ , then P is an internal point of the domain; if there exists a  $G_i$  for which  $\overline{PG_0} = \overline{PG_i} \leq \overline{PG_k}$  for all  $k \neq 0$  and  $k \neq i$ , then P is a boundary point of the domains of  $G_0$  and  $G_i$ .

Constructing these domains for all the grains the space is subdivided as illustrated in 2D in Figure 8. below. Notice the following properties:

- There is exactly one grain in each domain.
- The domains are contiguous.
- Grains on the boundary of the assembly have infinite domains while the domains are finite in the inside of the assembly.
- Common face of neighbouring domains (belonging to the grains  $G_1$  and  $G_2$ ) is the set of P points for which  $\overline{PG_1} = \overline{PG_2} \leq \overline{PG_k}$  for all  $k \neq 1$  and  $k \neq 2$ .
- If two grains have a contact point, the corresponding domains must have a common face that contains the contact point itself.

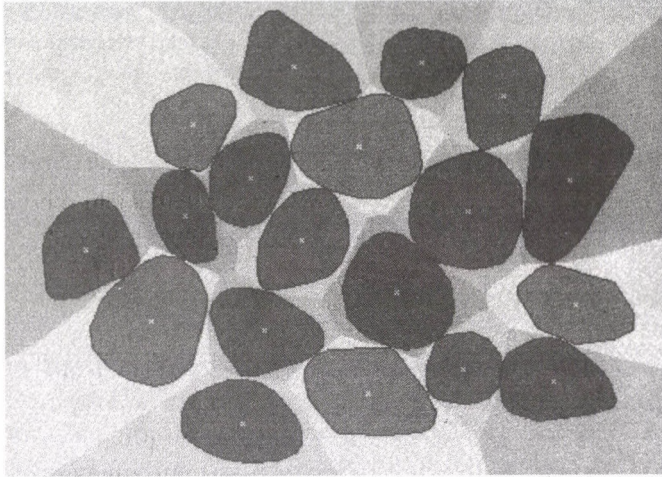


Fig. 8.

These domains will be referred to as *material cells*, and the total system given by them as the *material cell system*. The following terminology will be used in their characterisation:

- a) in 3D: The common points of neighbouring cells form *faces*; faces join each other on edges (note that if the face  $\overline{PG_1} = \overline{PG_2} \leq \overline{PG_k}$  and the face  $\overline{PG_2} = \overline{PG_3} \leq \overline{PG_k}$  have a common edge, then the face  $\overline{PG_1} = \overline{PG_3} \leq \overline{PG_k}$  also joins this edge); edges intersect with each other on *nodes*.
- b) in 2D: The common points of neighbouring cells will also be referred to as *faces* (however, they are one-dimensional lines in this case); they intersect with each other on *nodes*.

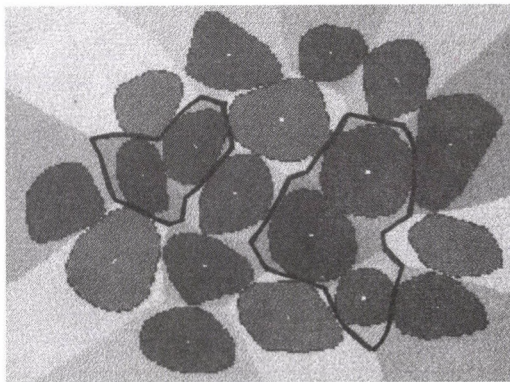


Fig. 9.

Any set of grains having finite material cells will be called as a *finite sub-assembly* (the cells are not required to form a contiguous domain). The boundary of the finite sub-assembly consists of finite faces, forming one or more closed surfaces (curves in 2D). Figure 9. illustrates a possibility in 2D.

## 2.2. Modify the topology: Triangularisation

It may happen in case of special geometry that there are *multiple nodes or edges* in the system: in 3D, more than 4 edges joining a node or more than 3 faces joining an edge; in 2D, more than 3 faces joining a node. In this case a small disturbance has to be added to the system to destroy the speciality of the geometry. The original topology (number of nodes/edges/faces, which node joins which edges, which grains are neighbouring etc.) has to be replaced by the topological data of this new assembly with disturbed geometry, and from now on, the new topology is considered in any further analysis. The multiple edges and nodes are *split up* this way.

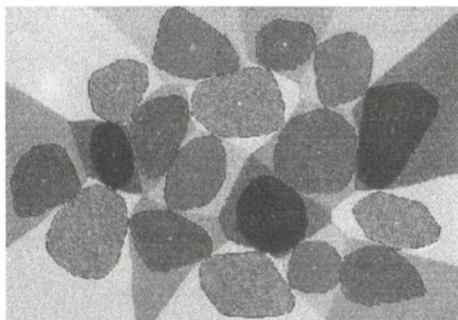


Fig. 10.

Figure 10. illustrates an assembly where there is a node belonging to 4 edges (in the middle), and an other node belonging to 5 edges (left down). Figure 11. shows the effect of small disturbance.

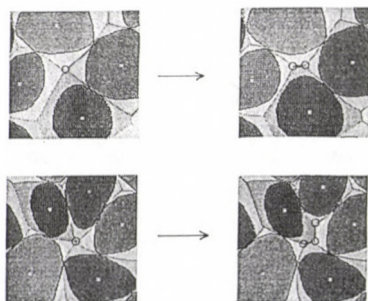


Fig. 11.



### 2.3. Add an imaginary grain in the infinity

This imaginary grain will be the neighbour of all those grains that have an infinite material cell (i.e. the grains on the boundary). So a boundary grain has a common face (in the infinity) with the imaginary grain; edges and nodes belonging to it are also formed (in the infinity) in the usual way.

This could be visualized for a 2D system in the following way: imagine the in-plane assembly as occupying a small portion of the surface of a sphere with infinitely large radius, and the imaginary grain is on the opposite side of the sphere; then the cell system is formed on the surface of the sphere. (Unfortunately, this is much more difficult to imagine for a 3D assembly, but the concept is just the same.)

The geometry and topology of the material cell system is now defined. It has to be noted that *every point of the space - apart from the set of nodes, edges and faces, a zero-measure set - is covered exactly once by the material cell system*. This fact makes the system suitable for the role of equivalent continuum of a granular assembly.

## **3. The space cell system**

### 3.1. Definition

The construction of space cell system is directly based on the above definitions and characteristics. Starting from an assembly and its material cell system, the space cell system is defined by the following algorithm in 3D:

1. *Nodes* of the system are the grain centres (they correspond to the material cells).
2. If two material cells have a common face, the corresponding grain centres are connected with a straight line that will serve as an *edge* in the space cell system.
3. Consider now an edge in the material system. Three faces intersect on this edge. The three faces define three edges in the space cell system in such a way that they form a closed triangle; this triangle will be a *face* in the space cell system.
4. Similarly, consider next a node in the material cell system; and consider the edges joining this node. As shown before, the edges in the material system correspond to faces in the space system; if the material-edges belong to the same node, the space-faces form a closed *cell* in the space system corresponding to the material-node.

The definition is of course shorter in the 2D case:

1. *Nodes* of the space cell system are the grain centres (they correspond to the material cells).
2. Where two material cells have a common face, the corresponding grain centres should be connected by a straight line, similarly to the 3D case; these lines will be the *edges* in the space cell system.
3. A node in the material system is the common point of three joining faces; the three corresponding edges in the space system form a closed *cell*.

Note that in 2D the cells are triangles, and in 3D the cells are tetrahedrons so the space cell system consists of *simplexes* in any case. Figure 12. shows the space cell system of a 2D assembly.

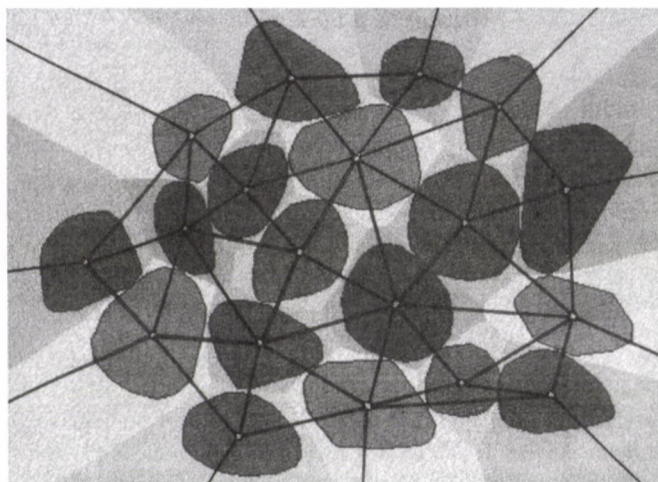


Fig. 12.

### 3.2. Positive and negative space cells

The definition of positive and negative space cells is illustrated in 2D in Figure 13. First, consider the  $N_1$  material-node, and the corresponding space cell that is given by the  $(n_1, n_2, n_3)$  space-nodes:

- going clockwise around material node  $N_1$ : the material cells are found in the order  $C_1 - C_2 - C_3$
- going clockwise in the corresponding space cell the space nodes are found in the order  $n_1 - n_2 - n_3$ . This is the same as the order of the corresponding material cells:  
*this is a positive space cell*

Now consider the  $N_2$  material-node, and the corresponding space cell that is also given by the  $(n_1, n_2, n_3)$  space-nodes:

- going clockwise around material node  $N_2$ : the material cells are found in the order  $C_1 - C_3 - C_2$
- going clockwise in the corresponding space cell the space nodes are found in the order  $n_1 - n_2 - n_3$ . This is the opposite of the order of the corresponding material cells:  
*this is a negative space cell*

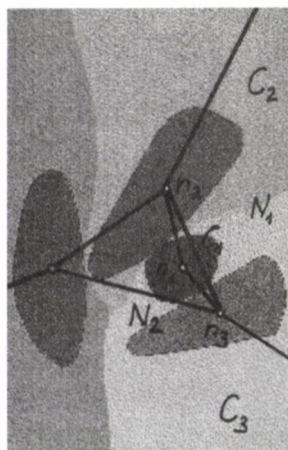


Fig. 13.



The definition is more complicated in the 3D case:

Consider a material node in the material cell system of a 3D assembly, which belongs to the  $(G_1, G_2, G_3, G_4)$  grains. There are four material-edges joining this node:  $E_1$  that belongs to the  $(G_2, G_3, G_4)$  grains;  $E_2$  that belongs to the  $(G_1, G_3, G_4)$  grains;  $E_3$  that belongs to the  $(G_1, G_2, G_4)$  grains; and  $E_4$  that belongs to the  $(G_1, G_2, G_3)$  grains. Arbitrarily select one of the edges:  $E_4$  for example. As the first step, from the direction of the analysed material-node look along the selected edge on an infinitesimally short distance. Three faces join this edge, separating the cells of the  $(G_1, G_2, G_3)$  grains. Now (looking still from the direction of the node) go clockwise around the edge, and remember the order of the material cells. Second, consider the space cell corresponding to the analysed material-node. Find the face corresponding to the previously selected material-edge (in our case this is the space-face determined by the centres of  $(G_1, G_2, G_3)$  grains). From the fourth node, look towards this face, and go clockwise around its nodes. Remember the order of space-nodes (grain centres).

Now compare the two orders. If they are the same, the space cell is *positive* by definition; if they are the opposite, the space cell is said to be *negative*.

Usually most of the space cells in an assembly are positive. But if the grains have a very elongated shape or the difference between the minimum and maximum size of the grains is very large, there may be several negative cells in the structure and they may even be embedded into each other several times. However, the following characteristic of the space cell system can be recognised:

*Every point of the space - apart from the set of nodes, edges and faces - is covered exactly once by the space cell system, in the sense that every point is covered by one more positive than negative cell.* This property makes the space cell system suitable for playing the role of an equivalent continuum of granular assemblies.

#### 4. Continuum-mechanical state variables

This chapter summarises the definition of stress and strain tensors: these macro-variables will be expressed with the help of micro-variables, namely, the *branch vector*, the *complementary area vector*, the *contact force* transmitted between neighbouring grains, and the *relative translation* of neighbouring grains (see [1] or [2] for details). First let us introduce the two geometrical micro-variables.

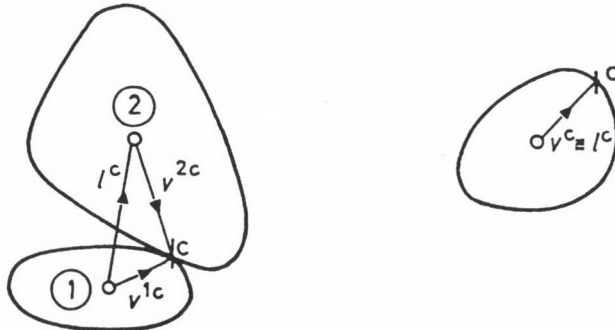


Fig. 14.

Consider a finite sub-assembly, and a material cell in it around a grain. Some of the faces of the cell belong to 'real' grain-grain contacts; a face like this contains the contact point itself. The rest of the faces are considered to belong to 'virtual' contacts (there is no contact between the grain and its neighbour, but their material cells have a common face); in these cases an arbitrary internal point of the face has to be chosen as virtual contact point. The vector showing from the centre of grain into the grain's  $c$ -th (real or virtual) contact point will be denoted as  $v_i^c$ .

Assume that the  $G_1$  and  $G_2$  grains have a (real or virtual) contact,  $c$ . The vectors  $v_i^{1c}$  and  $v_i^{2c}$  show from the corresponding grain centres to the contact point. The *branch vector* assigned to the contact is defined as

$$l_i^c = v_i^{1c} - v_i^{2c}$$

as illustrated in Figure 14. for 2D. In the special case when  $c$  is on the boundary of the sub-assembly (so it is a contact between a grain and the neighbourhood of the sub-assembly), the branch vector is defined to be equal to  $v_i^c$  (see Figure 14. again).

Now consider a space cell and number its nodes as 1, 2, ..., (D+1). (The cell is a simplex so it has (D+1) nodes.) Denote the faces of the cell by the number of that node which is not contained by the face (i.e. the  $k$ -th face contains all the nodes except the  $k$ -th node). Assign a vector  $b_i^k$  to each face in the following way:

→ The magnitude of  $b_i^k$  is equal to the area of the face (or length in 2D).

→ The direction of  $b_i^k$  is normal to the face, pointing outwards.

(It can easily be proved that  $\sum_{k=1}^{D+1} b_i^k = 0$  for any cell, for 2D and 3D too.)

The next vector,  $a_i^k$ , is defined as

$$a_i^k = -\frac{1}{D} b_i^k.$$

These vectors are illustrated in Figure 15. in 2D.

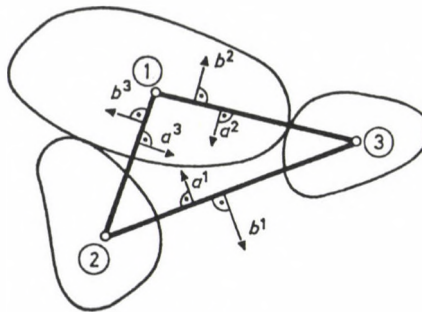


Fig. 15.

The  $a_i^k$  vector is the basis of the definition of the most important geometrical micro-variable of space cell system: the so-called *complementary area vector*. To construct it, consider a pair of grains,  $G_1$  and  $G_2$ , that have a (real or virtual) contact, so the two grain

centres, 1 and 2, are connected in the space cell system. Collect now *all* those space cells that contain this edge. Assume that altogether  $T$  cells were found; denote them as cell(1), cell(2), ...cell( $t$ ), ...cell( $T$ ). In the next step calculate the difference  $a_i^{1(t)} - a_i^{2(t)}$  separately in each cell from  $t=1$  to  $T$ ; after summing up for all space cells containing the 1-2 edge, the complementary area vector

$$d_i^{12} = \frac{1}{D+1} \sum_{t=1}^T (a_i^{1(t)} - a_i^{2(t)})$$

is given. (Its dimension is area in 3D assemblies and length in 2D.) This vector characterizes the local geometry of the neighbourhood of 1-2 edge.

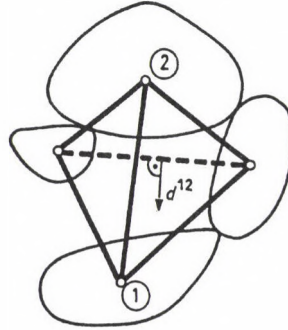


Fig. 16.

Figure 16. is an illustration in 2D where the 1-2 edge belongs to two cells shown by solid lines. The direction of  $d_i^{12}$  is as shown, and its magnitude is equal to the one-third of the dotted length. Similar - though more difficult to visualize - meaning can be found for  $d_i^{12}$  in 3D too.

The two geometrical micro-variables are of complementary nature in the sense that considering any finite sub-assembly, the sum of their scalar product for all edges of the space cell system is equal to the total volume of the space cells:

$$\sum_{(c)} l_i^c d_i^c = V$$

Now the state variables can already be defined. The stress tensor, the volume-weighted average of the stresses of material cells, can be expressed in terms of the branch vectors and the contact forces:

$$\bar{\sigma}_{ij} = \frac{1}{V} \sum_{(L)} v^L \bar{\sigma}_{ij}^L = \frac{1}{V} \sum_{(c)} l_i^c F_j^c$$

$L$ : index of material cells

$c$ : index of the faces of material cell system

$F_j^c$ : contact force between neighbouring grains

$l_j^c$ : branch vector



Note that the summation with index 'c' runs above all faces of the material cells of the considered finite sub-assembly (above the boundary faces too).

The average displacement gradient tensor of the space cells is, in terms of the complementary area vector and the relative translations:

$$\bar{e}_{ij} = \frac{1}{V} \sum_{(K)} V^K \bar{e}_{ij}^K = \frac{1}{V} \sum_{(c)} \Delta u_i^c d_j^c$$

K: index of space cells

c: index of the edges of space cell system

$\Delta u_i$ : relative translation of neighbouring grains

$d_j$ : complementary area vector

Note that the summation with index 'c' runs above all space edges of the considered finite sub-assemblies (all those edges whose both ends are within the sub-assembly).

The antimetric part of the average displacement gradient tensor expresses the average rigid-body rotation of the system. The symmetric part is related to the deformations of the material: by definition, this is the strain tensor of the assembly.

The two forms show a strong duality. Summation with index c runs through the same contacts in both cases (except from the boundary, see below); contact forces belong to the same pairs of grains as the relative displacements; and the two geometrical parameters are also dual to each other. But the duality is not complete since the two expressions can not belong to the same domain. The strain tensor is the average of the strains in the space cells; so the boundary of a domain where strain is meaningful has to go through the centres of the particles. On the other hand, the stress tensor is the average of stresses in the material cells so any domain where stress is defined is built up of material cells. As considering more and more grains and increasing the two domains further and further, the deviation between them, compared to the domain size, decreases. In the limit to infinity the difference tends to zero as the granular assembly tends to the continuum.

The fact that the two types of domains differ from each other seems to be a fundamental characteristic of granular assemblies in contrast to continua. Its physical meaning is that while the deformations of the material are carried on principally by the voids between the grains, the loads and stresses are transmitted by the grains themselves.

## 5. Summary

Two complementary cell systems, the material- and space-cell system, were introduced in this paper for the geometrical representation of granular assemblies. Both of them cover nearly-all points of the space exactly once; this property enables them to serve as equivalent continua to the material. The material cell system is the basis of stress definition while the strain is defined with the help of the space cell system. This is in good agreement with the theoretical expectations, and also with the physical experience that while the forces acting on an assembly are resisted by the grains (represented by the material cells), the deformations are carried on by the internal structure and the voids (reflected by the distortions of the space cells).

Beside their theoretical significance, the two systems can be applied for the modelling of any material having particulate internal structure. Their advantage to the recently used models (Voronoi etc.) is their much more general applicability.

#### ACKNOWLEDGEMENTS

This research has been supported by the Hungarian Scientific Research Fund (OTKA), under grant No. F-7641. Special thanks to Dr. J. Füzy for the valuable comments.

#### REFERENCES

1. Bagi, K.: Stress and strain in granular assemblies. *Mechanics of Materials*, 22, 165-177 (1996)
2. Bagi, K.: Analysis of granular assemblies. OTKA report, grant No. F-7641, Hungarian Academy of Sciences (1996)
3. Christoffersen J. - Mehrabadi, M. M. - Nemat-Nasser, S.: A micro-mechanical description of granular material behavior. *Journal of Applied Mechanics*, 48, 339-344 (1981)
4. Rothenburg, L. - Selvadurai, A. P. S.: A micromechanical definition of the Cauchy stress tensor for particulate media. *Procs. Intl. Symp. on the Mechanical Behavior of Structured Media*, Ottawa, ed. Selvadurai, Part B, 469-486 (1981)
5. Tonti, E.: The reason for analogies between physical theories. *Appl. Math. Modelling*, 1, 37-50 (1976)
6. Satake, M.: Discrete-mechanical approach to granular media. *Powders and Grains 93*, ed. C. Thornton, Balkema, 3-9 (1993)
7. Satake, M.: Discrete-mechanical formulation of granular materials. *Mechanics of Granular Materials* (as editor), report of ISSMFE, New Delhi, 1-6 (1994)
8. Schaefer, M.: Analysis der Motorfelder im Cosserat-Kontinuum, *ZAMM* 47. No. 5., 319-328 (1967)
9. Gellatly, B. J. - Finney, J. L.: Characterisation of multicomponent amorphous metals: the radical alternative to the Voronoi polyhedron. *Journal of Non-Crystalline Solids*, 50., 313-329 (1982)
10. Mindlin, R. D.: Micro-structure in linear elasticity. *Arch. Ration. Mech. Anal.*, 51-78 (1964)
11. Füzy, J. - Páti, Gy. - Vas, J.: Conclusion drawn from the analysis of a continuum mechanical model of four degrees of freedom proposed previously for the description of the behaviour of granular materials. *Acta Technica Acad. Sci. Hung.*, 101. No.1, 49-58 (1988)
12. Voronoi, G.: Recherches sur les paralléloèdres Primitives. *J. Reine Angew. Math.*, 134., 198-287
13. Dirichlet, G. L.: Über die Reduktion der positiven Quadratischen Formen mit drei anbestimmten ganzen Zahlen. *J. Reine Angew. Math.*, 40, 209-227



## NUMERICAL ESTIMATION OF THE ELASTIC PROPERTIES OF CEMENTED GRANULAR MATERIALS WITH MICROSTRUCTURE

BOJTÁR, I.\*

(Received 20 October 1996)

Elastic properties in the constitutive equations have a key role in the description of granular materials. This paper attempts to give an estimation for the value of elastic shear modulus ( $G$ ) and Young-modulus ( $E$ ) through the numerical analysis of microstructure in case of regular internal arrangements. The numerical method applied in the experiments, the geometrical structural models, and finally the results on the calculation of  $E$  and  $G$  will be introduced.

### 1. Introduction

Classical fenomenological (elastic, plastic, viscoplastic, fracture etc.) models of cemented granular materials require the knowledge of strength parameters that reflect the answer of the material to the external mechanical effects. Among these parameters the Young modulus ( $E$ ) and the shear modulus ( $G$ ) are generally considered to be the most important ones.

Physicists and engineers working in the field of material science usually accept that these macro-level parameters - as well as the other ones - are some kind of "averaged" characteristics of the behaviour of the microstructure, and they are determined by the micro-level strength parameters and the internal geometry. From that point of view of material science the trouble is that the problem of setting the "exact link" between the internal microstructure, and the macro-level behaviour known by experience and considered in the practical engineering work has not been solved yet. The recently existing methods contain very strong estimations and are rather questionable. This is basically due to the complexity of the microstructure: even in the case of a theoretically "regular" single-crystal there are several dislocations that influence the behaviour and modify, distort macro-level elastic parameters to deviate from the theoretically predicted values. The situation is, of course, even much more difficult in the case of a random irregular granular material: While for crystals at least an upper limit can be estimated assuming perfect geometrical order and considering the physical characteristics of the atoms, since not even an upper estimation can be given for the macro-level parameters.

Several authors have been dealing with the above problems. Among them, let us first mention the works of Bonzel whose monography [5] analysed different strength parameters

---

\* Bojtár, Imre, H-1181 Budapest, Barcsay u. 12.



of cemented granular materials, though without paying special attention on the internal granular structure. Carpinteri [6] concentrated on the effect of size on the strength parameters; Bagi [3] calculated the effect of the ratio of internal parameters of granular assemblies. Experiments and publications of Füzy [4] aimed to estimate the size of “representative volume” and to analyse the Cosserat-type behaviour of granular materials.

This paper concentrates on a special problem of this area. We would like to find a simple method to give an estimation for the macro-level strength parameter of a cemented granular material in terms of the micro-level geometrical data and the elastic characteristics of the contacts between the individual grains. The macro-level parameter gained this way could already be applied in phenomenological constitutive equations. Our analysis here will deal with the two parameters mentioned above: the Young-modulus ( $E$ ) and the shear modulus ( $G$ ).

This problem has an aspect even more important from theoretical point of view. A fundamental question of micromechanics of granular materials is to find the  $\Delta V$  representative volume (see for example Bazant, Oh [7]) that has a finite size in contrast to the infinitesimally small volume element of continuum-mechanics, since in the granular case an estimation is needed on how many grains should be considered in the analysis to reliably reflect the main characteristics of the material. If the size of domain necessary to predict the macro-level strength were found, this  $\Delta V$  could be estimated and then be applied as a material characteristic in the further engineering calculations.

Our researches were supported by numerical experiments executed by the algorithm PFC 2P (see chapter 2. for details). As a first try, different regular structures were simulated. The results verified the expectations: macro-level parameters could really be estimated on the basis of microstructural characteristics.

## 2. The numerical model

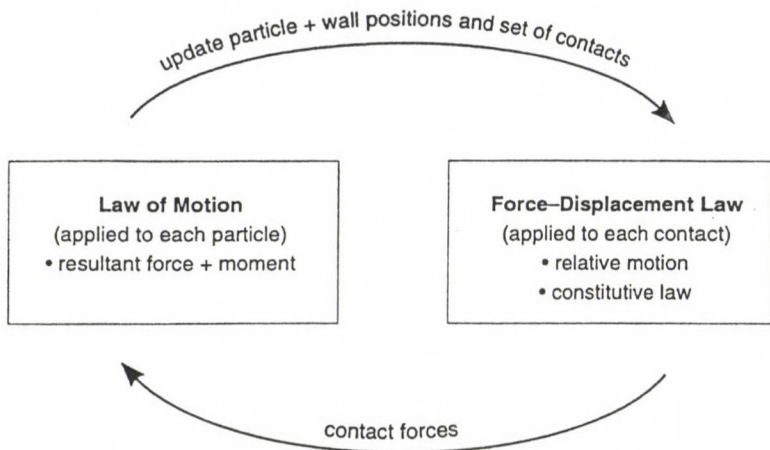
There were several possibilities for the numerical analysis of this problem. Two- or three-dimensional versions of either a quasi-static model developed by ourselves, or a commercial dynamic model made by P. Cundall [1], [2] were at our disposal.

Since our aim was to decide whether an estimation of macro-parameters is possible at all or not, the simplest and fastest version was chosen: the two-dimensional dynamic model. Its main advantage is its high executional speed: assemblies of thousands of grains can be analysed with about an order of magnitude faster than by the quasistatic model. It has to be emphasized that the 2D model was selected only for the sake of simplicity and higher speed. In the future we are planning to run 3D tests too that would probably make our results more exact.

The material is modelled as an assembly of discs. Its behaviour is analysed with the help of Newton's II. law; motion of each grain is followed in subsequent time-steps, according to the equations of dynamics. The micro-level strength and elastic parameters describing the materials of individual grains are considered as stiffness characteristics of the contacts between the grains. Tensional-compressional, shear and bending strength were assumed in our experiments. Walls may be defined to give boundaries of the assembly; they also have strength parameters, similarly to the grains.

The loads acting on the assembly cause contact forces between the grains according to the laws of Newtonian mechanics. Their calculation requires several (sometimes thousands

of) iterating time-steps, but the compilation and frequent modification of stiffness matrices of the quasistatic algorithms is avoided this way.



Constitutive behaviour for contact occurring at a point:

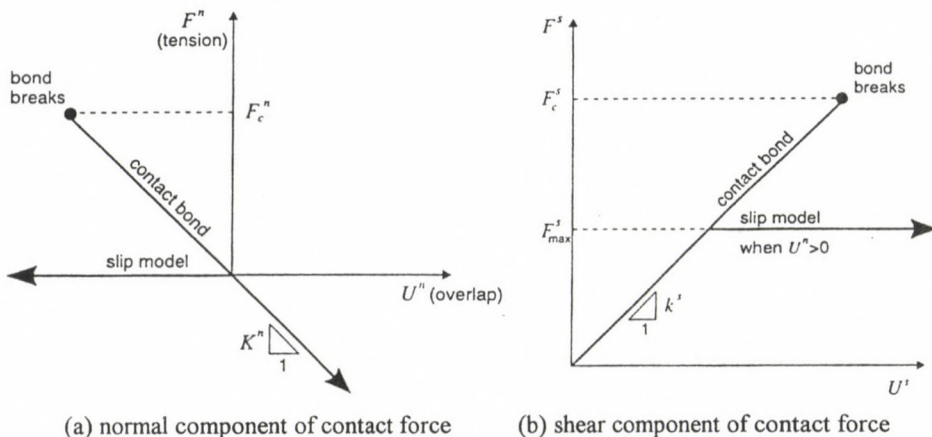


Fig. 1.  
The numerical model

Figure 1. illustrates the above principles (definition of grain-grain and grain-wall contacts, basic equations of the model, micro-level constitutive equations characterising the contacts etc.), bending stiffness of the contacts represent the effect of relative rotation of neighbouring grains.

Note that the contacts are, naturally, broken if the contact forces exceed the corresponding strength limit; in this case the cemented material can fall into smaller parts

having only compressional and frictional resistance between them. But this type of failure had no significance in our case: the analysis remained under the strength limit and concentrated on the behaviour of the intact material.

### 3. Numerical experiments

The 2D numerical simulations focused on the analyses of shear elastic modulus ( $G$ ) and tensional elastic modulus ( $E$ ): so shear and tensional tests were run. The shear model is shown in Figure 2.:

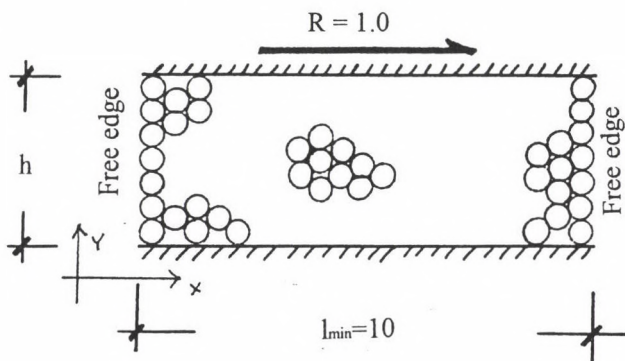


Fig. 2.  
Shear analysis

and the tensional model in Figure 3.:

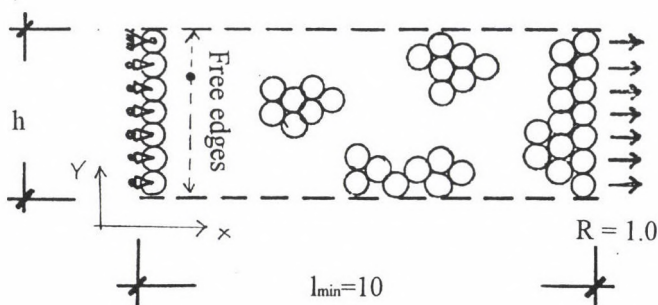


Fig. 3.  
Tensional analysis

As shown by the above sketches, the resultant of the external forces ( $R$ ) is equal to 1 in both cases. The boundary conditions: free edges, rigid fixed walls, and simple supported hinges, as denoted in the figures. Stiffness characteristics of the contacts are the same in all examples:



$$K_N = 10^{10}; \quad K_S = 10^9; \quad K_M = 10^8.$$

Length of the samples were taken in such a way that - to provide "sufficiently long" samples - the ratio  $L/h$  would always be higher than 10. Different thickness of sample ( $h$ ) and different radius of grains ( $r$ ) were applied in the different tests. This enabled us to analyse nearly continuous transition from the very thin-layered sample to the material consisting of huge number of grains. The uniform diameter ( $2r$ ) of the monosize assemblies took the values 0.1, 0.2, 0.3 and 0.4 in the different tests; the sample thickness ( $h$ ) was chosen in such a way that there would be at least two layers of grains in the material. The effect of internal geometry was also analysed: two different regular structures were applied. Figure 4/a illustrate the quadratic, 4/b the hexagonal packing.

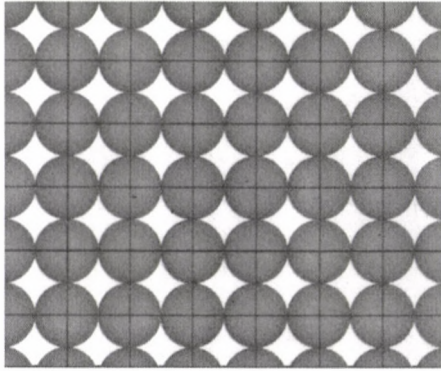


Fig. 4/a

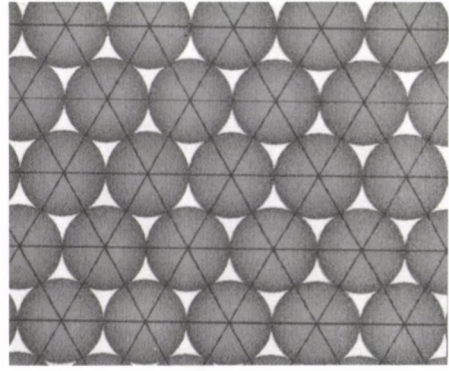


Fig. 4/b

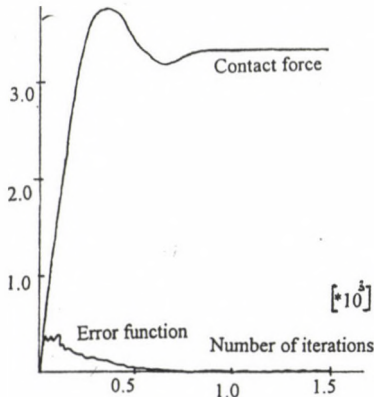


Fig. 5/a

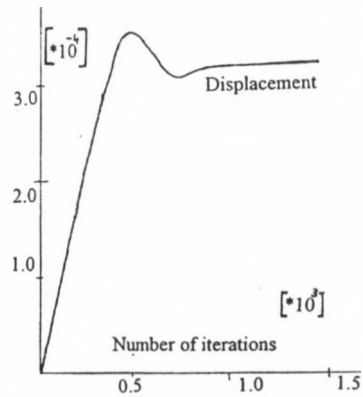


Fig. 5/b

During all tests, the iterations to find the equilibrated state were continued until the equilibrium error of contact forces (the magnitude of unbalanced forces acting on the individual grains) decreased under a pre-defined value ( $10^{-4}$  in our cases). Figure 5/a shows

a typical example on it. In one of the tensional tests 1342 iteration steps were necessary to approach the equilibrated state with the required exactness. The lower curve illustrates the decreasing of unbalanced forces (norm of equilibrium error) while the upper curve shows the evolution of average normal force (the stable state is reached when the norm of equilibrium error becomes smaller than the acceptable error norm magnitude). Figure 5/b belongs to the same test: displacement of a chosen grain is shown, and as the error norm decreases, the displacement becomes more and more stabilized near its final value. Horizontal axes show the number of iteration steps.

Next, examples will be introduced on how the internal forces change during the loading process. Results of a shear test on a hexagonal sample can be seen in Figure 6/a. Contact forces on the left part of the material are shown with their direction and magnitude; positive or negative sign of normal forces also distinct. Figure 6/b belongs to the same sample under tension; at both examples the middle of the sample was given.

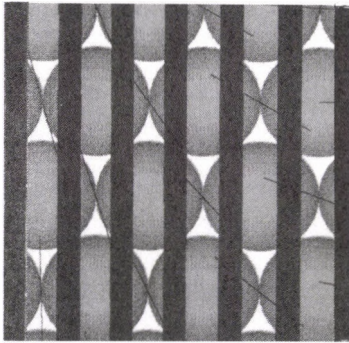


Fig. 6/a

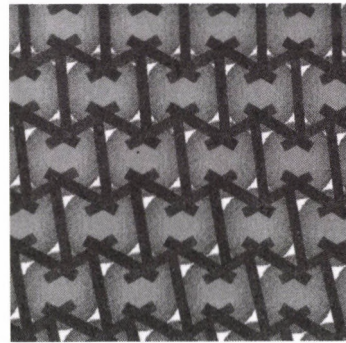


Fig. 6/b

Contact force distribution

The evaluation of the results will start with the elastic shear modulus,  $G$ . Figure 7. belongs to the hexagonal arrangements:

Value of  $G$  was calculated from the horizontal displacement of the upper layer, due to unit horizontal shearing force. In case of minimal thickness (two layers) the value of  $G$  was orders of magnitudes higher than in all other cases (see the arrows in the graph); then, after a "wave", as increasing the number of layers the value of  $G$  tended to the same limit, independently of the diameter of grains.

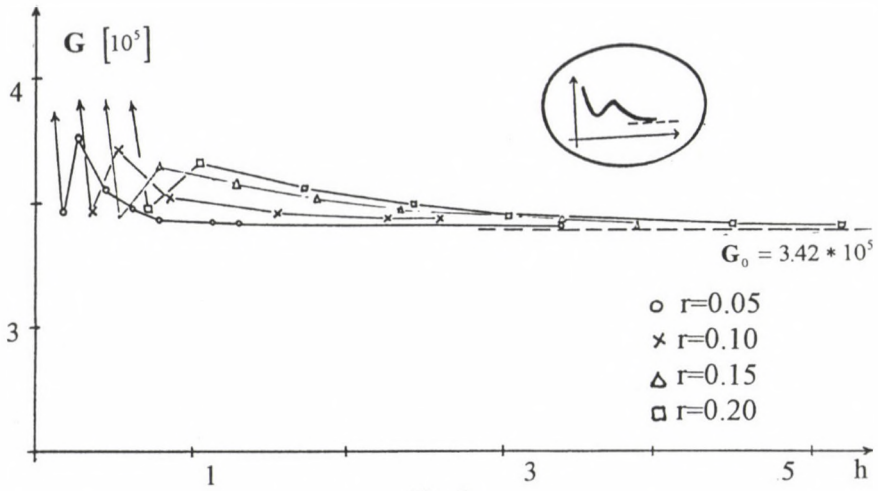


Fig. 7.  
Hexagonal arrangement; shear test

The experiences in case of quadratic arrangement are shown in Figure 8.:

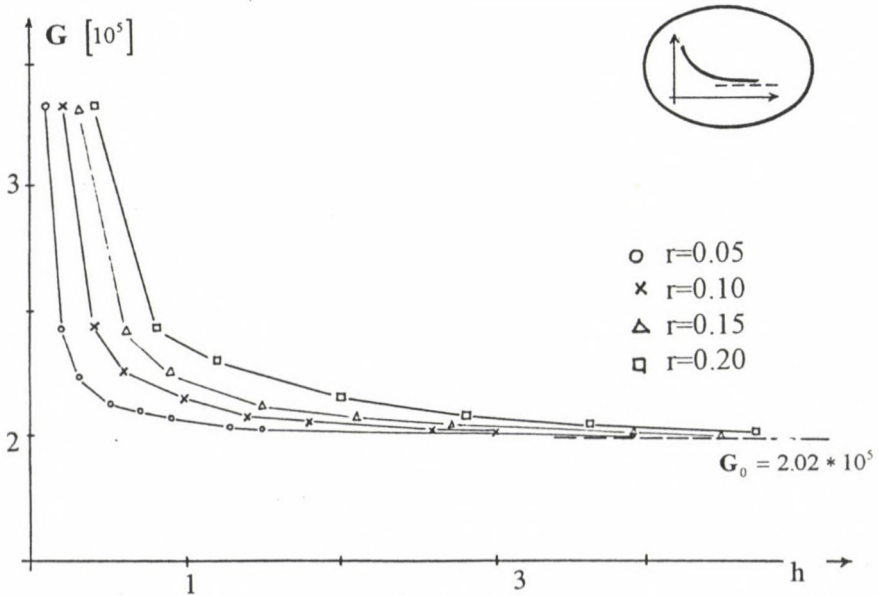


Fig. 8.  
Quadratic arrangement; shear test



The changes are smoother here and all values are about the same order of magnitude. The limit value of macro-level shear modulus is the same for all diameters. However, magnitude of  $G_0$  is smaller now than at the hexagonal arrangements.

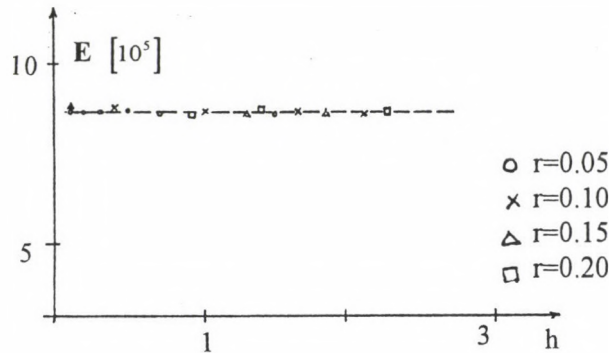


Fig. 9.  
Hexagonal arrangement; tension test

The next figures, Figure 9. and 10. introduce the results of the analysis of  $E$ , the tensional elastic modulus. The speciality of the hexagonal structure is that  $E_0$  was constant, independently of the layer thickness and diameter.

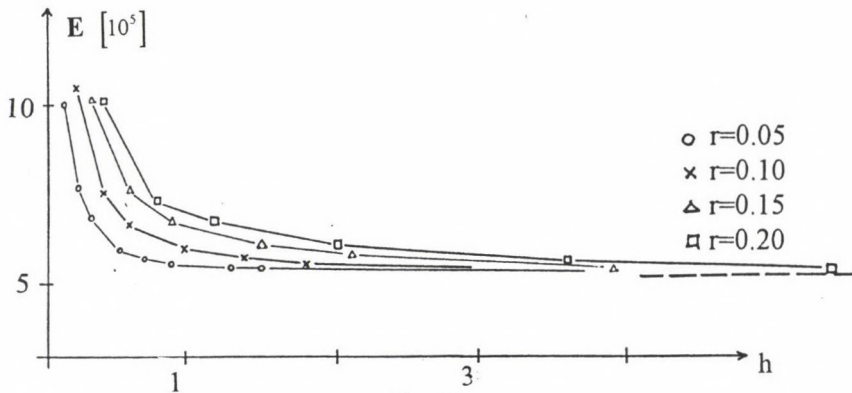


Fig. 10.  
Quadratic arrangement; tension test

The behaviour of rectangular structure (Figure 10.) is different: it was similar to that in Figure 8., and  $E_0$  was smaller here than at the hexagonal system, similarly to what was experienced about the shear modulus. It has to be emphasized that the results in Figure 10. do not follow from the Cosserat theory, and cannot derive from the relative rotations of neighbouring grains in diagonal directions. The explanation of this behaviour needs further analysis.

#### 4. Concluding remarks

The above numerical simulations led us to the following conclusions:

→ The main macro-level material strength parameters can be estimated on the basis of internal geometrical and strength characteristics of the microstructure.

→ In case of regular packings a definite  $f(d, h, K_N, K_T, K_M)$  function can be found for the calculation of E and G ( $K_N, K_T, K_M$  are the local stiffness characteristics of the grain-grain contacts).

→ The hexagonal microstructure has a more rigid macro-level behaviour than the rectangular arrangement.

→ The hyperbolic-type functions found at the rectangular structures are very similar to the functions of parameter estimation of Cosserat continua. A detailed theoretical analysis may be found for this type of behaviour in the paper of Füzy [4]. The limit-layer parameter of Cosserat, a basic variable in that theory, is unambiguously provided by our results.

→ Our experiments were restricted for the simplest cases only. Detailed analyses of the effect of micro-level strength parameters ( $K_N, K_T, K_M$ ) is indispensable in the further work. Even more important and complicated is the analysis of randomly packed irregular assemblies with the evaluation of the effect of grain-size distribution is not sufficient to consider; the internal geometry should probably be characterised with some kind of special geometrical variable like, for instance, the fabric tensor.

→ A different question - leading already towards the analysis of the behaviour of non-elastic characteristics - is to see how the above experiences change if the contacts may break when reaching the strength limit, making microcracks in the material.

#### ACKNOWLEDGEMENT

The above experiments were supported by the Hungarian Scientific Research Fund (OTKA), under grant No. T-014270.

#### REFERENCES

1. Cundall, P. A.: Distinct Element Models of Rock and Soil Structure. Analytical and Computational Methods in Engineering Rock Mechanics, ed. E. T. Brown (1987)
2. Cundall, P. A.: Numerical Experiments on Granular Assemblies. Deformation and Failure of Granular Materials, ed. H. Luger (1982)
3. Bagi, K. - Bojtár, I.: Numerical analysis of state changing processes in granular assemblies. Periodica Polytechnica, 33 (3-4), 183-196 (1989)
4. Füzy, J. - Vas, J.: Proposed continuum model for simulating the behaviour of granular materials. Acta Technica, 95, 49-53 (1982)
5. Bonzel, A.: Variation of strength parameters in concrete. Scientific report, IUTAM, London (1965)
6. Carpinteri, A.: Size effect in fracture toughness testing. Proc. Int. Conf. Analytical Fracture Mechanics, Rome, 23-27 (1980)
7. Bazant, Z. P. - Oh, B. H.: Crack band theory of concrete. Mater. Struct., RICEM 93, 155-177 (1993)





## COMPARISON OF TENT STRUCTURES CALCULATION IN HUNGARY AND FRANCE

GALASKÓ, GY.<sup>\*</sup> - GÁSPÁR, ZS.<sup>\*</sup> - NOURI-BARANGER, T.<sup>\*\*</sup>  
LEON, J.C.<sup>\*\*\*</sup> - TROMPETTE, P.<sup>\*\*</sup> - VERON, P.<sup>\*\*\*</sup>

(Received: 26 November 1996)

Tent structures are an interesting field of civil engineering, There exist special methods for their computation. There are only a few companies which are specialized in the design and constructing of tent structures in Europe. This paper compares the basic ideas of two working groups, each of which developed its own methods of calculation from the idea of the architect until the final cutting patterns. One example demonstrates the results of the two methods. The F 96 program between the French and Hungarian government helped the authors of this paper.

### 1. Calculation of tent structures in Hungary

At the Technical University of Budapest there exists a small group which developed algorithms and computer programs applying the following theory.

Textile is not able to keep pressure therefore the tent structures need to be designed in such a way that they can carry every possible load without pressure. In practice the meteorological and other possible loads are definitely more than the dead load, so the initial form is designed considering the unloaded structure. Under different loads, in some parts of the structure the normal stress may be reduced. Wrinkles can be avoided if the unloaded structure has a state of self-stress. This is possible in the case of statically indeterminate (and kinematically overdeterminate) structures. (The kinematic overdeterminacy is released by the self stress field's secondary stiffening effect.)

Tent structures are membranes which are not able to carry bending moments, hence they are generally kinematically indeterminate (statically overdeterminate) structures and self-stresses exist only in the case of a special geometry. The architect has a preliminary idea about the form of the tent which can satisfy the different functional, aesthetical and other requirements. The first step is to find a form which is nearest to the desire of the architect and has its own satisfactory stress field.

The question is, which stress distributions are satisfactory? The answer is that those which can avoid pressure in every part of the textile under all possible load combinations and the maximum stress is less than the limit. During the previous step (form finding) only static equilibrium was considered; in this step, involving the computation of the state-change of the structure, already requires the consideration of constitutive and geometrical equations. The

<sup>\*</sup> Technical University of Budapest, Department of Structural Mechanics, H-1521 Budapest, Hungary

<sup>\*\*</sup> Laboratoire de Biomécanique du Mouvement, Université Claude Bernard de Lyon 69622 Lyon, France

<sup>\*\*\*</sup> Laboratoire Sols, Solides, Structures, Domaine Université BP53X Grenoble Cedex 9, France

form finding generally results in a linear equation system, however, the computation of the state-change produces a non-linear system and requires information about the materials, as well.

The engineering practice suggests initial values for these parameters. After determining the values of the state variables in the first step results are checked and if modification of the input data is necessary then the calculation starts from the beginning.

If the above process yields a satisfactory form and corresponding self-stress state, starting from which all new configurations (corresponding to various loads) are satisfactory, it still remains an open question, how to produce the initial form. By cutting and gluing pieces of planar textile, the prescribed initial form can not be reached. By applying suitable changes to the locations of boundary points, stresses can be reduced to zero, however, the resulting surface is not developable. (The shape of the stress-free textile is not uniquely determined because of the kinematic indeterminacy, however, the admissible displacements do not change the developability.) The only realistic goal can be to find a cutting pattern resulting in a shape close to the previously determined initial shape.

During the last decades the research group at the TUB developed a program system for the solution of the above problems. This package contains computation and visualization modules. Computation modules can be divided into the following three groups:

- a) the initial form
- b) state change due to loads
- c) cutting pattern and shape.

Let us review the models applied for the different goals, computation methods, their advantages and disadvantages.

### 1.1 Calculation of the initial form

We applied a developed version of the method suitable for the determination of equilibrium of a cable net, the projection of which is an orthogonal grid. (This method characterizes the shape of the surface by the coordinates of the cable nodes, other points are computed by interpolation.)

In the first version [18,6] we applied the cable net model as well, with a relatively large choice of boundary conditions (rigid boundary with arbitrary shape, masts with prescribed locations, boundary cables). In those times the small memory capacity and low speed of computers represented a serious handicap, so the algorithm was based on the before mentioned assumption of the orthogonal projection. In this special case the horizontal equilibrium of the nodes yielded immediately the result that the horizontal components of the cable forces are constant. Fixing this two components the vertical equilibrium can be expressed by the following linear equations system

$$A X + X B = Q$$

where the matrices A and B depend on the distances between the cables and the prescribed horizontal force components. Matrix X describes the unknown heights of the nodes and Q is the load matrix. The size of this four matrices is  $n \cdot m$  where n and m are the number of the



cables in the two horizontal directions  $x$  and  $y$ . The spectral separation of the matrix  $A$  and  $B$  produces  $n*m$  simple equations for the components of  $X$ .

If the fix boundary is not rectangular and/or we apply masts as well, then the problem can be reduced to the previously described one by the repeated application of the method of singular loads. In the case of edge cables before the solution of the above mentioned process the horizontal coordinates of the edge cables can be computed from the horizontal equilibrium of the network and the magnitude of the horizontal components of the cables. Knowing the  $x$  and  $y$  coordinates of the edge cables the singular loads method produces the height at the surface as well.

Our present method relies on the higher speed and memory capacity of the computers. It abandons the singular load method and the uses one big linear equation system, which originates from the vertical equilibrium of every point (including the points of the edge cable), and provides the heights of the nodes. This program permits a large variety of different boundary conditions.

The results (the calculated points of the surface) are visualized with AUTOCAD program system giving good possibility to the architect and the structural engineer for spotting the weak parts of the structures and for the modification of it until the satisfactory shape is reached. Several independent ways of modification exist: The fix point coordinates (where the rigid boundary or/and the ring around the mast intersect a cable of the rectangular network); The coordinates of the endpoints of the edge cables; ordinates of the endpoints; The angle between the straight line connecting the endpoints of the edge cable and the initial tangent of the cable; The magnitude of the horizontal (stretching) forces. The last possibility, which allows different magnitude of each cable in the  $x$  or/and  $y$  direction, is useful in the case of valley cable(s). The program can handle two symmetry axes, reducing the computer time.

A mast means a singularity point of the tent surface, because the concentrated force there can be equalized very near to this point only with infinite stresses. Therefore in practice the textile is carried by a ring which hangs on the top of the mast. In the case of our cable network model every cable which intersects this ring, provides a boundary point with the given height of the mast. In a typical (usual) case four points describe the top area well but often more points are applied. A usual joint point division can not produce enough smooth surface if the cutting patten calculation program requires the interpolation between the internal points.

To reduce this effect an other program was developed which calculates the initial form in the vicinity of the mast more precisely. Instead of the rectangular cable network a polar system is applied. The inner points are the coordinates of the rigid ring with their fixed (previously defined) points, and the farrest are interpolated points of the first shape. In this model the magnitude of the projected cable forces are prescribed as well (the forces of the rings and one radial direction are mathematically independent). From this input data every other force magnitude is computable by using the horizontal equilibrium of the structure. The vertical equilibrium of this polar system produces the height of the points at the intersection points of the radial and annual cables. This method results better form near to the mast but sometime produce a compatibility error at the end of the polar coordinate system. Sometime the increasing of the number of points in the rectangular system provides sufficient precision without applying this second program. At the same time this program offers the possibility of the calculation of the exact position of the mast. Because the points of the top ring are prescribed from the directions and magnitudes of the radial cables, the coordinates of the resultant are computable, and this is the theoretical position of the mast.

Nowadays an other program is being developed for the shape finding ,applying triangulation of the membrane surface. This model uses the dynamic relaxation as the method [2].



## 1.2 Computation of the state-change

The initial form is generally calculated without external loads, in spite of the fact that the algorithm allows the effect of arbitrary vertical force. This does not mean that it is sufficient to run the shape finding program twice (firstly on the unloaded structure, secondly the actual loads are calculated), because the two results describe two different problems. Therefore an other program is necessary to calculate the stresses and displacements of a tent structure.

Our model is the same as in the previous step: it is a cable network, furthermore this program allows fully general geometry from different materials. In the case of a real cable network the shear forces between the cables are supposed to be small, consequently, instead of the orthogonal arrangement the principal curvature [9] or geodesic line [17] applications are advisable. When modeling textile structures with cable nets the shear stresses are not very important, thus the simplest way is to keep the rectangular cable network of the form finding, because the same data structure offers big advantage.

What are the drawbacks of the cable net model? In this model one member of a cable has constant cross section, although the width of a textile strip is not constant. Theoretically it does not result an error because in the whole model the total elongation of the member is counted and the average width describes it in this respect well. In practice the width is calculated from the top projection plan, in spite of the fact that the real cross section area depends on the vertical angle too, however, this effect is neglected.

The stress-strain diagram of a textile is generally nonlinear. A short part of this diagram is nearly linear, so a polygon is applied in stead of the real diagram.

The material of the textile is anisotropic but in the stage of the calculation the fiber directions are unknown. The angle between this two main directions of the material and the direction of the member are neglected. The perpendicular cables cover the surface twice: the same material works in x and y directions as well and the behavior of the material in this two directions is not independent: the Poisson number gives the connection.

A rectangular part of the network may deform into rhomboids under load and shear stiffness of the textile influences this deformation, while in the applied model only the cable forces secondary stiffening effect limits this deformation. (The model allows diagonal members to model the shear effect but in practice they are not used because the most common textile in Hungary has low shear modulus. [20])

A part of the problems mentioned above originate from the model itself, therefore these problem can not be eliminated. Until recently we did not pay much attention to the applied width of the strip because the uncertainty of the materials and the loads are high. In the textile only a few member types have been considered, but the different steel cables (edge cables or anchorage) have been described more precisely.

An other interesting question is the modeling of the connection of the steel edge cable and the textile. The cable is placed at the pouch of the textile. The textile is fixed at two ends and only the friction between the textile and the cable prevents the relative displacement. Hoping that the stresses from the textile are almost perpendicular to the cable, the friction has been neglected.

The program allows concentrated force with arbitrary spatial directions at the joints. The two most important loading cases are the wind and the snow. The program automatically calculates the effect of the wind from arbitrary directions with an approximate form coefficient.

The actual status of the structure is computed by a modified Newton-Raphson iteration method [18]. Every step of the iteration calculates the unbalanced force from the equilibrium



equations and the incompatibility of the members. The iteration is finished when both of the two errors of the calculation are less than a small given number. If the errors exceed the limits then the incompatibility lengths are converted into unbalanced forces at the node and the tangent stiffness matrix of the structure is recalculated from the actual nodal coordinates and internal forces. (During the iteration process some cable forces may be negative - pressure - which decrease the stiffness of the member and may produce numerical problem. Therefore only the tension forces are taken into account in the tangent stiffness matrix. This neglection does not influence the final result because the error vectors are computed from the exact formulas.)

The tangent stiffness matrix and the summarized vector of the unbalanced forces produce a linear equation system. The solution of this system results new coordinates of the nodes. A useful improvement of the program is the limitation of the nodal displacement. If the equation system produces bigger displacement than the given limit then every displacement is reduced with same factor, which is computed from the ratio the maximum and the given limit. This simple modification speeds up the iteration.

The new internal forces of the cables are calculated from the new coordinates of the nodes with a linear formula. It produces definitely better convergence as if the cable forces were computed from the exact elongation.

### 1.3 Calculation of the cutting patterns

In our opinion the calculation of the cutting patterns is reasonable only if the shape is suitable for the architect and the structural engineer as well, namely the tent satisfies every functional and static requirement.

The textiles, which are produced by factories, have a given width and the next goal is to find the most economic 2D cutting patterns of the 3D shape. In practice developable surfaces are rarely applied when the flattening is simple. A typical surface with double curvature is undevelopable and the goal is the best approximation. In Hungary the triangle method is the most widespread [7]. The essential idea of this method supposes that the space length between two points of the real surface is the same than the distance measured on the strip of the textile in the plane.

The method requires two zigzag lines on the surface which define a sequence of triangles. If the space line requires internal points, then they are interpolated by using linear approximation. The flattening is based on the common nodes of these triangles.

Two main tent groups exist from the viewpoint of the cutting pattern. Generally it is a more simple case when the textile is stretched between fix boundaries. It is a typical application of this structures when the loads are carried by parallel arches. In this case the direction of the cutting patterns may be parallel or perpendicular to the plane of the arches. (It depends on the ratio of the spans.) In that case the rectangular cable network model describes the real behavior very well and the calculated points provide directly the strips of the cutting patterns.

The other frequently used tent structure applies mast(s) to support the textile. In this case a polar coordinate system is used around the mast and the radial lines give the edges of a strip, meanwhile the zigzag connects them. Around the center (mast) generally one central angle is applied, which sometime results very different width at the other end of the strip. We prefer this solution against the other one, which tries to optimize the maximum width at the wider end or uses the length at the edge. Especially the second solution can produce big difference in the angle of the center, which is not very attractive.

The flattened cutting patterns often has big curvatures producing wider cutting pattern than the width of the material. In this case an other (smaller) central angel is necessary and the whole process starts again. Very small cutting patterns are not economic because of the lot of welding between the neighboring strips. This simple final step often requires much effort while the final form is developed.

## **2. Design and mechanical analysis of tent structures in France**

Due to the multiple actors involved in the design process of a textile structure i.e.: the customer, the architect who designs the structure, following the requirements of the customer, the textiles manufacturers who must satisfy the architect's aesthetic criteria and minimize the fabric wasted, the engineering office which has to control the whole process from mechanical and structured point of view, an integrated approach has been set up to carry out efficient dialogues between all these persons. This dialogues is centered on the tasks of the engineering office and try to use the information exchanges with partners. It has been proposed to divide the design process in three main steps: the first one related to the definition of three dimensional shape of the textile structure, the second one focuses on the definition of its manufacturing features and the third is centered on the structural analysis of the textile and the metallic parts. The main idea is to build in next future a software package which makes possible the generation of a real design loop where the different users can interfere at each stage to modify their interior data and choices even if they are located at different geographic places.

### 2.1 Three-dimensional shape finding

At first a topological model of the textile structure is defined both from its fixed parts: anchorage points, masts, bars, arches which contribute to maintain the membrane, and from the panels created from the previous fixed parts. Then this pure geometric support, node coordinates and connected nodes are the basic data of a topological mesh on which the force density method [16] is applied. This means that the equilibrium position of a cable net, which has the same topological definition than the geometric support, is obtained giving force densities in the cables, and constraining boundary points to be fixed or to move along the giving boundary. Afterwards, this cable net shape may be modified by the architect by changing the force densities values, until a convenient final shape is reached. Finally this polyhedral model is automatically converted into a C1 continuous surface model approximation containing Bezier patches. The surface decomposition into patches is related to the topological meshes initially created as each of the patches is computed to provide the best approximation of its bar network subset counterpart. Thus this surface model is defined as the basis of the subsequent phases.

### 2.2. Cutting Patterns

The second phase focuses on the creation of the manufacturing features of the textiles structure. The surface model of the textile structure previously defined allows the results of the flattening technique to be independent from the parameter used during the first phase (for example the discretization used to build the polyhedral model). Thus the 3D cutting patterns boundaries are directly defined over the surface model according to the requirement of the



architect, the designer and the manufacturer choices. Then an automatic flattening process produces 2D contours that will be cut in the fabric rolls. The user can choose one among several methods ranging from fast to robust ones in accordance with the quality desired [1, 7, 19]. Improvements of the quality of the cutting patterns obtained may be required when 2D shapes have global curvatures that are not acceptable from a manufacturing point of view, because they would generate large wastes of textiles or reduce the load capacity of the structure. A direct and interactive correction process applied to the boundaries of 3D cutting pattern helps the user to modify the flattened 2D cutting patterns in the area of interest. Finally the engineering drawings are made after a length adjustment of the different pieces has taken place. Because the flattening process is carried out on an independent basis for each cutting pattern, the 2D boundaries obtained for two adjacent cutting patterns usually have different lengths. This justifies the length adjustment process previously mentioned. The two sets of 3D and 2D cutting pattern boundaries constitute the manufacturing model of the textile structure. Like in the first phase, the second one also makes it possible to go back to any of the previous states and to modify various parameters. The interactive and real time improvement of the 2D cutting patterns is an important process to obtain the validation of the cutting patterns by different actors.

### 2.3. Mechanical behavior of the tent structure

Here again the surface model of the textile is used as an important parameter in order to ensure the independence of the data structures generated during the first phase with respect to the mesh used for the finite element computations. The latter is a subset of mechanical models of the membrane part of the textile structure. The first step consists of checking the equilibrium position of the textile stressed by different climatic loading, taking into account the initial stresses produced when we join all 2D cutting patterns and build 3D shapes. To this end, a finite element (F.E.) mesh is created from the surface model and the stress state computation helps to verify the textile structure which should be everywhere under an adequate level of tension. Then the boundary forces obtained from the previous computations are used for the dimensioning process of the metallic structure and civil engineering parts. This step can be made can be precise in the sense the F.E. calculations take into account the large membrane displacement theory, the initial stresses and the anisotropy of the constitutive fabric material.

### 2.4. Software architecture

Seeking compromises of agreements between the actors at various stages of the design process through meetings and discussions requires a specific environment which include computer supported collaborative work tools [10] capable of exploiting computer objects to achieve collaborative work between actors [6]. The various sites are interconnected through a computer network. These collaborations are carried out with a multimedia software tool, that allows the users to simultaneously share a 3D graphic scene displayed on their respective workstations. The collaborations between actors that have different working themes require the use of dialogue objects which must be as realistic as possible and incorporate the data placed at the center of the mediation.

### 2.5. Conclusion and perspectives

The design methodology and the data architecture elaborated fit into the design process of textile structures and incorporate the actors involved in this process as well as their know-how. The coupling between the geometry of the structure and the mechanical stress state of its fabric has been treated through the decomposition of the knowledge of each actor. This decomposition has been worked out with the use of appropriate approximation methods and automatic conversions of the data between successive models. The identification and adaptation of the specific models (or views) of the structure to each of its definition levels and to each of the actors using it, is a key point of the methodology introduced. The collaborative environment developed participates in the convergence of the design process and improves the efficiency of the dialogues between the actors by the use of shared graphic objects and annotation tools. However, further progress can be made in this area through the use of true design objects to ease the modifications specified during a dialogue. Future work will aim at developing an effective mechanical analysis view of the overall structure and an integrated view of the whole design environment.

### 3. Example

The following example, so called "Chinese Hat", was calculated by the two program systems. Fig. 1. shows that the boundary conditions are simple: the lengths of the fix edges are 4 meter and the height of the hoop is 2 meter. The diameter of this top ring is 50 cm.

The Hungarian program requires the magnitude of the horizontal components of the stretching forces. They are the same value: 1 kN in X and Y direction. The form was calculated by rectangular mesh while the Fig. 1. shows the interpolated points in polar system.

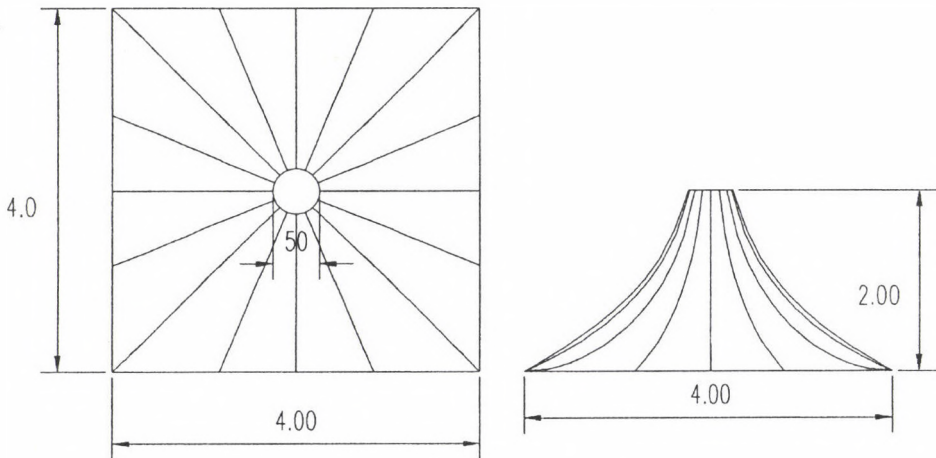


Fig. 1.

The model of the French program consisted initially 9 circles and 16 radial lines. The force density of the element at circles was always 1 and at the other elements is denoted by  $d$ . Fig. 2. shows a few variations.  $d=0.5$  force density produced almost the same cross section at



the diagonal of the tent as the Hungarian method. It may be an interesting new research area to compare the form finding methods. The  $d=0.5$  form was taken into account at the calculation of the cutting patterns.

The cutting patterns are shown in Fig. 3. Their forms are significantly different at the middle line. The reason originated from the different technics of the two calculation methods. The Hungarian method divided the surface into two elements with a straight line in the top plan while the French program has a curve. It would be an other interesting topic of our future common research to compare and develop together the calculation methods. Interesting aspects are the area and perimeter of the current cutting patters which are shown in the next tabulation.

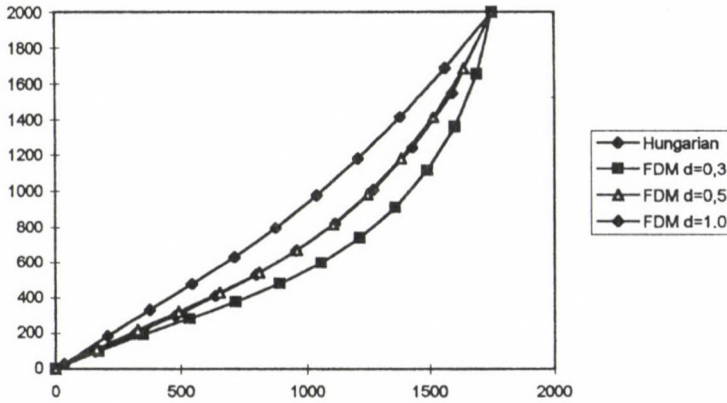


Fig. 2.

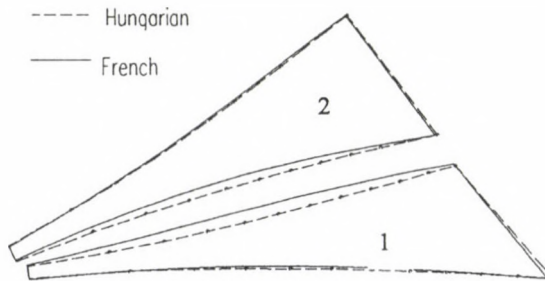


Fig. 3.

	Hungarian			French		
	Pattern 1	Pattern 2	Total	Pattern 1	Pattern 2	Total
Area [m <sup>2</sup> ]	1.209	1.313	2.522	1.313	1.212	2.525
Perimeter	7.543	6.803		7.514	6.772	



### Acknowledgement

The research reported here was partly supported by the Hungarian Scientific Research Foundation (OTKA Grant no: T015351) and the scientific and technological cooperation between the Hungarian and French government (sponsored by OMFB and Le Misteré des Affaires Etrangères).

### REFERENCES

1. Allera, R.: Mise en forme des structures textiles tendues, PhD thesis, Institut Polytechnique de Grenoble, Grenoble, France, 1992
2. Barnes, M.R.: Form finding and analysis of prestressed nets and membranes. *Journal of Computers and Structures* 1988
3. Beziers, P.: Mathematical and practical possibilities of UNISURF, in *Computer Aided Geometric Design*, 127-152, edited by R.E. Barnhill and R.F. Riesenfeld, Academic Press, 1974
4. Coons, S.A.: Surfaces for computer aided design of space forms, MIT Project MAC-TR-41, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1967
5. Farin, G.: *Curves and Surfaces for Computer Aided Geometric Design*, Academic Press, 1988
6. Favela, J., Xong, A., Chakravarthy, A.: Supporting Collaborative Engineering Design, *Engineering with Computers*, 9 (1993), 125-132
7. Galaskó, G., Popper, G.: Identification in the Design Process of Tent Structures. XI. Polish Conference on Computer Methods in Mechanics. (1993), 273-278
8. Galaskó, G., Trompette P.: Calculation of cutting patterns using triangles, Proc. of the Euromech 334 Int. Conf. on Textiles composites and textiles structures. Lyon, France, 15-17 May (1995), 327-337
9. Gáspár, Zs.: Nach Hauptkrümmungslinien konstruiertes Seilnetz. *Acta Techn. Hung.* (1976), 103-116
10. Gomez-Molinario, V., Vilanova, J., Grave, M., Lang, U.: Use of Computer Supported Collaborative Work (CSCW) in Structural Analysis and Mechanical Testing of Spacecraft, *Spacecraft Structures, Materials and Mechanical Testing Conference, ESA/ESTEC Noordwijk, March 1996*
11. Haber, R.B., Abel, J.F., Greenberg, D.P.: An integrated design system for cable reinforced membranes using interactive computer graphics, *Computers and Structures*, 14 (1981), 261-280
12. Hangleiber, H., Gruending, L.: Cutting pattern for structural membranes, Proc. of the first Int. Conf. on Lightweight Structures in architecture, Sydney, 1986
13. Léon, J.C.: Modelisation et construction de surfaces pour la C.F.A.O., Ed. Hermes, Paris, 1991
14. Noel, F., Léon, J.C., Trompette, P.: A new approach to free-form surface mesh control in CAD environment, *Int. J. for Num. Methods in Eng.*, 38 (1995), 3121-3142
15. Noel, F.: Mailleur auto-adaptatif pour des surfaces gauches en vue de la conception intégré, PhD thesis. Institut National Polytechnique de Grenoble, France, October 1994
16. Scheck, H. J.: The force density method for form finding and computation of general networks, *Comp. Meth. in Applied Mech. and Eng.*, 3 (1974), 115-134
17. Szabó, J., Gáspár, Z.: Berechnung des auf Randkabel gespannten rechtwinkligen Seilnetzes. *Acta Techn. Hung.* (1974), 365-384
18. Szabó, J., Kollár, L.: *Structural Design of Cable-Suspended Roofs*. Akadémiai Kiadó. Budapest. 1984
19. Shimada, T., Yoshida, A.: Development of a curved surface using a finite element method, *Int. Conf. on Computer Aided Design of Structures: recent advances*, Southampton, 1989
20. Trompette, P., Quang, N.T., Denes, H.: Application of the off-axis tensile test for the determination of elastic coefficients of prestressed coated fabric. II. *Textile Composites in Building Construction Lyon*. (1992) 193-202

## APPLICATION OF GRÖBNER BASIS THEORY TO FIND GLOBAL EQUILIBRIUM PATHS OF THE SIMPLE ARCH

GÁSPÁR, ZS.\* - KÁROLYI, GY.\*\* - POPPER, GY.\*\*\*

(Received: 12 August 1996)

Application of the Gröbner basis theory is presented for finding the global equilibrium paths of the model of a simple arch. Gröbner basis calculations (with respect to pure lexicographic term ordering) make the solution of polynomial equation systems easier and may provide us with analytical results. This makes possible the determination of the global equilibrium paths as explicit functions, and the examination of their exact parameter dependence. Even for this simple model unexpected solutions were found with symmetry breaking property or with negative bar length. The physical meaning of the latter is explained by means of non-homogenous bars with a certain internal structure.

### 1. Introduction

Structural engineers tend to consider the loads of a certain structure to be *one-parametered*. Starting from a known, stable equilibrium state, they try to determine the properties of the equilibrium states as a function of the unique load parameter. Generally they are satisfied by following the equilibrium path of the starting state up to the *first critical* (limit or bifurcation) *point*. For the characterization of the *sensibility to the imperfectness*, the determination of the type of the critical point and of the starting branches of the postcritical equilibrium paths may provide useful information.

Several recent papers [1,2] are devoted to the determination of the *global equilibrium paths*, when not only the starting branches of the postcritical paths are aimed by the calculations, but also their further segments together with the possible new limit or bifurcation points. Besides, it may be chosen as a goal to find, for a certain range of the state variables, the equilibrium paths not connected to the original one.

To this end we may apply e.g. the *simplex method* [1,2], or the *Gröbner basis* calculations presented in this paper. While the simplex method is a numerical method applicable for a wide range of problems, the method based on the Gröbner basis theory is useful only for some special problems, but with the advantage of *exact, analytical results*, thus the examination of the parameter dependence also becomes possible.

Searching for the global equilibrium paths we must often cope with polynomial equation systems of several unknowns. Since we are interested in *all the solutions* of these equations, and we want to examine the dependence of the solutions on *several parameters* characterizing the

---

\* Gáspár, Zsolt, H-1025 Budapest, Kapy u. 40/b, Hungary

\*\* Károlyi, György, H-1126 Budapest, Ugozca u. 11, Hungary

\*\*\* Popper, György, H-1016 Budapest, Szirtes u. 28/a, Hungary



system, we have to find the results using symbolical computations. The well-known traditional methods [3] can not give these results, since they can not handle undetermined parameters. However, application of the theory of *Gröbner basis* can provide us with all the solutions of the polynomial equations and also allows us to carry out the computations without prescribing the parameter values in advance [4-7].

Elimination methods based on iterative exclusion of the variables from a set of polynomial equations has the property that "false roots" can occur in the final univariate equation which are not solutions of the original system. Gröbner basis calculations, based on recent results of modern algebra, assure that the transformed system and the original one are equivalent in the sense that they have *only common solutions*. Applying a special ordering of the terms of the polynomials, the so-called *pure lexicographic term ordering*, the Gröbner basis of the original set will consist of equations that can be solved one by one, since all of them contains only one new unknown. The theory of the Gröbner basis is not explained here in detail, it can be found e.g. in [7].

For finding equilibrium paths of structures described by polynomial equations the Gröbner basis calculation can be a good candidate. With its application one can obtain *analytical* results for the global equilibrium paths of structures, which enables us to examine their behaviour's parameter dependence. To demonstrate it we show the calculation of the global equilibrium paths of a model of the simple arch [8,9]. For this structure only some of the symmetric equilibrium layouts have been known so far. With the application of the Gröbner basis we show that for a range of the parameter values we can have loaded shapes with symmetry breaking. We also show that allowing the concept of negative bar length these solutions always exist and more symmetric solutions can be found. We also explain the physical meaning of the negative bar length by means of non-homogenous bars with a certain internal structure.

In Section 2 we introduce the model of the simple arch together with the mathematical equations describing its equilibrium states, and compute the corresponding Gröbner basis. Then we interpret the solutions in Section 3. Section 4 deals with the parameter dependence of the global equilibrium paths, while in Section 5 we draw our conclusions.

## 2. The model of a simple arch

Consider the structural system shown in Fig. 1 consisting of two elastic bars with equal initial length  $l$  and equal stiffness  $k$ , being pinned to each other and to rigid supports by ideal hinges. The distance between the abutments is taken as  $a$ . We want to find all the possible positions of the central hinge subjected to a vertical dead load of arbitrary magnitude  $F$ . Denote the horizontal position of the hinge by  $x$  and the vertical position by  $y$ .

Due to the reflection symmetry of the structure and of the load with respect to the vertical axis  $x = a/2$  we may expect solutions, i.e. equilibrium states of the structure, provided with the same symmetry property. However, it is not straightforward that other solutions with symmetry breaking do not exist, and even the number of symmetric solutions is not apparent.

Denoting the new lengths of the beams after loading by  $L_1$  and  $L_2$ , the geometrical equations, expressing that the bars will connect the abutments to the central hinge also after loading, are

$$x^2 + y^2 = L_1^2, \quad (1a)$$

$$(a - x)^2 + y^2 = L_2^2. \quad (1b)$$



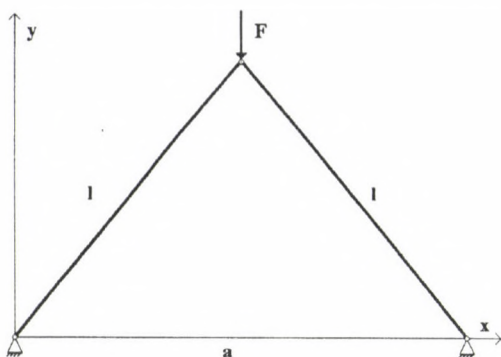


Fig. 1. Layout of the model of the simple arch before loading. The bar lengths are  $l$ , the distance between the abutments is  $a$ .

By the material equations we express that Hooke's law holds for arbitrarily large deflections. In this case it means a linear dependence between the elongation and the internal force of the bars in the following form:

$$k(L_1 - l) = S_1, \quad (2a)$$

$$k(L_2 - l) = S_2. \quad (2b)$$

Here  $S_1$  and  $S_2$  are the internal forces of the bars. The fact that the deflections are not restricted in Hooke's law will lead to very interesting results, in fact it will increase the number of solutions, as we shall see later. The equilibrium equations

$$F + \frac{y}{L_1} S_1 + \frac{y}{L_2} S_2 = 0,$$

$$\frac{x}{L_1} S_1 - \frac{a-x}{L_2} S_2 = 0$$

express the equilibria of the vertical and the horizontal force components at the central hinge, respectively. These equations can be rewritten in the form of polynomial equations as

$$FL_1L_2 + yS_1L_2 + yS_2L_1 = 0, \quad (3a)$$

$$xS_1L_2 - aS_2L_1 + xS_2L_1 = 0. \quad (3b)$$

We note that these equations are equivalent with the former ones, except for  $L_1 = 0$  or  $L_2 = 0$ . For convenience from now on we allow these values also by using Eqs. (3), although at this stage the meaning of such solutions, if there are any, is not apparent.

To prepare this system of six algebraic equations of six unknowns for solving, the Gröbner basis of the six polynomials was calculated with respect to the pure lexicographic term ordering. This provides us with six univariate equations each with only one unknown [7]. By means of the Gröbner basis theory, it is expected that the solutions obtained from the Gröbner basis are exactly the same as those of the original polynomial equation system. This enables us to look for the roots of the Gröbner basis which is a significantly easier problem.

The Gröbner basis of this polynomial equation system was calculated by the computer algebra system MAPLE V [10]. The application of the command "GSOLVE" of MAPLE V leads us to the following collection of reduced *lexicographic* Gröbner basis:

$$[2x - a, 4y^2 + a^2, S_1 + kl, S_2 + kl, L_1, L_2] \quad (4a)$$

$$\begin{aligned} & [2x - a, -4k^2L_2^3 + 2klFy + (k^2a^2 + F^2)L_2 + 4k^2lL_2^2 - k^2a^2l, S_1 - kL_2 + kl, \\ & S_2 - kL_2 + kl, -L_2 + L_1, -8k^2lL_2^3 + 4k^2L_2^4 + 2k^2a^2lL_2 + (-k^2a^2 + 4k^2l^2 - F^2)L_2^2 - k^2a^2l^2] \end{aligned} \quad (4b)$$

$$[x - a, y, S_1 - kL_1 + kl, S_2 + kl, L_1^2 - a^2, L_2] \quad (4c)$$

$$[x, y, S_1 + kl, S_2 + kl - ka, L_1, L_2 - a] \quad (4d)$$

$$[x, y, S_1 + kl, S_2 + kl + ka, L_1, L_2 + a] \quad (4e)$$

$$\begin{aligned} & [(4lk^4a^3 + 4ak^2lF^2)L_2 + (2k^4a^4 + 4F^2k^2a^2 + 2F^4)x - k^4a^5 - 2F^2k^2a^3 - 4a^3l^2k^4 - aF^4, \\ & (2k^5a^4 + 4k^3a^2F^2 + 2kF^4)y - k^4a^4F + 4k^4a^2F^2 - 2F^3k^2a^2 - F^5, \\ & (k^2a^2 + F^2)S_1 + (k^3a^2 + kF^2)L_2 - k^3a^2l + klF^2, S_2 - kL_2 + kl, \\ & (k^2a^2 + F^2)L_1 + (k^2a^2 + F^2)L_2 - 2k^2a^2l, \\ & -8a^2lk^4L_2 + (4k^4a^2 + 4k^2F^2)L_2^2 - k^4a^4 + 4a^2l^2k^4 - 2F^2k^2a^2 - F^4] \end{aligned} \quad (4f)$$

Here six subsystems between brackets are given. Each of them represents a system of six equations, in which the given polynomials are supposed to be equal to 0.

The first subsystem gives imaginary result for  $y$ , thus it is physically irrelevant. Subsystems (4c), (4d) and (4e) are included in (4f) for special  $F$  values. Thus we must examine only subsystems (4b) and (4f) to find the global equilibrium paths of the model of Fig. 1.

The remaining two subsystems can be solved more easily than the original equation system. This is the consequence of the application of the *pure lexicographic term ordering* in the calculation of the Gröbner basis. In the last equation of the subsystems there is  $L_2$  as the only unknown, thus it can be easily determined. Since these equations are of degree 4 and 2 in the unknown  $L_2$  in (4b) and (4f), respectively, explicit algebraic solutions can be obtained. This provides us with the opportunity of the examination of the parameter dependence of the global equilibrium paths. After determining  $L_2$  from the last equation of the subsystems, and substituting it into the preceding equation we again find a univariate equation for the only unknown  $L_1$ . Carrying on this procedure finally we can obtain the solution for all the unknowns  $x, y, S_1, S_2, L_1, L_2$  as explicit functions of the parameters  $F, a, k, l$ . The formulas determined this way are collected in the Appendix.

### 3. Global equilibrium paths of the simple arch

The new shape of the model of the simple arch after loading may be described by the coordinates  $(x, y)$  of the central hinge. From the Gröbner basis we have calculated in the former section, one could determine them as explicit functions of the parameters (see Appendix). For a detailed examination of the equilibrium paths we set these parameter values as  $l = 5$ ,  $a = 6$  and  $k = 1$  in arbitrary units. Thus we can follow the changing of the position  $(x, y)$  of the central

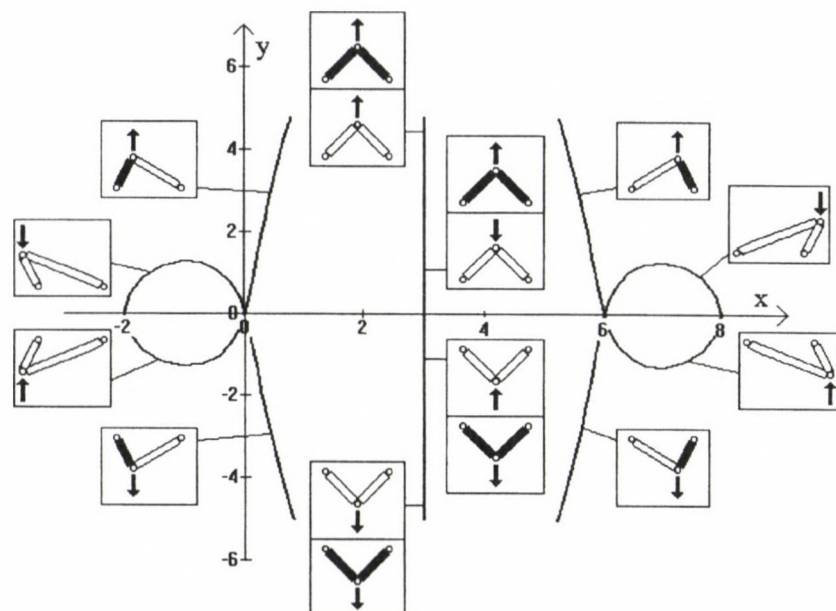


Fig. 2a. Global equilibrium paths of the simple arch for  $l = 5$ ,  $a = 6$  and  $k = 1$ , together with the corresponding layouts. Bars with negative length are painted black.

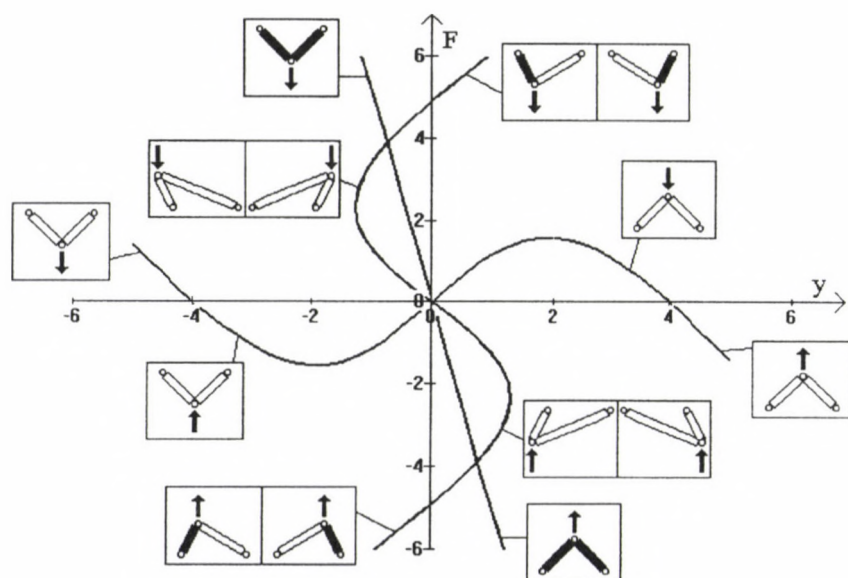


Fig. 2b. Dependence of the vertical coordinate of the central hinge on the load parameter for  $l = 5$ ,  $a = 6$  and  $k = 1$ , together with the corresponding layouts. Bars with negative length are painted black.



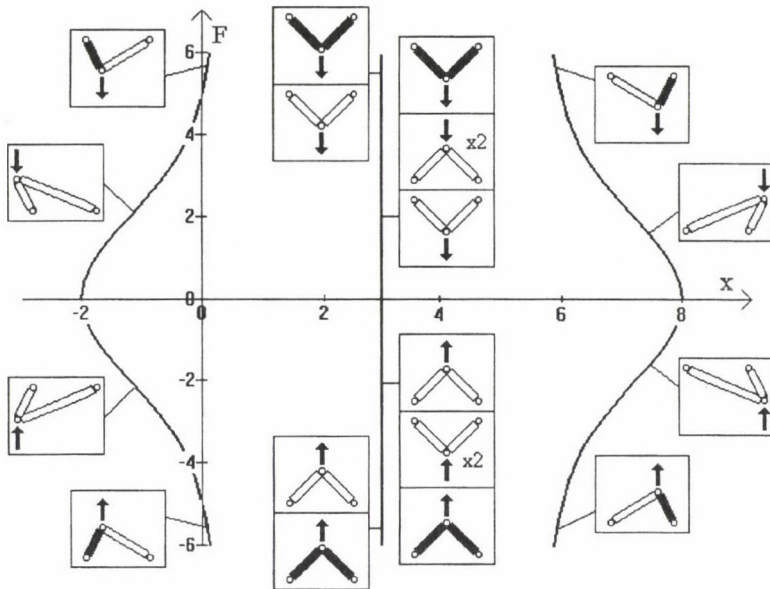


Fig. 2c. Dependence of the horizontal coordinate of the central hinge on the load parameter for  $l = 5$ ,  $a = 6$  and  $k = 1$ , together with the corresponding layouts. Bars with negative length are painted black.

hinge by varying the value of the load  $F$ . As the load varies between minus and plus infinity, the central hinge moves along one of its equilibrium paths. This is illustrated in Fig. 2a. Here the possible equilibrium positions of the central hinge are shown as a function of  $F$  together with the corresponding structural layouts. It is obvious that there are solutions breaking the reflection symmetry of the original structure with respect to the  $x = a/2$  line.

In fact the solutions shown in Fig. 2a are projections of the six solutions of the Gröbner basis to the  $(x, y)$  plane; recall that the two equations for  $L_2$  were of order two and four, thus they can have six different real solutions. The equations do have six real solutions for certain values of  $F$ , and have four real solutions for other load values. This is illustrated in Figs. 2b and c, where the equilibrium coordinates of the central hinge are shown as a function of the load  $F$ .

Bars coloured black in the small layout figures indicate a special property of the corresponding solutions obtained from the Gröbner basis: for some load values these bars have large enough compression to decrease their length below zero:  $L_i \leq 0$  (see Fig. 3). This property is included in the original equations describing the model, this is the result of the fact that the deflections were not restricted. The question arises whether these solutions may have physical meaning. In fact they can have, if the bars are not homogeneous, but have a certain internal structure illustrated in Fig. 4a. Here the supporting fixed hinge is connected to a tube somewhere along it, while an elastic spring connects the central hinge via the open end of the tube to its other end. Thus, applying large compressing load to the "end of the bar", i.e. to the free end of the spring, it can move inside the tube further than the supporting hinge as is illustrated in Fig. 4b. If we define the new length of the bar as the original length  $l$  plus the deflection  $\Delta l < -l$ , we clearly get a bar with negative length in accordance with the equations (1-3) describing the equilibrium state of the system.

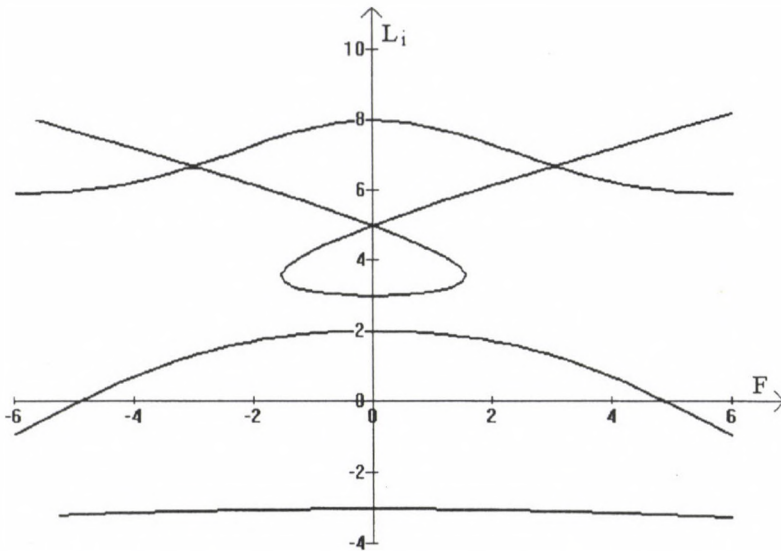


Fig. 3. Dependence of the bar lengths on the load parameter for  $l = 5$ ,  $a = 6$  and  $k = 1$ . Obviously for some parts of the global equilibrium paths the length of the bars may be negative.

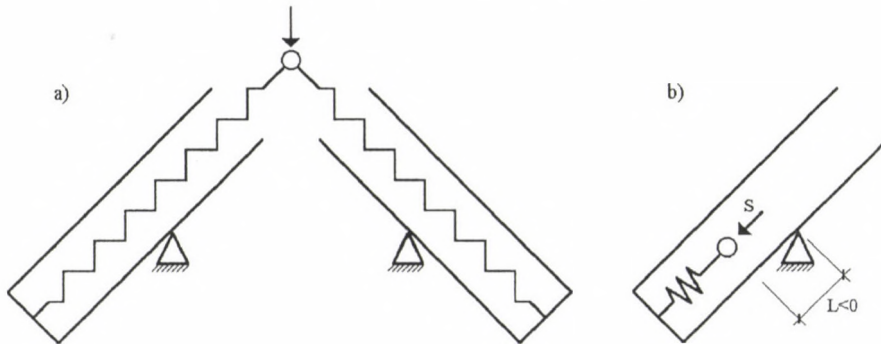


Fig. 4. Possible layout of the internal structure of a bar, which can have "negative length". a) The initial layout of the structure set together from such bars. b) The length of the bar with substantial compression may decrease below zero.

#### 4. Parameter dependence of the global equilibrium paths

Recall that executing the solution procedure of Sec. 2 finally we were provided with all the unknowns as explicit functions of the parameters. Changing the value of the parameter  $a$  denoting the distance between the supports, we only set the length scale of the structure. Choosing different values for the stiffness  $k$  only the unit of force is adjusted. Thus any change of  $a$  and  $k$  can be turned into a similarity transformation only by the appropriate choice of the values for  $l$  and  $F$ . Thus the system is uniquely described by the two parameters  $l$  and  $F$ .

Varying  $F$  for fixed value of  $l$  we scan the global equilibrium paths of the structure the way as we did it in Sec. 3. This section is devoted to the changing of the equilibrium states as  $l$  varies.

For  $l \geq a/2$  the arch can be in equilibria in its original position for  $F = 0$  without any internal force. For  $l < a/2$  it is not possible, since without stretching the bars starting from the supports they can not reach each other. At the crossover value  $l = a/2$  the global equilibrium paths of the simple arch also change dramatically. This value is the borderline between two different states, where the number of the solutions is different. As we have already seen in Sec. 3 for  $l \geq a/2$  we could find six or four different real solutions, depending on the value of  $F$ . However, for  $l < a/2$  we definitely have only four different real solutions for any value of the load  $F$ . This is illustrated in Fig. 5 for the parameter values  $l = 2.5$ ,  $a = 6$  and  $k = 1$ . Comparing Figs. 2b and 5b it can be understood that those curves of Fig. 2b which have intersection with the axis outside the origin will be tangent to the same axis at the origin at the crossover value  $l = a/2$ .

The four different lines of Fig. 5 illustrate all the possible positions of the central hinge in equilibrium state. Three of the lines now correspond to layouts where at least one of the bars has negative length. Note, that the loop of Fig. 2a according to the non-symmetric shapes now disappeared; for all the non-symmetric solutions now one of the bars must be of negative length after loading. The border between the existence and non-existence of the loop at the non-symmetric solutions is again  $l = a/2$  where the zero length is approached for  $F \rightarrow 0$ , for other  $F$  values one of the bars must be of negative length.

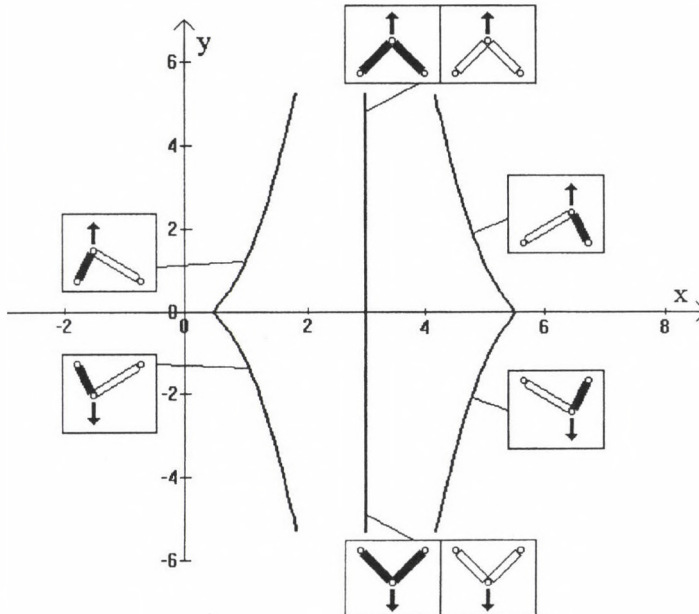


Fig. 5a. Global equilibrium paths of the simple arch for  $l = 2.5$ ,  $a = 6$  and  $k = 1$ , together with the corresponding layouts. Bars with negative length are painted black.



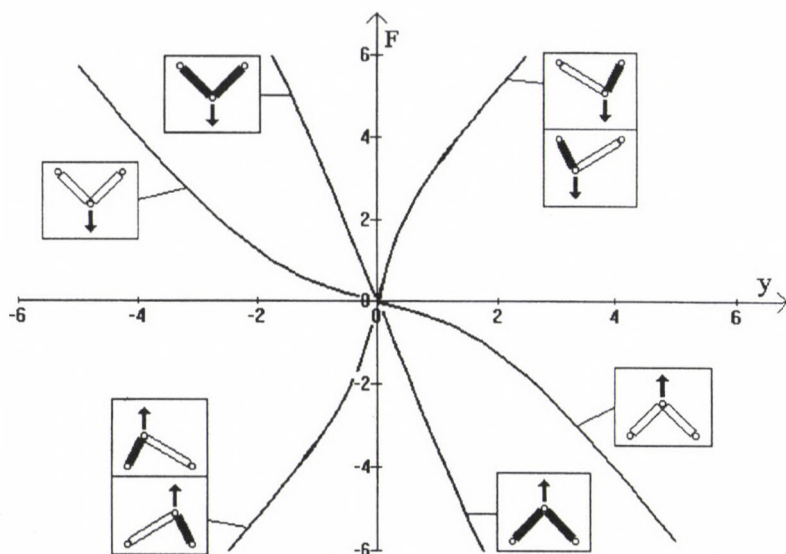


Fig. 5b. Dependence of the vertical coordinate of the central hinge on the load parameter for  $l = 2.5$ ,  $a = 6$  and  $k = 1$ , together with the corresponding layouts. Bars with negative length are painted black.

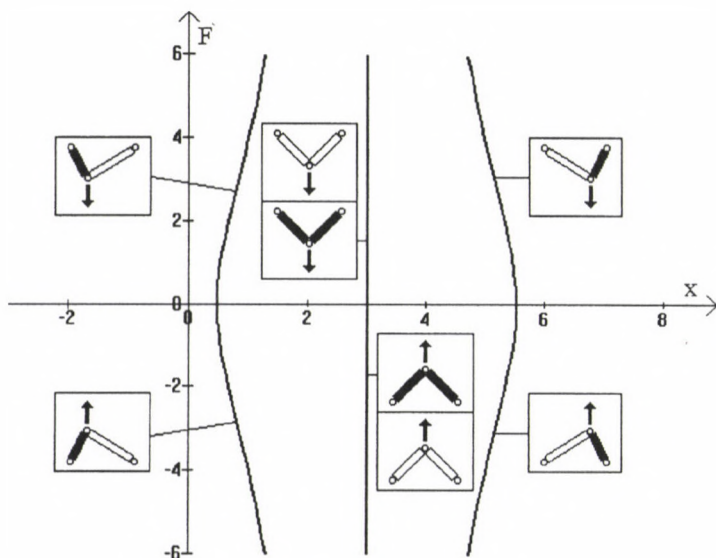


Fig. 5c. Dependence of the horizontal coordinate of the central hinge on the load parameter for  $l = 2.5$ ,  $a = 6$  and  $k = 1$ , together with the corresponding layouts. Bars with negative length are painted black.

## 5. Conclusions

Application of Gröbner basis theory for determining global equilibrium paths was presented through a simple example. Due to the analytical results provided by the Gröbner basis calculations considerable insight was possible into the behaviour of the model of the simple arch. Even for this very simple example quite unexpected results were found such as equilibrium states with symmetry breaking property or with negative bar length. The explicit knowledge of the analytical solution thus could lead to deeper understanding of the behaviour of this kind of bar structures.

However, it must be noted, that the results of the Gröbner basis calculations often lead to equations with degree higher than four. These equations can not be solved analytically, thus we lose one great advantage of this solution technique. But even in such cases the numerical solution is considerably easier due to the fact that the equations to be solved are univariate, and a drawback of other elimination methods is excluded, namely that we do not get any false solutions, i.e. the roots of the original polynomial system and of the Gröbner basis are the same.

General theorems guarantee that the Gröbner basis always exist for any polynomial equation system. However, it does not mean that it is easy to calculate, since the necessary computer time increases exponentially with the number of the variables [4-7].

In spite of these difficulties, the application of the Gröbner basis calculations can be suggested for examples where the number of variables is not too high, and analytical results are required.

## Acknowledgements

This work was supported by the Hungarian Scientific Research Foundation under Grants No. OTKA T014287 and T015860.

## Appendix: Solutions of the univariate equations of the Gröbner basis

The sixth of Eqs. (4f) is of second degree for the unknown  $L_2$ . Solving it we get two solutions:

$$L_2 = \frac{2k^3a^2l \pm \sqrt{k^6a^6 + F^6 + 3k^2a^2F^4 + 3k^4a^4F^2 - 4k^4a^2F^2l^2}}{2k^3a^2 + 2kF^2}$$

These two  $L_2$  values are quite complicated, but explicit functions of the parameters. Substituting this expression into the fifth of Eqs. (4f) we can obtain the corresponding  $L_1$  values as

$$L_1 = \frac{2k^3a^2l \mp \sqrt{k^6a^6 + F^6 + 3k^2a^2F^4 + 3k^4a^4F^2 - 4k^4a^2F^2l^2}}{2k^3a^2 + 2kF^2}$$

From the third and fourth of Eqs. (4f) the corresponding  $S_1$  and  $S_2$  values may be derived, respectively, while from the second equation the value of  $y$  for both  $L_2$  values is

$$y = \frac{F^5 + 2F^3k^2a^2 - 4F^2k^4a^2 + Fk^4a^4}{2k^5a^4 + 4k^3a^2F^2 + 2kF^4}$$

Finally, from the first of Eqs. (4f)  $x$  can be expressed as

$$x = \frac{k^4a^5 + 2F^2k^2a^3 + aF^4 \mp 2alk \sqrt{k^6a^6 + F^6 + 3k^2a^2F^4 + 3k^4a^4F^2 - 4k^4a^2F^2l^2}}{2k^4a^4 + 4F^2k^2a^2 + 2F^4}$$

These solutions correspond to the equilibrium states of the system with symmetry breaking.

For the Eqs. (4b) the same procedure could be followed. However, it is more advantageous to express the variables  $x, L_1, L_2, S_1, S_2$  and  $F$  with  $y$  as the main parameter. The reason for this is that we always have two  $F$  values for any  $y$  value, but we may get two or four (real)  $y$  values for some  $F$ , and the latter is quite difficult to handle during plotting the graphs of the resulting functions.

The results obtained from Eqs. (4b) in this manner are as follows:

$$x = a/2,$$

$$F = -2ky \pm \frac{4kly}{\sqrt{a^2 + 4y^2}}$$

$$L_1 = L_2 = \pm \frac{1}{2} \sqrt{a^2 + 4y^2}$$

From the first of these equations it is obvious that these are the symmetric solutions.

#### REFERENCES

1. Domokos, G.: Global description of elastic bars. ZAMM 74 (1994) T289-T291
2. Domokos, G.-Gáspár, Zs.: A global, direct algorithm for path-following and active static control of elastic bar structures. Mech. Struct. & Mach. 23 (1995) 549-571
3. Moses, J.: Solution of a System of Polynomial Equations by Elimination. Comm. ACM 9 (1966) 634-637
4. Geddes, K. O.-Czapor, S. R.-Labahn, G.: Algorithms for Computer Algebra. Kluwer Academic Publishers, 1992
5. Davenport, J. H.-Siret, Y.-Tournier, E.: Systems and algorithms for algebraic computation. Computer Algebra. Academic Press, 1993
6. Becker, T.-Weispfenning, V.: Gröbner Bases: a Computational Approach to Commutative Algebra. Springer-Verlag, 1993
7. Popper, Gy.: Solving systems of polynomial equations using Gröbner basis calculations with applications to mechanics. Accepted for publication by Comp. Ass. Mech. Eng. Sci., 1996
8. Thompson, J. M. T.-Hunt, G. W.: A General Theory of Elastic Stability. John Wiley & Sons, 1973
9. Thompson, J. M. T.-Hunt, G. W.: Elastic instability phenomena. Wiley, Chichester, 1984
10. Char, B. W.-Geddes, K. O.-Gonnet, G. H.-Leong, B. L.-Monagan, M. B.-Watt, S. M.: Maple V, Library Reference Manual. Springer-Verlag, 1991





## USE OF FINITE ELEMENT METHOD FOR TESTING STRUCTURES FOR DYNAMICAL EFFECTS CAUSED BY MOVING VEHICLES

GYÖRGYI, J.\*

(Received: 11 July 1996)

In tests of bridge behaviour under influence of moving load, a quasi-harmonic vibration task should be solved if we want to take into consideration also the mass of the moving body. Use of finite element method requires analysis of systems of several degrees of freedom and this calls for appropriate numerical techniques. Both the effect of the external damping proportional to the velocity should be computed and the procedure must be able to consider also the structural damping. The method to be outlined in this paper allows the analysis of simultaneous effect of both the mass and stiffness matrix changing with time, and the external damping proportional to velocity and frequency-independent structural damping. The computation algorithm refers to a general case, however, the effect of moving mass regarded as a material point, the two-axle road vehicle and multiaxle railway vehicle is separately discussed.

### 1. Introduction

Among engineering vibration tasks, test of bridges under the effect of a moving load is an old problem. If we want to take into consideration also the mass of the moving body during tests, a quasi-harmonic vibration task should be solved. Before the appearance of computers, this problem could be analysed only by essential simplifications. One of the possible simplifications is neglect of the structure's mass. In another case a harmonic vibration task will be solved neglecting the mass of the moving body. The starting point for this test was the differential equation of the vibrating bar, and the solution of this equation was sought. The structural model could not be changed until the appearance of computers with sufficient capacity, however, already Frýba [1] analysed the simultaneous effect of load and mass. In these tests, the bridge was simulated by a beam of constant stiffness and tested as continuum. Later, as an advance, the vehicle was simulated not only as a moving mass but as a dynamical system of several degrees of freedom - like Green and Cebon [2] did it for road vehicles.

Application of the finite element method requires test of dynamical systems of several degrees of freedom and this calls for appropriate numerical methods. For computing matrix differential equations of constant coefficient for given starting conditions, several methods have been developed, among others the Wilson- $\Theta$  method described by Bathe and Wilson [3]. In their original forms these methods are - as to be illustrated below - for matrix differential equations with time-dependent coefficients rather computation-intensive, and fit only for solving very simple tasks. Inbanathan and Wieland [4] modelled the vehicle as a single moving mass, and the effect of damping was neglected both for structure and vehicle. However, it is well known that the dynamical overload is greatly influenced by damping, too. The known methods - reported e.g. by Olsson [5] - are able to compute the effect of the external damping

---

\*Györgyi, József, H-1221 Budapest, Arany János út 96/b, Hungary

proportional to the velocity, however, are not fit to take the structural damping into consideration. Györgyi [6] suggested a method for computing the structural damping as an external damping for which, as preparation, the eigenvalue problem belonging to the undamped dynamical system of several degrees of freedom must be solved. It is necessary because - as proved by Györgyi [7] - the appropriate size of the integration step can be determined only this way. All this makes the computations rather sophisticated and unfit for testing large-size systems. The problems mentioned necessitated the preparation of the method to be described below. The technique published in this paper allows the simultaneous analysis of both the mass and stiffness matrix changing with time and the external damping proportional to velocity and the frequency-independent structural damping.

## 2. Mechanical models of the task

### 2.1 Effect of a moving mass regarded as a material point

The simplest case of the task under test can be seen in Fig. 1 where the vehicle is modelled by a material point. Furthermore, in tests we assume that when moving, the vehicle is continuously supported by the system, thus, the mass point does not separate from the structure.

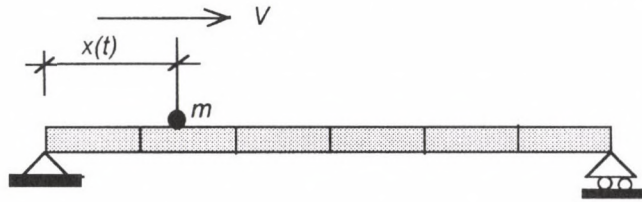


Fig. 1. Mechanical model of a moving material point

In this case, the mass matrix of the structure will be extended according to the vehicle mass. Considering that the mass point moves, the supplementary mass will act on different elements in time. If the mass point is within one element, the relevant mass forces should be reduced to the neighbouring nodes. If we comprise in matrix  $N_t$  the values belonging to the given state of the mass point of the displacement functions describing the relation between the displacement of the element's internal points and the displacement of the element's different nodes, we can write the time-dependent supplementary elementary mass matrix in the following form:

$$\mathbf{M}_{et} = \mathbf{N}_t^T \mathbf{M}_0 \mathbf{N}_t .$$

Here  $\mathbf{M}_0$  is the diagonal matrix in a size corresponding to the number of components of the internal point displacement, the value of elements belonging to displacements is  $m$ , while the others are zero. It often occurs when the mass moves on a horizontal surface in the direction of the axis of the bridge to be modelled as a straight beam. In this case, we can analyse a planar task. If the test aims at analysing bending in vertical plane, only one displacement component in each node should be reckoned with, thus, only one element of  $\mathbf{M}_0$  matrix will be other than zero.



The weight force corresponding to the moving mass will appear at those nodes of the element where the mass point is. This is the elementary vector of force:

$$\mathbf{r}_{ef} = \mathbf{N}_t^T \mathbf{r}_0.$$

In vector  $\mathbf{r}_0$ , components of displacement's direction of the weight force corresponding to the given mass  $m$  are not zero. For a mass moving on a horizontal surface, the force will be perpendicular to the structure. Then, the differential equation of the task shown in Fig. 1 will be as follows:

$$(\mathbf{M}_B + \mathbf{M}_1(t))\ddot{\mathbf{u}} + \mathbf{K}_B \mathbf{u} = \mathbf{r}_1(t). \quad (1)$$

In this expression  $\mathbf{M}_B$  is the mass matrix of the structure,  $\mathbf{K}_B$  is the stiffness matrix of the structure and  $\mathbf{x}$  is the vector of node displacements. The moving mass did not change the task's sizes, its effect is expressed by mass matrix  $\mathbf{M}_1(t)$  and vector of force  $\mathbf{r}_1(t)$ . Only those elements of them differ from zero in which the components of matrix  $\mathbf{M}_{ef}$  and  $\mathbf{r}_{ef}$  are displaced depending on the mass state. We can see that the mass matrix consists of two parts. If we introduce in the complete dynamical test the notation without index and the following expression:

$$\mathbf{M}(t) = \mathbf{M}_B + \mathbf{M}_1(t)$$

for the matrices and vectors of load, equation (1) appears in the following form:

$$\mathbf{M}(t)\ddot{\mathbf{u}} + \mathbf{K}\mathbf{u} = \mathbf{r}(t). \quad (2)$$

## 2.2 Dynamical equation for a two-axle road vehicle

The vehicle model will be arranged after Green and Cebon [2] as depicted in Fig. 2. The vehicle is supported by the structure in points **a** and **b**, the vertical displacements of these points are identical with vertical displacements of a given point of the structure.

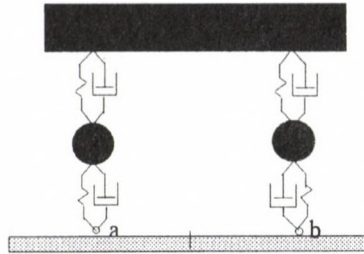


Fig. 2. Mechanical model of a road vehicle

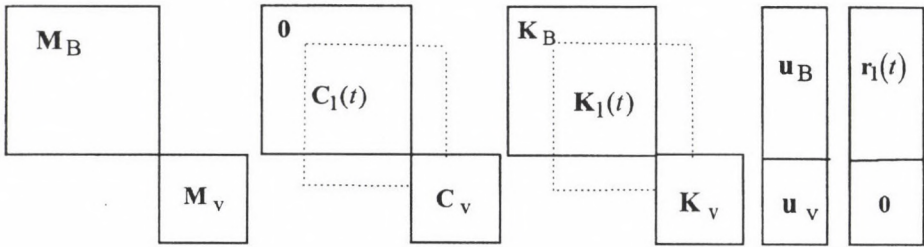
We assume that points **a** and **b** do not displace and write the dynamical equations of the vehicle vibrating around the equilibrium state.

$$\mathbf{M}_v \ddot{\mathbf{u}}_v + \mathbf{C}_v \dot{\mathbf{u}}_v + \mathbf{K}_v \mathbf{u}_v = \mathbf{0}. \quad (3)$$

Here  $\mathbf{M}_v$  is the mass matrix of the vehicle,  $\mathbf{C}_v$  the damping matrix corresponding to hydraulic dampers, and  $\mathbf{K}_v$  the stiffness matrix of the vehicle. In this case, the displacement vector  $\mathbf{u}_v$  includes only the displacement components of the mass points, the rows and columns corresponding to supports can be deleted from the matrices. We obtain the following dynamical equations for the structure:

$$\mathbf{M}_B \ddot{\mathbf{u}}_B + \mathbf{K}_B \mathbf{u}_B = \mathbf{r}_1(t). \quad (4)$$

In the matrix equation the whole weight of the vehicle in the nodes of elements containing points **a** and **b** will be given by the help of the elementary force vectors  $\mathbf{r}_{et} = \mathbf{N}_t^T \mathbf{r}_0$  as load in vector  $\mathbf{r}_1$ . When arranging the complete dynamical system, the displacements of the structure come in the first block of the displacement vector, and in the second one the displacement vector of the vehicle. Analysing the whole system, points **a** and **b** in the vehicle model will displace, and this displacement will be identical to the displacement of the given point of the finite element corresponding to the vehicle state. Accordingly, both matrix differential equations placed next to each other will join and the matrices and vectors will have the following structure:



In matrices  $\mathbf{C}_1(t)$  and  $\mathbf{K}_1(t)$  place of elements will change with time according to the vehicle's movement. These matrices can be obtained by compilation from the elementary matrices  $\mathbf{C}_{et}^a$ ,  $\mathbf{C}_{et}^b$  and  $\mathbf{K}_{et}^a$ ,  $\mathbf{K}_{et}^b$ , respectively, belonging to the individual nodes. For computation of the elementary matrices, matrix  $\mathbf{N}_t$  already described above can be used, its elements are, of course, different in points **a** and **b**. For point **a**, e.g.:

$$\mathbf{K}_{et}^a = k^a \begin{array}{|c|c|} \hline \mathbf{N}_t^T \mathbf{N}_t & -\mathbf{N}_t^T \\ \hline -\mathbf{N}_t & \\ \hline \end{array} \quad \mathbf{C}_{et}^a = c^a \begin{array}{|c|c|} \hline \mathbf{N}_t^T \mathbf{N}_t & -\mathbf{N}_t^T \\ \hline -\mathbf{N}_t & \\ \hline \end{array}$$

Here  $k^a$  and  $c^a$  mean the stiffness of the spring joining in point  $a$ , and damping coefficient of the damper, respectively. Thus, the matrix differential equation of the task demonstrated in Fig. 2. is as follows:

$$\mathbf{M}\ddot{\mathbf{u}} + (\mathbf{C}_C + \mathbf{C}_1(t))\dot{\mathbf{u}} + (\mathbf{K}_C + \mathbf{K}_1(t))\mathbf{u} = \mathbf{r}_1(t). \tag{5}$$

In this relationship, subscript C designates the matrices with constant elements. We note that in this model the mass matrix does not change with time. Writing the equation by the help of summed up matrices, we obtain

$$\mathbf{M}\ddot{\mathbf{u}} + \mathbf{C}(t)\dot{\mathbf{u}} + \mathbf{K}(t)\mathbf{u} = \mathbf{r}(t). \tag{6}$$

2.3 Analysis of the effect of multi-axle railway vehicles

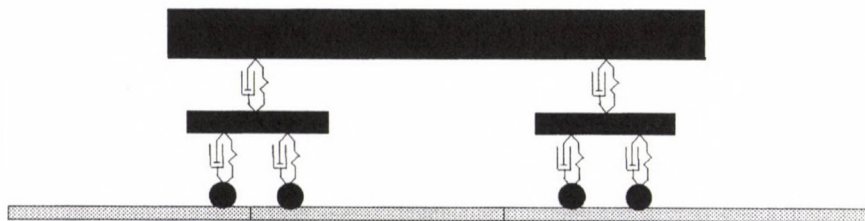
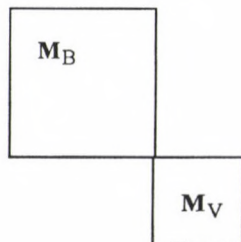


Fig. 3. Mechanical model of the case of a railway vehicle

In this task, there are mass points of the vehicle also in contact points of the structure. If we want to apply again the procedure described before, in the matrix equation of the vehicle these masses do not appear in the case of unmoving contact points. Nevertheless, as shown in point 1.1, the effect of these mass points will be expressed in the mass matrix of the structure as a time-dependent component. Consequently, the dynamical equation system of the coupled system is the following:

$$(\mathbf{M}_C + \mathbf{M}_1(t))\ddot{\mathbf{u}} + (\mathbf{C}_C + \mathbf{C}_1(t))\dot{\mathbf{u}} + (\mathbf{K}_C + \mathbf{K}_1(t))\mathbf{u} = \mathbf{r}_1(t). \tag{7}$$

$\mathbf{M}_C$  is the time-independent mass matrix shown in 2.2:





$\mathbf{M}_1(t)$  and  $\mathbf{r}_1(t)$  are matrix and load vectors of equation (1), respectively, containing, of course, several components  $\mathbf{M}_{e_t}$  and  $\mathbf{r}_{e_t}$  because there are several mass points. If we write the equation by summed up matrices as shown above, we obtain

$$\mathbf{M}(t)\ddot{\mathbf{u}} + \mathbf{C}(t)\dot{\mathbf{u}} + \mathbf{K}(t)\mathbf{u} = \mathbf{r}(t). \quad (8)$$

The last one out of the three models introduced is the general model. In this case, all the interacting matrices and the load vector are time-dependent.

### 3. Solution of the matrix differential equation by direct integration

The dynamical analysis aims at solution of the matrix differential equation under initial conditions  $\dot{\mathbf{u}}_0, \mathbf{u}_0$  belonging to a given time  $t_0$ , and at computation of dynamical stresses knowing displacements. Inbanathan and Wieland [4] and Olsson [5] successfully applied the Newmark procedure for computing displacements in simpler tasks, while Györgyi [8] showed that it was expedient to use the Wilson- $\theta$  procedure for solution. Wilson [3] assumes linear acceleration change between moments  $t$  and  $t + \theta \Delta t$  - the procedure's convergence can be assured if  $\theta=1.4$ . In this case:

$$\ddot{\mathbf{u}}_{t+\tau} = \ddot{\mathbf{u}}_t + \frac{\tau}{\theta \Delta t} (\ddot{\mathbf{u}}_{t+\theta \Delta t} - \ddot{\mathbf{u}}_t), \quad (9)$$

$$\dot{\mathbf{u}}_{t+\tau} = \dot{\mathbf{u}}_t + \ddot{\mathbf{u}}_t \tau + \frac{1}{2} \frac{\tau^2}{\theta \Delta t} (\ddot{\mathbf{u}}_{t+\theta \Delta t} - \ddot{\mathbf{u}}_t), \quad (10)$$

$$\mathbf{u}_{t+\tau} = \mathbf{u}_t + \dot{\mathbf{u}}_t \tau + \frac{1}{2} \ddot{\mathbf{u}}_t \tau^2 + \frac{1}{6} \frac{\tau^3}{\theta \Delta t} (\ddot{\mathbf{u}}_{t+\theta \Delta t} - \ddot{\mathbf{u}}_t). \quad (11)$$

From the above we can deduce

$$\ddot{\mathbf{u}}_{t+\theta \Delta t} = \frac{6}{(\theta \Delta t)^2} (\mathbf{u}_{t+\theta \Delta t} - \mathbf{u}_t) - \frac{6}{\theta \Delta t} \dot{\mathbf{u}}_t - 2\ddot{\mathbf{u}}_t, \quad (12)$$

$$\dot{\mathbf{u}}_{t+\theta \Delta t} = \frac{3}{\theta \Delta t} (\mathbf{u}_{t+\theta \Delta t} - \mathbf{u}_t) - 2\dot{\mathbf{u}}_t - \frac{\theta \Delta t}{2} \ddot{\mathbf{u}}_t. \quad (13)$$

Assuming linear change of  $\mathbf{r}(t)$  for the given time period, we get

$$\mathbf{r}_{t+\theta \Delta t} = \mathbf{r}_t + \theta (\mathbf{r}_{t+\Delta t} - \mathbf{r}_t)$$

and at the moment  $t + \theta \Delta t$  displacements can be obtained from the following relationship:

$$\left( \mathbf{K} + \frac{6}{(\theta \Delta t)^2} \mathbf{M} + \frac{3}{\theta \Delta t} \mathbf{C} \right) \mathbf{u}_{t+\theta \Delta t} = \mathbf{r}_{t+\theta \Delta t} + \mathbf{M} \left( \frac{6}{(\theta \Delta t)^2} \mathbf{u}_t + \frac{6}{\theta \Delta t} \dot{\mathbf{u}}_t + 2\ddot{\mathbf{u}}_t \right) + \mathbf{C} \left( \frac{3}{\theta \Delta t} \mathbf{u}_t + 2\dot{\mathbf{u}}_t + \frac{\theta \Delta t}{2} \ddot{\mathbf{u}}_t \right). \quad (14)$$

Afterwards, displacements, velocities and accelerations belonging to the moment  $t + \Delta t$  can be calculated.

For constant  $\mathbf{M}$ ,  $\mathbf{C}$ ,  $\mathbf{K}$ , the coefficient matrix

$$\mathbf{A} = \left( \mathbf{K} + \frac{6}{(\theta\Delta t)^2} \mathbf{M} + \frac{3}{\theta\Delta t} \mathbf{C} \right)$$

should be decomposed only once after having taken the time step  $\Delta t$ , while the right-side vector is time-dependent.

We have seen that in the tasks analysed, at least one matrix was time-dependent, thus, matrix  $\mathbf{A}$  will be time-dependent, and the time-consuming decomposition should be carried out at every time step. Inbanathan and Wieland [4] (who tested the model shown in 2.1 for a simply supported structure and one single mass point) took this way. If we want prevent time-dependence of matrix  $\mathbf{A}$ , time-dependent components in equation (7) should be taken to the right-hand side of the equation. Further on, neglecting subscript  $\mathbf{C}$  in the constant element matrices, we can write the dynamical task in the following form

$$\mathbf{M}\ddot{\mathbf{u}} + \mathbf{C}\dot{\mathbf{u}} + \mathbf{K}\mathbf{u} = \ddot{\mathbf{r}},$$

where

$$\ddot{\mathbf{r}} = \mathbf{r} - \mathbf{M}_1\ddot{\mathbf{u}} - \mathbf{C}_1\dot{\mathbf{u}} - \mathbf{K}_1\mathbf{u}.$$

In this case, displacements can be obtained from the relation

$$\begin{aligned} \left( \mathbf{K} + \frac{6}{(\theta\Delta t)^2} \mathbf{M} + \frac{3}{\theta\Delta t} \mathbf{C} \right) \mathbf{u}_{t+\theta\Delta t} &= \ddot{\mathbf{r}}_{t+\theta\Delta t} + \mathbf{M} \left( \frac{6}{(\theta\Delta t)^2} \mathbf{u}_t + \frac{6}{\theta\Delta t} \dot{\mathbf{u}}_t + 2\mathbf{u}_t \right) + \\ &+ \mathbf{C} \left( \frac{3}{\theta\Delta t} \mathbf{u}_t + 2\dot{\mathbf{u}}_t + \frac{\theta\Delta t}{2} \ddot{\mathbf{u}}_t \right), \end{aligned} \quad (15)$$

where

$$\begin{aligned} \ddot{\mathbf{r}}_{t+\theta\Delta t} &= \mathbf{r}_{t+\theta\Delta t} - \mathbf{M}_{1t+\theta\Delta t} \left( \frac{6}{(\theta\Delta t)^2} (\mathbf{u}_{t+\theta\Delta t} - \mathbf{u}_t) - \frac{6}{\theta\Delta t} \dot{\mathbf{u}}_t - 2\ddot{\mathbf{u}}_t \right) - \\ &- \mathbf{C}_{1t+\theta\Delta t} \left( \frac{3}{\theta\Delta t} (\mathbf{u}_{t+\theta\Delta t} - \mathbf{u}_t) - 2\dot{\mathbf{u}}_t - \frac{\theta\Delta t}{2} \ddot{\mathbf{u}}_t \right) - \mathbf{K}_{1t+\theta\Delta t} \mathbf{u}_{t+\theta\Delta t}. \end{aligned} \quad (16)$$

It can be seen that on the right-hand side of the equation, vector  $\mathbf{u}_{t+\theta\Delta t}$  appears and its computation calls for an iteration procedure. The solution is now faster than decomposing matrix  $\mathbf{A}$  in every time step, however, Györgyi [8] experienced that convergence of the method could only be assured if the supplementary mass was not too large as compared to the structure's mass.

#### 4. Use of the modal analysis

##### 4.1 Effect of a moving mass regarded as material point and that of internal damping

Györgyi [8] showed that the differential equation of motion was proportional to internal damping (when the damping constant of all the structural elements is the same):

$$\mathbf{M}\ddot{\mathbf{u}} + \left( \nu \mathbf{M}\mathbf{V} \left\langle \frac{1}{\omega_{ru}} \right\rangle \mathbf{V}^T \mathbf{K} \right) \dot{\mathbf{u}} + u \mathbf{K}\mathbf{u} = \mathbf{r} - \mathbf{M}_1 \ddot{\mathbf{u}}. \quad (17)$$

Here

$$\nu = \frac{4\gamma}{4 + \gamma^2}; \quad u = \frac{4 - \gamma^2}{4 + \gamma^2}; \quad \omega_{ru} = \frac{\omega_r}{\sqrt{1 + \frac{\gamma^2}{4}}}; \quad \gamma = \frac{\varrho}{\pi};$$

In this relation,  $\varrho$  is the logarithmic decrement of damping,  $\omega_r$  can be computed from the  $r$ -th eigenvalue of the eigenvalue problem  $\mathbf{K}\mathbf{v} = \omega^2 \mathbf{M}\mathbf{v}$  belonging to the undamped case, while  $\mathbf{V}$  is the matrix containing  $\mathbf{M}$ -normed eigenvectors. Apparently, in the case of internal damping the direct integration must be preceded by the solution of an eigenvalue problem. Knowing the structure and the loading force, on the basis of Györgyi's paper [7] the number of eigenvectors needed for appropriately correct solution and the integration steps can be determined. All this supplies grounds for considering it when selecting the solution method for the dynamical task, and for trying to apply the modal analysis. In knowledge of the  $\mathbf{M}$ -normed eigenvectors of the eigenvalue problem  $\mathbf{K}\mathbf{v} = \omega^2 \mathbf{M}\mathbf{v}$ , the solution can be sought in the form of  $\mathbf{u} = \mathbf{V}\mathbf{x}$ . After substitution and multiplying by transported matrix  $\mathbf{V}^T$  from the left, we obtain

$$\mathbf{V}^T \mathbf{M}\mathbf{V}\ddot{\mathbf{x}} + \nu \mathbf{V}^T \mathbf{M}\mathbf{V} \left\langle \frac{1}{\omega_{ru}} \right\rangle \mathbf{V}^T \mathbf{K}\mathbf{V}\dot{\mathbf{x}} + u \mathbf{V}^T \mathbf{K}\mathbf{V}\mathbf{x} = \mathbf{q}, \quad (18)$$

where

$$\mathbf{q} = \mathbf{V}^T \mathbf{r} - \mathbf{V}^T \mathbf{M}_1 \mathbf{V}\ddot{\mathbf{x}} = \mathbf{f} - \mathbf{B}\ddot{\mathbf{x}}. \quad (19)$$

Due to orthogonality, we may analyse  $n$  pieces of equations with one unknown. The  $r$ -th equation is:

$$\ddot{x}_r + \gamma \omega_{ru} \dot{x}_r + \omega_{ru}^2 x_r = q_r. \quad (20)$$

In this case

$$q_{r_t + \theta \Delta t} = f_{r_t + \theta \Delta t} - \mathbf{b}_{r_t + \theta \Delta t}^T \ddot{\mathbf{x}}_{t + \theta \Delta t}, \quad (21)$$

$$f_{r_t + \theta \Delta t} = f_{r_t} + \theta (f_{r_t + \Delta t} - f_{r_t}),$$

$$\mathbf{b}_{r_t + \theta \Delta t} = \mathbf{b}_{r_t} + \theta (\mathbf{b}_{r_t + \Delta t} - \mathbf{b}_{r_t}).$$



According to the considerations before, we get

$$\left( \omega_{ru}^2 + \frac{6}{(\theta\Delta t)^2} + \frac{3}{\theta\Delta t} \gamma\omega_{ru} \right) x_{r_t+\theta\Delta t} = q_{r_t+\theta\Delta t} + \frac{6}{(\theta\Delta t)^2} x_{r_t} + \frac{6}{\theta\Delta t} \dot{x}_{r_t} + 2\ddot{x}_{r_t} + \gamma\omega_{ru} \left( \frac{3}{\theta\Delta t} x_{r_t} + 2\dot{x}_{r_t} + \frac{\theta\Delta t}{2} \ddot{x}_{r_t} \right). \quad (22)$$

In equation (21) it can be seen that  $q_{r_t+\theta\Delta t}$  contains the vector  $\ddot{x}_{r_t+\theta\Delta t}$ , thus, an iteration procedure must be used. According to Györgyi's experience [8] the iteration method does not converge always, therefore, it would be expedient to apply the direct solution of the following m-order equation system:

$$\left[ \tilde{\mathbf{D}} + \frac{6}{(\theta\Delta t)^2} \mathbf{B}_{t+\theta\Delta t} \right] \mathbf{x}_{t+\theta\Delta t} = \mathbf{f}_{t+\theta\Delta t} + (\mathbf{B}_{t+\theta\Delta t} + \mathbf{E}) \left( \frac{6}{(\theta\Delta t)^2} \mathbf{x}_t + \frac{6}{\theta\Delta t} \dot{\mathbf{x}}_t + 2\ddot{\mathbf{x}}_t \right) + \mathbf{G} \left( \frac{3}{\theta\Delta t} \mathbf{x}_t + 2\dot{\mathbf{x}}_t + \frac{\theta\Delta t}{2} \ddot{\mathbf{x}}_t \right). \quad (23)$$

Here,  $\tilde{\mathbf{D}}$  and  $\mathbf{G}$  are diagonal matrices with the  $r$ -th element of  $\omega_r^2 + \frac{6}{(\theta\Delta t)^2} + \frac{3}{\theta\Delta t} \gamma\omega_{ru}$  and  $\gamma\omega_{ru}$  respectively.

#### 4.2 Solution in the case of combined external and internal damping

Out of the tasks analysed, the one in 2.3 is the most general. There are several problems in analysis. One of this is that the damping matrix is not proportional to the mass or stiffness matrix, thus, it cannot be diagonalized. Neither the coefficient of internal damping can be considered constant. Of course, its typical value is not identical for the structure and the vehicle. Determination of the number of eigenvectors necessary for analysing the structure's displacement with proper accuracy is sophisticated by the appearance of the dynamical factors of the vehicle in the system. However, all these problems can be controlled if the matrix differential equation is generated in the way described in 2.3. Let us rewrite the relationship (7) in the following form:

$$\mathbf{M}_C \ddot{\mathbf{u}} + \mathbf{K}_C \mathbf{u} = -\mathbf{C}_C \dot{\mathbf{u}} + \mathbf{r}_1(t) - \mathbf{M}_1(t) \ddot{\mathbf{u}} - \mathbf{C}_1(t) \dot{\mathbf{u}} - \mathbf{K}_1(t) \mathbf{u} \quad (24)$$

The matrices on the right-hand side are hyperdiagonals. Accordingly, also the eigenvalue problem  $\mathbf{K}_C \mathbf{v} = \omega^2 \mathbf{M}_C \mathbf{v}$  splits in two parts. If we generate some smallest eigenvectors both for the structure and the vehicle ( $m_B$  for the structure and  $m_V$  for the vehicle).

Collect the eigenvectors in matrix  $\mathbf{V}$ , also matrix  $\mathbf{V}$  becomes a hyperdiagonal (its blocks will be quadratic only in that case when all the eigenvectors both for the structure and the vehicle were computed).

If we seek the solution again in the form of  $\mathbf{u} = \mathbf{V}\mathbf{x}$ , the task becomes again diagonalizable. Nothing will hamper the consideration of the proportional internal damping - different for the bridge structure and the vehicle.

$$\mathbf{V} = \begin{array}{|c|c|} \hline \mathbf{V}_B & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{V}_V \\ \hline \end{array}$$

As the individual blocks of the stiffness and mass matrices will be only multiplied by the  $\mathbf{V}$  matrix blocks belonging to them due to the structure of  $\mathbf{K}_C$  and  $\mathbf{M}_C$  and  $\mathbf{V}$ , the internal friction may be different for the blocks. The relationship for calculation of vector  $\mathbf{x}$  is the following:

$$\begin{aligned} \left[ \hat{\mathbf{D}} + \frac{6}{(\theta\Delta t)^2} \hat{\mathbf{B}}_{t+\theta\Delta t} + \frac{3}{\theta\Delta t} (\hat{\mathbf{H}}_C + \hat{\mathbf{H}}_{t+\theta\Delta t}) + \hat{\mathbf{P}}_{t+\theta\Delta t} \right] \mathbf{x}_{t+\theta\Delta t} = \\ = \hat{\mathbf{f}}_{t+\theta\Delta t} + (\hat{\mathbf{B}}_{t+\theta\Delta t} + \mathbf{E}) \left( \frac{6}{(\theta\Delta t)^2} \mathbf{x}_t + \frac{6}{\theta\Delta t} \dot{\mathbf{x}}_t + 2\ddot{\mathbf{x}}_t \right) + \\ + (\hat{\mathbf{G}} + \hat{\mathbf{H}}_C + \hat{\mathbf{H}}_{t+\theta\Delta t}) \left( \frac{3}{\theta\Delta t} \mathbf{x}_t + 2\dot{\mathbf{x}}_t + \frac{\theta\Delta t}{2} \ddot{\mathbf{x}}_t \right). \end{aligned} \tag{25}$$

Here vector  $\mathbf{x}$  has the size of  $m_B + m_V$  and  $\mathbf{E}$  is a unit matrix. Matrices  $\hat{\mathbf{D}}$  and  $\hat{\mathbf{G}}$  are hyperdiagonal matrices, in their first element  $m_B$  with the characteristics of the structure, then those of the vehicle:

$$\hat{\mathbf{D}} = \left\langle \dots, \omega_r^2 + \frac{6}{(\theta\Delta t)^2} + \frac{3}{\theta\Delta t} \gamma_B \omega_{ru}, \dots, \omega_s^2 + \frac{6}{(\theta\Delta t)^2} + \frac{3}{\theta\Delta t} \gamma_V \omega_{su}, \dots \right\rangle,$$

$$\hat{\mathbf{G}} = \left\langle \dots, \gamma_B \omega_{ru}, \dots, \gamma_V \omega_{su}, \dots \right\rangle.$$

Effect of masses in the supports of the vehicle is contained in  $\hat{\mathbf{B}}(t)$ . It is clear that only block  $ii$  generated by the eigenvectors of the structure of the size  $m_B \times m_B$  is not zero.

$$\hat{\mathbf{B}} = \begin{array}{|c|c|} \hline \begin{array}{c} i \\ \mathbf{V}_B^T \mathbf{M}_1(t) \mathbf{V}_B \\ \mathbf{0} \end{array} & \begin{array}{c} j \\ \mathbf{0} \\ \mathbf{0} \end{array} \\ \hline \end{array} \begin{array}{l} i \\ j \end{array}$$

The damping system of the vehicle proportional to the velocity appears, on the one hand, in matrix  $\hat{\mathbf{H}}_C$  containing constant elements, on the other hand, in matrix  $\hat{\mathbf{H}}(t)$  expressing effect of dampers joined to vehicle supports. Only block  $jj$  of matrix  $\hat{\mathbf{H}}_C$  differs from zero, and this block can be computed from the following relationship:

$$\hat{\mathbf{H}}_{Cjj} = \mathbf{V}_V^T \mathbf{C}_C \mathbf{V}_V.$$

The block  $jj$  of matrix  $\hat{\mathbf{H}}(t)$  is zero, but this can be computed from the

$$\hat{\mathbf{H}}(t) = \mathbf{V}^T \mathbf{C}_1(t) \mathbf{V}$$

relationship. Matrix  $\hat{\mathbf{P}}(t)$  expressing the effect of the vehicle springs joined to vehicle supports is similar to matrix  $\hat{\mathbf{H}}(t)$  :

$$\hat{\mathbf{P}}(t) = \mathbf{V}^T \mathbf{K}_1(t) \mathbf{V}.$$

Also vector  $\hat{\mathbf{f}}(t)$  can be partitioned in two parts. The  $i$ -th block is time-dependent, while  $j$ -th block will be zero:

$$\hat{\mathbf{f}}(t)_i = \mathbf{V}_B^T \mathbf{r}_1(t).$$

The road vehicle demonstrated in 2.2 will be simpler than this general model highlighted. In that case, matrix  $\hat{\mathbf{B}}(t)$  will be zero.

## 5. Numerical experience

### 5.1 Data of the structure and vehicle tested

During numerical tests, computations were carried out on a reality-like structure. Data of the bridge spanning 30 m are shown in Fig. 4. The internal damping factor  $\gamma$  was 0.1.

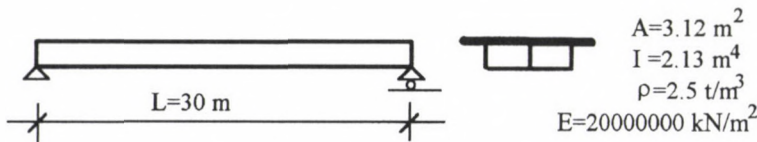
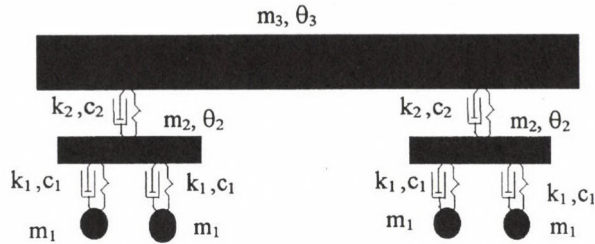


Fig. 4. Data of the structure tested

On the bridge, a railway vehicle was moving. Its internal friction factor was chosen 0.05, its stiffness, mass and damping data are contained in Fig. 5. Vehicle velocities were taken between 0 and 50 m/s in the different cases.





$$m_1 = 2,5 \text{ t}, m_2 = 10 \text{ t}, m_3 = 50 \text{ t}, \theta_2 = 8 \text{ tm}^2, \theta_3 = 1500 \text{ tm}^2$$

$$k_1 = 3000 \text{ kN/m}, k_2 = 2100 \text{ kN/m}, c_1 = 20 \text{ kNs/m}, c_2 = 150 \text{ kNs/m}$$

Fig. 5. Data of the railway vehicle

### 5.2 Eigenvectors of the structure needed for the modal analysis

The dynamical task was solved by modal analysis. For determination of the number of eigenvectors needed for properly correct solution, a numerical experiment was performed. Displacement of the central point of the structure was examined reckoning with increasing numbers of eigenvectors. Due to the nature of loading, role of components belonging to even number (antimetric) vibration forms in solution is not considerable, the size of overload changes essentially with the appearance of the new odd one. Table 1 shows the percentages of excess displacements generated by the dynamical effect at different velocities.

Table 1. Additional displacement due to dynamical effect

number of eigenvectors → velocities[m/s] ↓	1	3	5	7
10	2.1	1.0	0.8	0.9
20	2.3	1.3	1.2	1.2
35	5.5	4.6	4.4	4.4
50	20.2	18.8	18.7	18.7

The data above reveal that for the structure consideration of already five eigenvectors provides the additional dynamical displacements with sufficient accuracy, nevertheless, we calculated with 7 structural eigenvectors. As the vehicle model had six degrees of freedom, in this case we had to solve a system of equations with 13 unknowns in the individual steps, what did not require considerable computation time. The time step applied when solving the task was reckoned with by a value of  $\Delta t = \frac{T_p}{10}$  on the basis of the smallest vibration time belonging to the eigenvectors considered in the solution and in accordance with literature's recommendation. In the case tested, the vibration time belonging to the highest number eigenvector considered in the test was smaller than the smallest vibration time of the vehicle, thus, it was the structure that determined the time step size.

### 5.3 Influence of damping on additional dynamical displacements

During tests, it was analysed how the change of the vehicle's damping system influences the dynamical excess displacements. The first row of Table 2 contains the additional displacements computed with the initial damping value in the percentage of the static displacement. The second row shows the case when the coefficient of the vehicle's damping system decreased to the tenth of the original value. In the next row, the coefficient of the damping structure is given with tenfold value. Finally, the last row underlines the importance of considering the internal friction of the structure. Without it, essentially higher additional displacements would be found than in reality. Values of the table relate to two velocities. It can be seen that the damping effect is much stronger with a higher vehicle velocity than at an average speed. Tests revealed that the internal friction of the vehicle in cases tested did not influence the movements of the structure.

Table 2. Effect of the dampers of the moving vehicle and the structure's internal friction

velocity of the vehicle → damping ↓	v=20 m/s	v=50 m/s
$\gamma_B, \gamma_V, c$	1.2	18.7
$\gamma_B, \gamma_V, 0.1c$	0.6	16.2
$\gamma_B, \gamma_V, 10c$	1.3	19.2
$\gamma_B = 0, \gamma_V, c$	2.4	24.7

## 6. Summary

The paper discusses the test of bridges under moving vehicle applying the finite element method, and shows a procedure able to compute the influence of the external damping proportional to velocity and, at the same time, is fit for considering structural damping, too. The computation algorithm relates to a general case, however, the effect of moving mass regarded as a material point, effect of two-axle road vehicle and multi-axle railway vehicle is extra dealt with. Application of modal analysis essentially reduced the computation time. The numerical results cited proved the efficiency of the algorithm for joint test of moving vehicle and the structure as well as the importance of taking the influence of the internal friction of the bridge structure into consideration. As per the numerical results, the possible most correct analysis of the interaction of the vehicle's damping structure and the bridge structure is necessary mainly in the case of high-speed vehicles.

## Acknowledgement

The author thanks for supporting the research by the Hungarian Scientific Research Fund (OTKA No. 11499-009).

## REFERENCES

1. Frýba, L. : *Vibration of Solids and Structures under Moving Loads*. Academia. Prague 1972
2. Green, M. F. - Cebon, D.: *Dynamic Response of Highway Bridges to Heavy Vehicle Loads: Theory and Experimental Validation*. *Journal of Sound and Vibration* 170 (1994) 51-78
3. Bathe, K. J.-Wilson, E. L.: *Numerical Methods in Finite Element Analysis*. Prentice-Hall. New Jersey 1976
4. Inbanathan, M. J.- Wieland, M.: *Bridge Vibrations due to Vehicle Moving Rough Surface*. *Journal of Structural Engineering* 113 (1987) 1994-2008
5. Olsson, M.: *Finite Element, Modal Co-ordinate Analysis of Structures Subjected to Moving Loads*. *Journal of Sound and Vibration* 99 (1985) 1-12
6. Györgyi, J.: *Viscous and Hysteretic Damping in Vibration of Structures*. *Periodica Polytechnica Civil Engineering* 29 (1985) 23-31
7. Györgyi, J.: *Calculation for the Vibration of Structures: A partial eigenvector problem solution*. *Acta Technica Hung.* 99 (1986) 103-123
8. Györgyi, J.: *Dynamic Analysis of Structures supporting a Moving Mass, exposed to External and Internal Damping*. *Periodica Polytechnica Civil Engineering* 38 (1994) 387-398



## IMPERFECTION SENSITIVITY OF THE CLASSICAL STABLE-SYMMETRIC BIFURCATION PROBLEM MODIFIED BY BILINEAR MATERIAL

KURUTZ, M.\*

(Received: 15 July 1996)

In this paper, the effect of different cases of imperfections applied to the modified classical stable symmetric bifurcation problem is analysed. The classical bifurcation example is modified by bilinear elastic material behaviour, by taking Hencky-type elastic-plastic behaviour into consideration. Beside the classical geometric imperfections, certain material and some loading imperfections will also be considered. Perturbations in the elastic and plastic material characteristics, moreover, in the loading device is considered. Namely, it may happen in the practice that the nature of loading process is not perfectly "dead" being perfectly independent of the occurring deflections. The deviation from the dead loading program can as loading imperfection be considered.

The present analysis is based on the papers [3,4,5] of the author where the nonsmooth stability analysis have been introduced for those dead loaded structures which have polygonal material, consequently nonsmooth internal potential. In [6,7] these results have been extended to the cases of configuration-dependent conservative loading, by introducing polygonal loading functions yielding nonsmooth external potential.

The aim of this paper is to obtain the geometric, material and loading imperfection-sensitivity functions of the classical stable-symmetrical bifurcation problem modified by bilinear elastic material.

### 1. Introduction

The three classical examples of bifurcation, the stable and unstable symmetric and the asymmetric bifurcation classified by Koiter in 1945 are detailed in Thompson & Hunt and Bazant & Cedolin in [1, p. 5-20] and in [2, p. 238-256], respectively. These type of dead loaded simple structures seen in Fig. 1.a., b. and c., respectively, are assumed to be composed by a perfectly rigid element of length  $l$  pinned to a rigid foundation and connected to the support by linear elastic springs in which the material behaviour is concentrated.

The model of *stable-symmetric bifurcation* in Fig. 1.a., named "hinged cantilever" is supported by a rotational spring, while the model of *unstable-symmetric bifurcation* in Fig. 1.b., the "propped cantilever" is supported by a horizontal extensional spring. Finally, the model of *asymmetric bifurcation* in Fig. 1.c. is supported by a skew extensional spring. Fig. 1. shows the three classical curves of bifurcation, respectively, namely, the associated postbifurcation equilibrium paths  $\lambda(q)$  of the structures, respectively, by applying one parameter dead load  $F=F(\lambda)=\lambda F_0$  where  $F_0=1$ , and  $\lambda$  is the load parameter.

In this paper we deal with the first classical problem only, seen in Fig. 1.a.

---

\* Kurutz, Márta, H-1118 Budapest, Serleg u. 8, Hungary

The functional finitization needs to introduce the *generalised coordinates*, namely, *displacement parameters*. In these simple cases, the single generalised coordinate is the angle of rotation  $\vartheta=q$  at the support hinge, thus, the function  $u=l(1-\cos q)$  of the vertical displacement of the top of the cantilever is used as compatibility condition. To analyse imperfection-sensitivity, perturbation or *imperfection parameters*  $\varepsilon$  are introduced. The system corresponding to  $\varepsilon=0$  is perfect. Thus, for the potential energy function, we obtain  $\pi=\pi(q,\lambda,\varepsilon)$  having displacement, load and imperfection parameters as variables.

Three type of imperfection will be analysed: geometric, material and loading type one. Here we assume that the system is geometrically perfect in the sense that the springs are unstrained when the link is vertical. *Geometric imperfection*  $\varepsilon_g$  means an initial rotation at the support hinge.

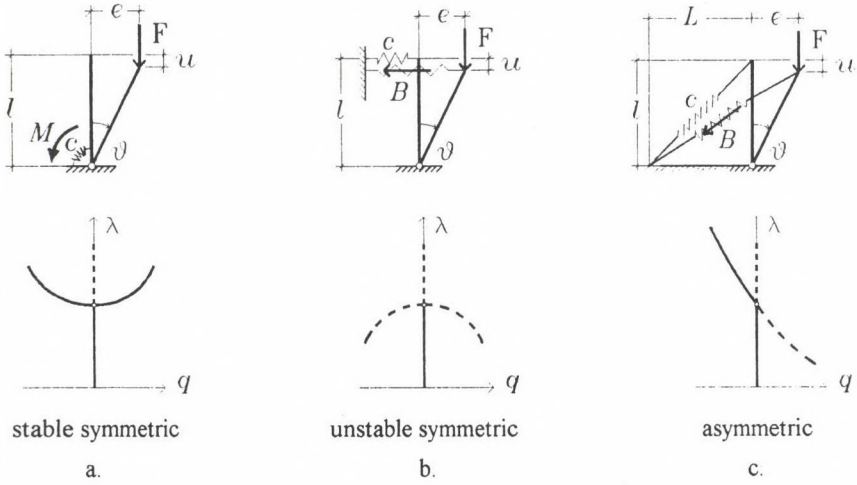


Fig. 1. The three classical bifurcation models

In the classical cases [1,2] the springs are linear and have a spring constant  $c$ . The spring moment represents the *stress* variable, while the deformation of the spring represents the *strain* variable. The classical cases assume that the elastic system is perfect if the stress-strain relation is  $M=c\vartheta$ . In [3] the equilibrium paths of linear elastic perfectly plastic, moreover, for perfectly rigid-plastic structures are detailed by using nonsmooth analysis. Thus, in this paper, the linear elastic-plastic or rigid-plastic system is assumed to be perfect, and in Fig 2.a. and b. the cases of material imperfections can be seen. The *material imperfection* can occur in two form: imperfection  $\varepsilon_e$  in the elastic modulus  $c$ , and imperfection  $\varepsilon_p$  in the plastic yield limit  $M_p$ . The material imperfections are specified by the functions (Fig.2.a. and b.)

$$M(\vartheta) = \begin{cases} (c - \varepsilon_e)\vartheta & \text{in elastic state} \\ M_p + \varepsilon_p & \text{in plastic state} \end{cases} \quad (1)$$

where  $\varepsilon_e$  and  $\varepsilon_p$  can be positive and negative, as well. Taking the *geometric imperfection*  $\varepsilon_g$  into account, the stresses in the elastic state form (Fig.2.c.)

$$M(\vartheta) = c(\vartheta - \varepsilon_g) \quad (2)$$



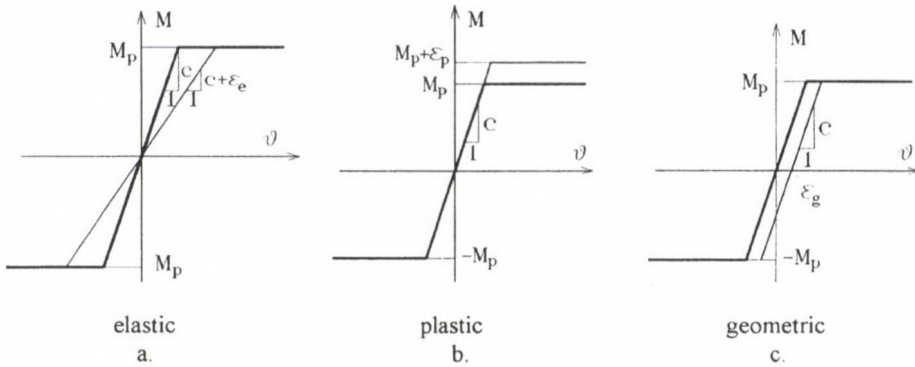


Fig. 2. Material and geometric imperfections

Consider now the *loading imperfection*, where  $F$  is the given external force load acting vertically on the top of the cantilevers in Fig. 1. If the force load  $F$  is independent of the occurring displacement  $u$ , it is a dead load, while if it is in interaction with the displacement  $u$ , it is a configuration-dependent load. Some loading devices, for example the hydraulic loading, show certain deformation-sensitive characteristics leading to changes in the classical postbifurcation behaviour of structures. Here we assume that the load is perfect in sense of dead load, and, it is imperfect if the load shows some deformation-sensitive characteristics.

Fundamental aspects and classification of loading types is detailed in [2, p. 207 and p. 224]. *Dead* type conservative loading device supposes the applied load to be independent of the occurring deflections. This kind of loading process can be characterised by a constant load-deflection function  $F(\lambda)=\lambda F_o$ , where  $\lambda$  is the load parameter and  $F_o$  is the initial value of the load (Fig.3.a.). *Variable* or *configuration-dependent* conservative loading process assumes the applied load to be dependent on the occurring deflections, but independent of the properties of the structure. This kind of loading device can be specified by a variable load-deflection function  $F(\lambda,u)=\lambda F_o+f(u)$ . In this case, the load is divided into two parts: the *controllable* (dead) part  $\lambda F_o$  governed by the load parameter  $\lambda$ , and the *deformation-sensitive* part  $f(u)$  specified as a linear or nonlinear function. In this paper, we consider linear loading functions  $F(\lambda,u)=\lambda F_o+fu$  only, by assuming the *loading modulus*  $f=\epsilon_1$  to be the *imperfection of the loading device*, thus (Fig.3.b.):

$$F(\lambda,u)=\lambda F_o+\epsilon_1 u \tag{3}$$

In [6] it has been shown that the deformation-sensitive part of the loading characterised by given load-deflection functions shows similarity to the material behaviour characterised by given stress-strain functions. Consequently, the variable loading can be handled in a similar way to that of the material. Obviously, in the tangent stiffness of the structure, beside the *tangent modulus of the material*, the *tangent modulus of the load* appears.

It has been proved by the author in [3,4,5] that by approximating the *nonlinear material* functions by polygonal, and by using *nonsmooth analysis*, the equilibrium paths can be obtained as the envelope of the individual equilibrium paths related to each segment of the material polygon. In [6,7] the results of nonlinear material have been extended to the case of *nonlinear loading* programs.



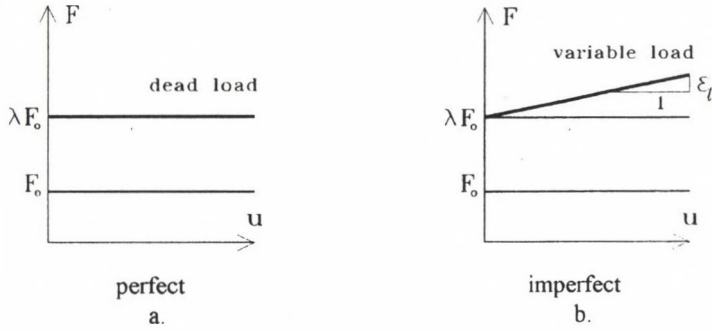


Fig. 3. Perfect and imperfect loading devices

Let us consider now the different cases of imperfection-sensitivity, namely, the modifications of the classical stable-symmetric postbifurcation equilibrium paths and the associated imperfection-sensitivity functions.

## 2. Geometric imperfection-sensitivity

The linear elastic material behaviour of the hinged cantilever seen in Fig.4.a. can be characterised by the stress-strain relation  $M=c\vartheta$  where  $c$  is the spring constant,  $M$  is the moment in the spring and  $\vartheta=q$  is the angle of rotation of the cantilever.

In the case of fully perfect structure, the postbifurcation equilibrium path is given by the classical expression

$$\lambda(q) = \frac{c}{F_0 l} \frac{q}{\sin q} \quad \lambda_{cr} = \frac{c}{F_0 l} \quad (4)$$

which latter represents the stable symmetric point of bifurcation at the critical equilibrium state at  $q=0$ . For the associated function  $K(q)$  of the tangent stiffness we obtain

$$K_t(q) = c \left( 1 - \frac{q}{\tan q} \right) \quad (5)$$

In the case of perfectly elastic-plastic structure, the modified postbifurcation equilibrium paths form a set consisting of the following nonsmooth function [3]

$$\lambda(q) = \frac{1}{F_0} \begin{cases} \frac{M_{p1}}{l \sin q} & \text{for plastic loading if } dq > 0 \\ \frac{c(q-q_0)}{l \sin q} & \text{for elastic loading - unloading} \\ \frac{M_{p2}}{l \sin q} & \text{for plastic loading if } dq < 0 \end{cases} \quad (6)$$

seen in Fig.4.b. by applying  $c=1$ ,  $M_{p1}=1$ ,  $M_{p2}=-2$ ,  $l=1$  and  $F_o=1$ . Here  $q_o$  represents the actual residual plastic rotation associated with the unloading/reloading material paths. In Fig.4.b. the interval  $-1.0 \leq q_o \leq 1.0$  is illustrated only, by the steps  $\Delta q_o=0.1$ . Since infinite number of unloading-reloading paths can occur in  $-\pi \leq q_o \leq \pi$ , the equilibrium paths form a set, bounded by the four curves associated with the plastic limits  $M_{p1}$  and  $M_{p2}$ .

The functions seen in Fig. 4. a. and b. represent the modified stable symmetric bifurcation problem, due to perfectly elastic-plastic material, detailed in [3]. However, in this analysis, the unloading is excluded, thus, a Hencky-type bilinearly elastic material is considered where the stress-strain diagram is similar to that of the elastic-plastic material, but here the unloading uses the paths of the loading, thus, no plastic strain occurs, as seen in Fig. 5. a.

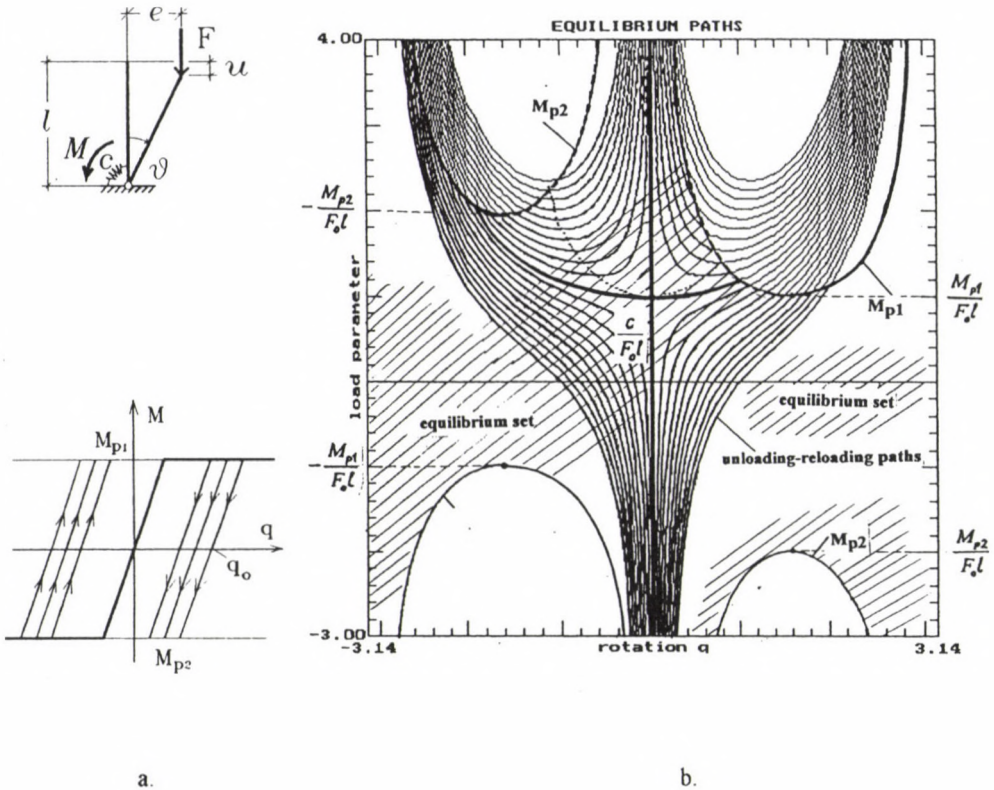


Fig. 4. Equilibrium paths of the perfect elastic-plastic stable symmetric bifurcation problem

By considering a *geometric imperfection*  $\varepsilon_g$ , the equilibrium paths are modified to

$$\lambda(q, \varepsilon_g) = \frac{1}{F_o} \begin{cases} \frac{M_{p1}}{l \sin q} & \text{for plastic states if } dq > 0 \\ \frac{c(q - \varepsilon_g)}{l \sin q} & \text{for elastic states} \\ \frac{M_{p2}}{l \sin q} & \text{for plastic states if } dq < 0 \end{cases} \quad (7)$$

seen in Fig. 5.b. for different values of  $-1 \leq \varepsilon_g \leq +1$ , by the steps 0.1 and by applying  $c=1$ ,  $M_{p1}=1$ ,  $M_{p2}=-2$ ,  $l=1$  and  $F_o=1$ . By comparing the geometric imperfection seen in Fig. 2.c. and the elastic-plastic unloading seen in Fig. 4.a., moreover, the expressions (6) and (7), the conclusion can be drawn that the two cases seem to be similar. Thus, the equilibrium paths of a geometrically imperfect Hencky-type structure can be obtained in the same way that of the elastic-plastic unloading/reloading paths.

In Fig. 5.b. the function of the classical elastic critical states are indicated by dotted lines, while the elastic-plastic critical states, namely, the intersection of the geometric imperfection paths and the plastic loading paths are indicated by heavy lines. Thus, the imperfection-sensitivity function consists of the classical and the plastic parts. By applying Taylor expansion for determining the critical states in term of the geometric imperfection, from the condition of the vanishing tangent stiffness, we obtain  $q_{cr} = \sqrt[3]{-3\varepsilon_g}$ , thus the *function of the geometric imperfection-sensitivity* reads

$$\lambda(\varepsilon_g) = \frac{1}{F_o} \begin{cases} \frac{M_{p1}}{l \sin(\varepsilon_g + M_{p1}/c)} & \text{for plastic states} \\ \frac{c(\sqrt[3]{-3\varepsilon_g} - \varepsilon_g)}{l \sin(\sqrt[3]{-3\varepsilon_g})} & \text{for elastic states} \\ \frac{M_{p2}}{l \sin(\varepsilon_g + M_{p2}/c)} & \text{for plastic states} \end{cases} \quad (8)$$

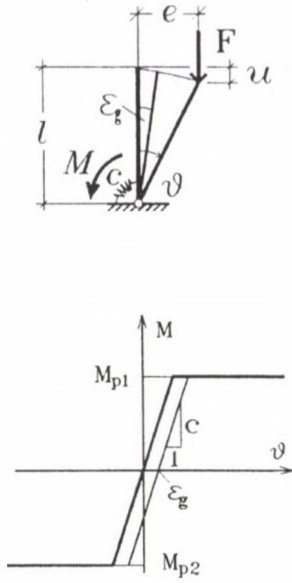
seen in Fig. 5.c. distinguished by heavy line. These functions are certain segments of (8), limited by the values

$$\varepsilon_{g1} = \frac{\pi}{2} - \frac{M_{p1}}{c} \quad \varepsilon_{g2} = -\frac{\pi}{2} - \frac{M_{p2}}{c} \quad (9)$$

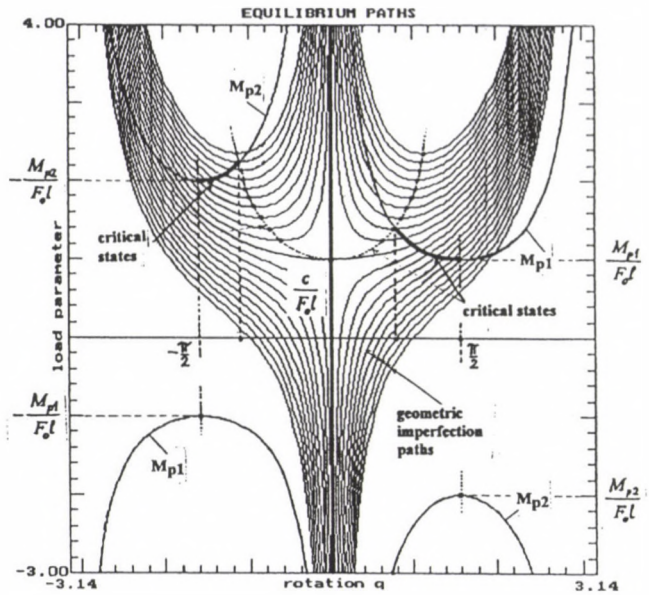
and by  $\varepsilon_{g3}$  and  $\varepsilon_{g4}$  which represent the tangent points of the classical elastic and the plastic geometric imperfection-sensitivity functions. It can be concluded that the elastic and plastic imperfection-sensitivity functions are tangential to each other.

This fact can be observed in a more attractive way by applying different plastic moment values. Thus, consider now the material imperfections.

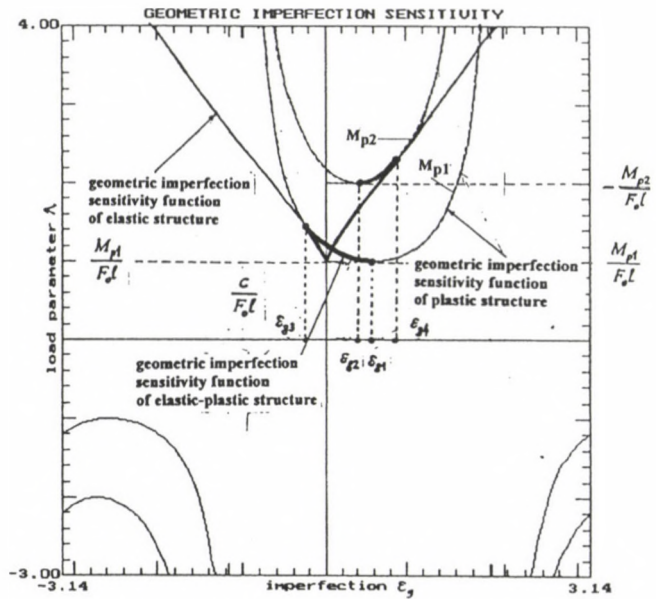




a.

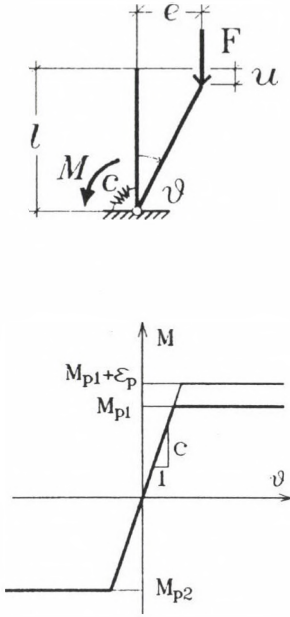


b.

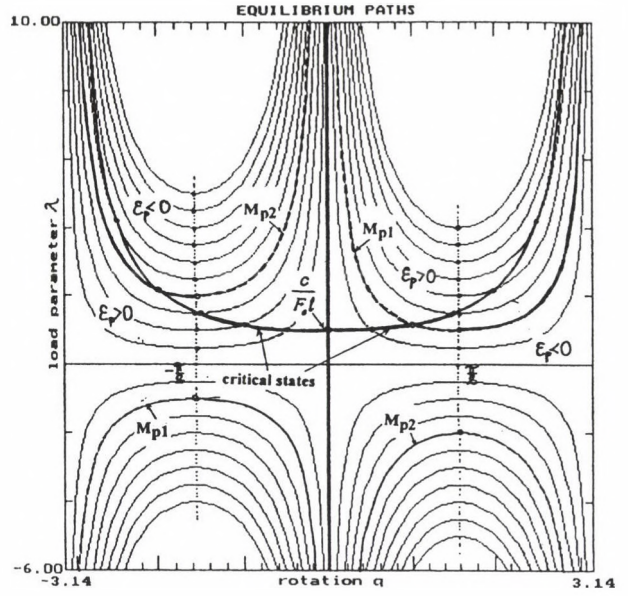


c.

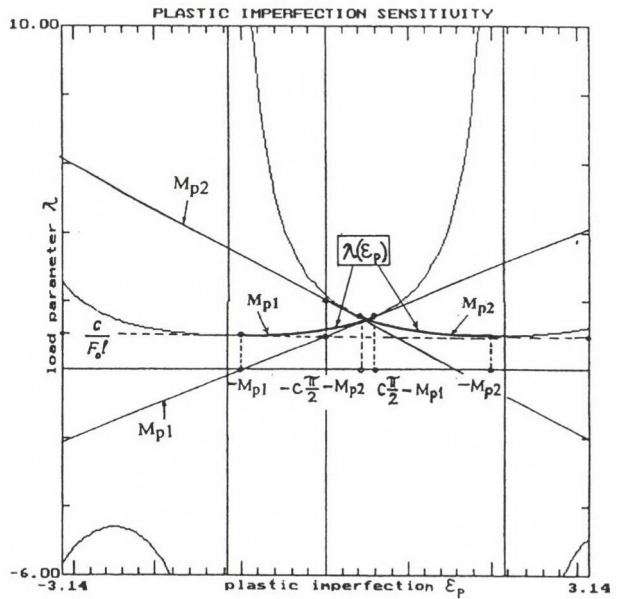
Fig. 5. Geometric imperfection-sensitivity



a.



b.



c.

Fig. 6. Plastic material imperfection-sensitivity

### 3. Material imperfection-sensitivity

Consider first the case of *plastic material imperfections* with perfect geometry, for an elastic-plastic system seen in Fig.6.a. In this case, the equilibrium paths (7) are modified to

$$\lambda(q, \varepsilon_p) = \frac{1}{F_o} \begin{cases} \frac{M_{p1} + \varepsilon_p}{l \sin q} & \text{for plastic states if } dq > 0 \\ \frac{cq}{l \sin q} & \text{for elastic states} \\ \frac{M_{p2} + \varepsilon_p}{l \sin q} & \text{for plastic states if } dq < 0 \end{cases} \quad (10)$$

seen in Fig.6.b. for  $-3 \leq \varepsilon_p \leq +3$ , by the steps 0.5, and by applying  $c=1$ ,  $M_{p1}=1$ ,  $M_{p2}=-2$ ,  $l=1$  and  $F_o=1$ , as previously. The *plastic imperfection-sensitivity function* then reads

$$\lambda(\varepsilon_p) = \frac{1}{F_o} \begin{cases} \frac{M_{p1} + \varepsilon_p}{l \sin((M_{p1} + \varepsilon_p)/c)} & \text{for } -M_{p1} \leq \varepsilon_p \leq \frac{c\pi}{2} - M_{p1} \\ \frac{M_{p2} + \varepsilon_p}{l \sin((M_{p2} + \varepsilon_p)/c)} & \text{for } -\frac{c\pi}{2} - M_{p2} \leq \varepsilon_p \leq -M_{p2} \end{cases} \quad (11)$$

seen in Fig.6.c. where the fact can be seen that the imperfection-sensitivity functions are limited again, due to the possible critical states in the interval  $-\pi/2 \leq q \leq \pi/2$  seen in Fig.6.b.

Let us consider now the plastic material imperfection with a simultaneous geometric imperfection seen in Fig.7. By increasing the yield limits, the possible imperfection-sensitive domain - between the line A of classical limit points of geometric imperfection and the line B of limit points of plastic imperfection - changes, as seen in Fig.7.a. Consequently, the actual segment of the imperfection-sensitivity function changes, too. This can be seen in Fig.7.b. where imperfection-sensitivity functions belonging to different plastic moment values are illustrated. All the plastic curves are tangent to the classical elastic imperfection-sensitivity curve, or better, the elastic curve is the envelope of the plastic curves. Thus, the domain of valid imperfection sensitivity functions forms a set. The correspondence between the domain of possible critical states and that of the associated imperfection-sensitivity set, bordered both by lines A and B can clearly be seen in Fig.7.a. and b.

Consider now the *imperfection of the elastic modulus* of an elastic-plastic structure seen in Fig.8.a. In this case the equilibrium path forms

$$\lambda(q, \varepsilon_e) = \frac{1}{F_o} \begin{cases} \frac{M_{p1}}{l \sin q} & \text{for plastic states if } dq > 0 \\ \frac{(c + \varepsilon_e)q}{l \sin q} & \text{for elastic states} \\ \frac{M_{p2}}{l \sin q} & \text{for plastic states if } dq < 0 \end{cases} \quad (12)$$



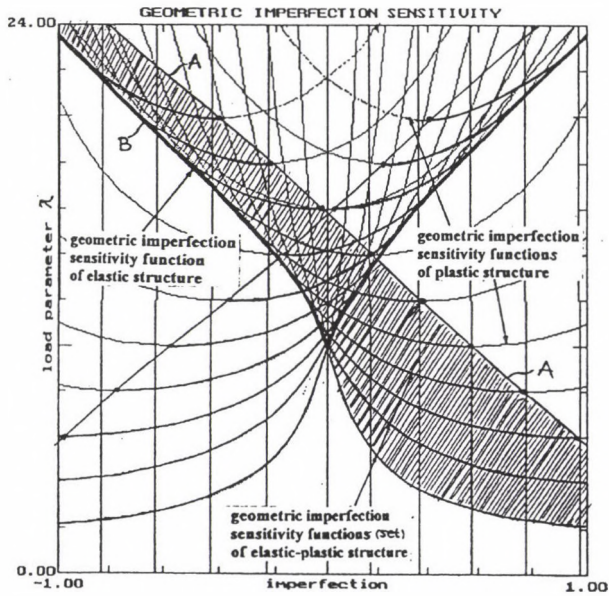
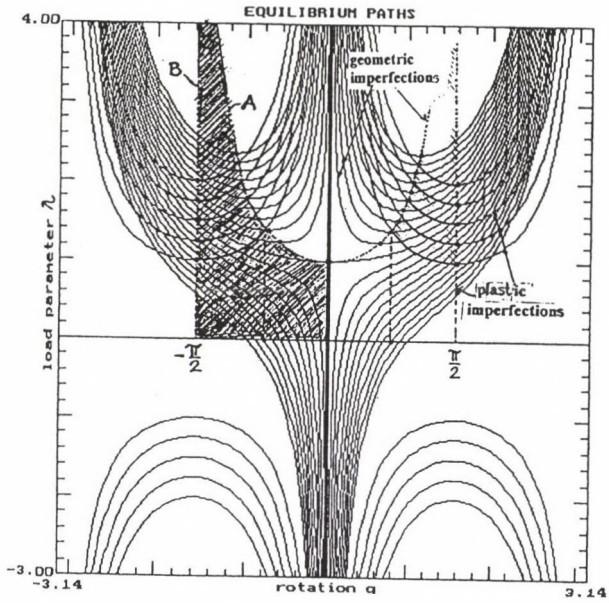
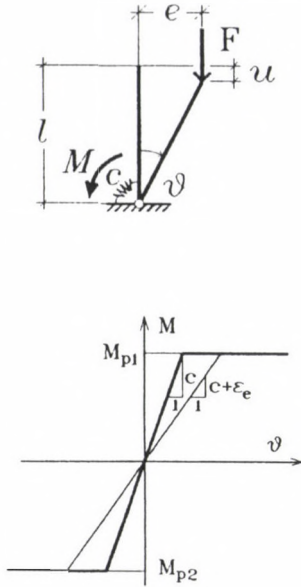
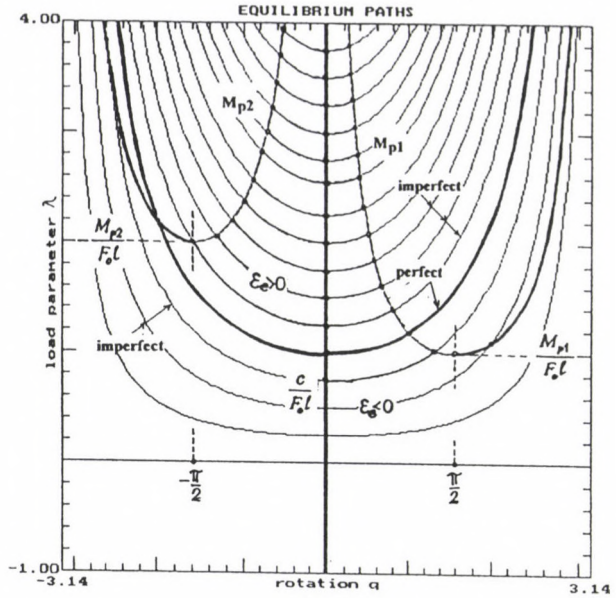


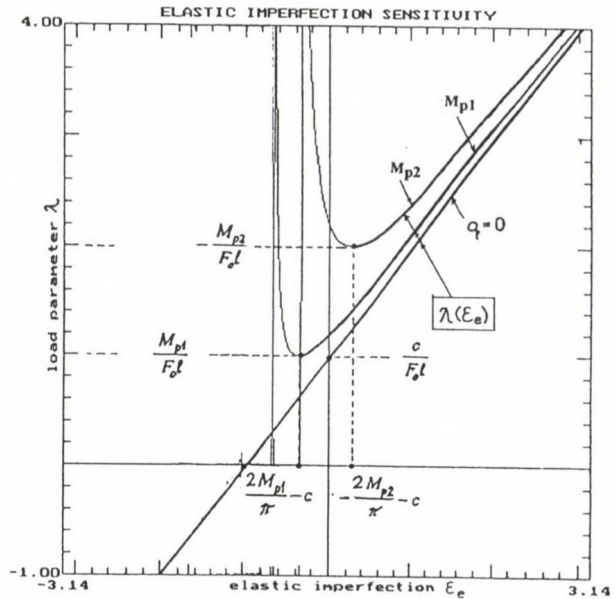
Fig.7. Geometric and plastic imperfection-sensitivity



a.



b.



c.

Fig. 8. Elastic material imperfection-sensitivity

seen in Fig.8.b. for  $-1 \leq \varepsilon_e \leq +3$ , by the steps 0.25 and by applying  $c=1$ ,  $M_{p1}=1$ ,  $M_{p2}=-2$ ,  $l=1$  and  $F_o=1$ , as previously. The associated *elastic imperfection-sensitivity functions* then reads

$$\lambda(\varepsilon_e) = \frac{1}{F_o} \begin{cases} \frac{M_{p1}}{l \sin(M_{p1}/(c + \varepsilon_e))} & \text{for } \frac{2M_{p1}}{\pi} - c \leq \varepsilon_e \leq +\infty \\ \frac{M_{p2}}{l \sin(M_{p2}/(c + \varepsilon_e))} & \text{for } -\frac{2M_{p2}}{\pi} - c \leq \varepsilon_e \leq +\infty \end{cases} \quad (13)$$

seen in Fig.8.c. Note that in this case the imperfection-sensitivity functions have a lower limit, due to the possible critical states in the interval  $-\pi/2 \leq q \leq \pi/2$  again, seen in Fig.8.b.

Consider now a Hencky-type perfectly rigid-plastic structure specified by the yield moments  $M_{p1}$  and  $M_{p2}$ , by assuming some *hardening or softening imperfections* seen in Fig.9.a. The hardening (softening) phenomenon is indicated by the positive (negative) sign of the hardening imperfection  $\varepsilon_h$ . The functions of equilibrium paths are as follows

$$\lambda(q, \varepsilon_h) = \frac{1}{F_o} \begin{cases} \frac{\varepsilon_h(q + \frac{M_{p1}}{\varepsilon_h})}{l \sin q} & \text{for hardening/softening states if } dq > 0 \\ (-\infty, +\infty) & \text{for } q = 0, dq = 0 \\ \frac{\varepsilon_h(q + \frac{M_{p2}}{\varepsilon_h})}{l \sin q} & \text{for hardening/softening states if } dq < 0 \end{cases} \quad (14)$$

seen in Fig.9.b. where the perfect structure is represented by the two curves belonging to  $M_{p1}=1$ , and  $M_{p2}=-2$ , and the illustrated imperfections are  $-1 \leq \varepsilon_h \leq +1$ , by the steps 0.2 and by applying the previously used constants. The limit points in Fig.9.b. determine the *functions of hardening imperfection-sensitivity*

$$\lambda(\varepsilon_h) = \frac{1}{F_o} \begin{cases} \frac{\varepsilon_h \left( 3\sqrt{\frac{3M_{p1}}{\varepsilon_h} + \frac{M_{p1}}{\varepsilon_h}} \right)}{l \sin \left( 3\sqrt{\frac{3M_{p1}}{\varepsilon_h}} \right)} & \text{for } -\frac{M_{p1}}{\pi} \leq \varepsilon_h \leq +\infty \\ \frac{\varepsilon_h \left( 3\sqrt{\frac{3M_{p2}}{\varepsilon_h} + \frac{M_{p2}}{\varepsilon_h}} \right)}{l \sin \left( 3\sqrt{\frac{3M_{p2}}{\varepsilon_h}} \right)} & \text{for } \frac{M_{p2}}{\pi} \leq \varepsilon_h \leq +\infty \end{cases} \quad (15)$$

seen in Fig.9.c. These functions are limited by the value of imperfections where the limit points vanish. Fig.10.a. illustrates the case of the hardening/softening imperfections if having an elastic-plastic material. In this case the equilibrium paths are modified to



$$\lambda(q, \varepsilon_h) = \frac{1}{F_o} \begin{cases} \frac{\varepsilon_h(q + \frac{M_{p1}}{\varepsilon_h})}{l \sin q} & \text{for hardening/softening states if } dq > 0 \\ \frac{cq}{l \sin q} & \text{for elastic state} \\ \frac{\varepsilon_h(q + \frac{M_{p2}}{\varepsilon_h})}{l \sin q} & \text{for hardening/softening states if } dq < 0 \end{cases} \quad (16)$$

seen in Fig.10.b. for the case of  $M_{p1}=1$ ,  $M_{p2}=-2$  with the softening/hardening imperfections as  $-1 \leq \varepsilon_h \leq +1$ , by the steps 0.2 and by applying other constants as used previously.

The function of the limit points associated with the imperfections seen in Fig.10.b. limits the domain of the possible critical points occurring as intersection points of the elastic and plastic behaviour. Thus, we can see again that the imperfection-sensitivity functions are limited in the tangent points  $\varepsilon_{h1}$  and  $\varepsilon_{h2}$  with the hardening/softening imperfection-sensitivity functions (15) of the rigid-plastic case, seen in Fig.10.c. The *hardening/softening imperfection-sensitivity functions* are as follows

$$\lambda(\varepsilon_h) = \frac{1}{F_o} \begin{cases} \frac{c M_{p1}}{(c - \varepsilon_h) l \sin(\frac{M_{p1}}{c - \varepsilon_h})} & \text{for } -\infty \leq \varepsilon_h \leq \varepsilon_{h1} \\ \frac{c M_{p2}}{(c - \varepsilon_h) l \sin(\frac{M_{p2}}{c - \varepsilon_h})} & \text{for } -\infty \leq \varepsilon_h \leq \varepsilon_{h2} \end{cases} \quad (17)$$

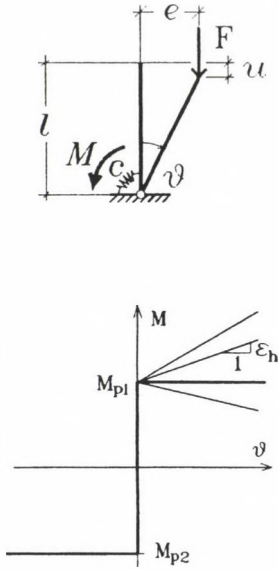
In Fig.10.c. the hardening imperfection-sensitivity functions are illustrated. Obviously, the system is sensible for the softening material behaviour, namely, for the damaging material behaviour with decreasing stress-strain diagram. That is why the imperfection-sensitivity functions are located mainly in the negative region of the imperfections.

#### 4. Loading imperfection-sensitivity

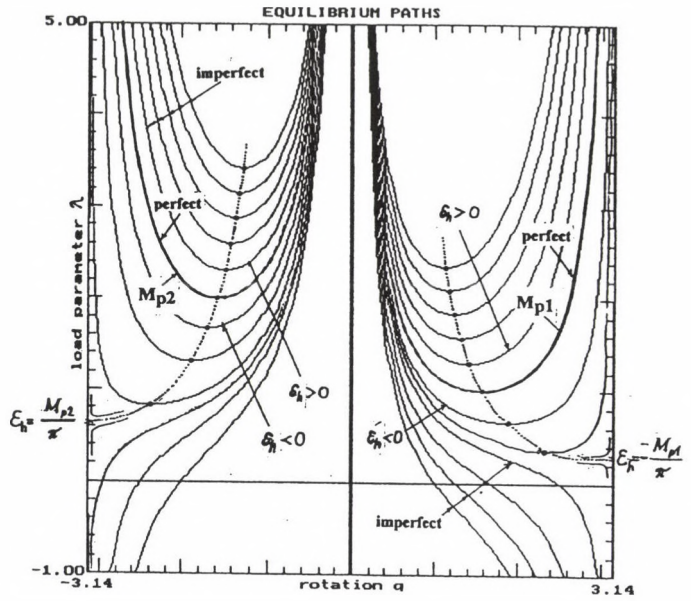
It happens in the practice that the loading device is not perfectly dead, namely it shows certain deformation-sensitive feature which may cause considerable change in the equilibrium paths of the structural system. In this paper, we take the variable characteristics as *loading imperfection* into consideration.

In the case of load imperfection seen in Fig.11.a., the material is assumed to be perfectly elastic. Thus, the equilibrium paths are smooth functions being written in the form [6,7]

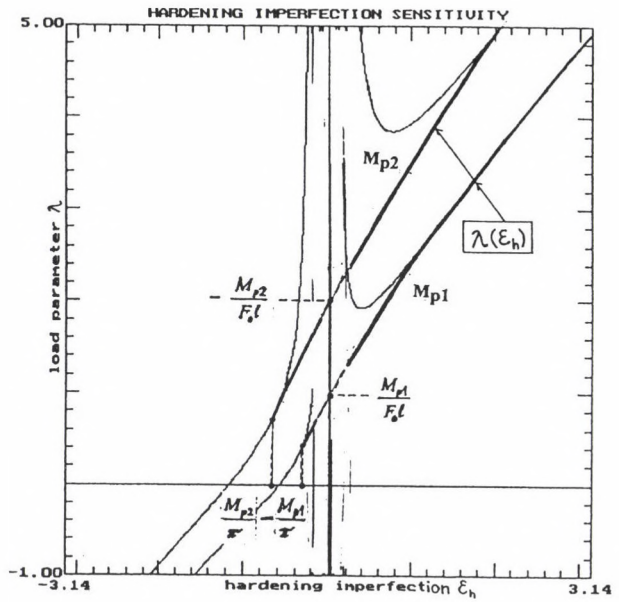
$$\lambda(q, \varepsilon_l) = \frac{1}{F_o} \left( \frac{c}{l \sin q} - \varepsilon_l l (1 - \cos q) \right) \quad \lambda_{cr} = \frac{c}{F_o l} \quad (18)$$



a.

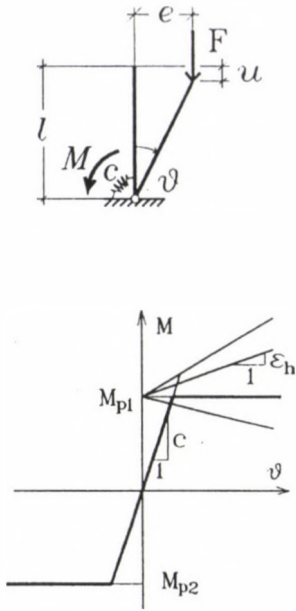


b.

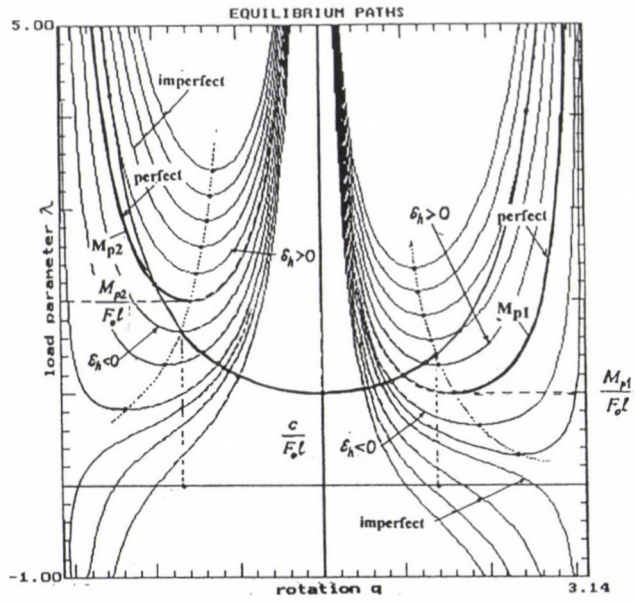


c.

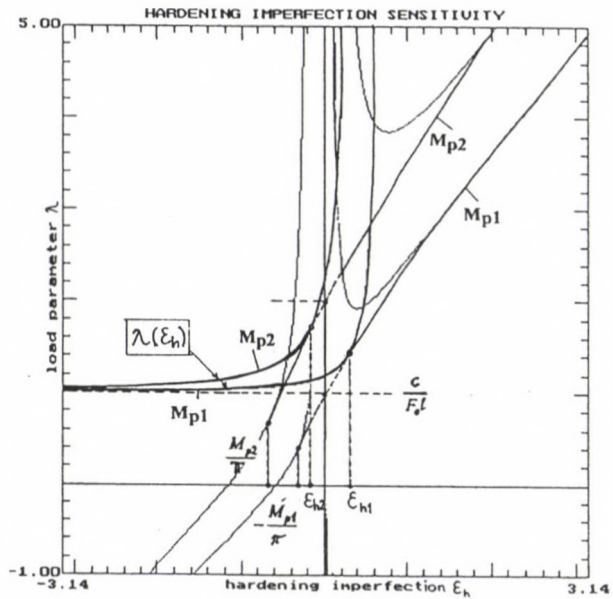
Fig. 9. Hardening/softening material imperfection-sensitivity of rigid-plastic structure



a.



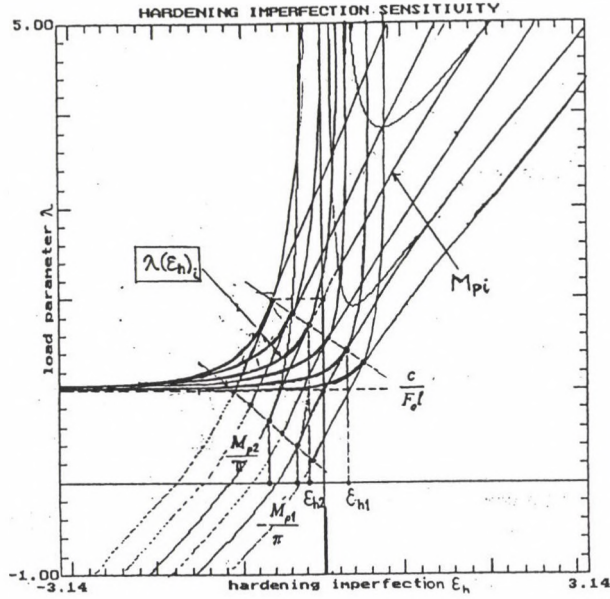
b.



c.

Fig. 10.a-c. Hardening/softening material imperfection-sensitivity of elastic-plastic structure





d.

Fig. 10.d. Hardening/softening material imperfection-sensitivity of elastic-plastic structure

with the same symmetric but not always stable point of bifurcation at the same critical point  $q=0$ , seen in Fig. 11. b. for different values of imperfection  $-1 \leq \epsilon_l \leq 3$  as load modulus, since  $\epsilon_l = f$ . From the graphs of the functions, the destabilising effect of the positive load modulus can be seen. Namely, from the vanishing associated tangent stiffness

$$K_t(q) = c \left(1 - \frac{q}{\tan q}\right) - \epsilon_l^2 \sin^2 q = 0 \quad (19)$$

the fact can be observed that at a certain positive value  $\epsilon_l = c/3l^2$  of the load imperfection (load tangent modulus), the classical stable bifurcation becomes unstable. This fact is illustrated in Fig. 11. b. where the important fact can evidently be observed that by increasing the load imperfection (load modulus) step by step by  $\Delta \epsilon_l = 0.2$ , the classical stable characteristic of the symmetric bifurcation tends to be lost, and the bifurcation becomes unstable by the value  $\epsilon_l = c/3l^2$  of the loading modulus. Thus, the value  $\epsilon_l = c/3l^2$ , by which the bifurcation mode changes, can as *critical load modulus* be considered. By connecting the critical (limit) points at  $q \neq 0$ , belonging to the different values of load moduli, the obtained function represents the *loading imperfection-sensitivity function* seen in Fig. 11. c.

$$\lambda(\varepsilon_l) = \frac{1}{F_o} \left( \frac{c}{l} \frac{\sqrt[4]{12 - \frac{4c}{\varepsilon_l l^2}}}{\sin\left(4\sqrt[4]{12 - \frac{4c}{\varepsilon_l l^2}}\right)} - \varepsilon_l l \left(1 - \cos\left(4\sqrt[4]{12 - \frac{4c}{\varepsilon_l l^2}}\right)\right) \right) \quad (20)$$

As a conclusion, we can state that the positive loading imperfection has a destabilising effect, while on the contrary, the negative loading perturbation has a stabilising effect. While the stable characteristic of the bifurcation mode changes due to the loading imperfection, Fig. 11.b. demonstrates that the symmetric characteristic of the bifurcation type does not change in any perturbation of the loading process. The effect of loading variability is analysed in details in other papers of the author [6,7].

However, the symmetric characteristic of the bifurcation mode rapidly changes in the presence of any small geometric perturbation. Let us compare Fig. 11.b. with Fig. 12.a. where a geometric imperfection is assumed. The illustrated equilibrium paths are as follows

$$\lambda(q, \varepsilon_l) = \frac{1}{F_o} \left( \frac{c}{l} \frac{(q - \varepsilon_g)}{\sin q} - \varepsilon_l l (1 - \cos q) \right) \quad (21)$$

where  $\varepsilon_g = 0.2$ . There is no critical state at  $q=0$  any more, and the single imperfection-sensitivity function in Fig. 11.c. is separated to different branches

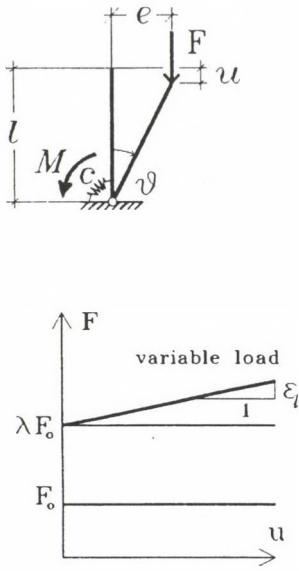
$$\lambda(\varepsilon_l) = \frac{1}{F_o} \left( \frac{c}{l} \frac{\sqrt[3]{\frac{\varepsilon_g}{\varepsilon_l l^2} - \frac{1}{c}} - \varepsilon_g}{\sin\left(\sqrt[3]{\frac{\varepsilon_g}{\varepsilon_l l^2} - \frac{1}{c}}\right)} - \varepsilon_l l \left(1 - \cos\left(\sqrt[3]{\frac{\varepsilon_g}{\varepsilon_l l^2} - \frac{1}{c}}\right)\right) \right) \quad (22)$$

seen in Fig. 12.b. Thus, we can say that a structure with loading imperfection is very sensitive for any geometric perturbation.

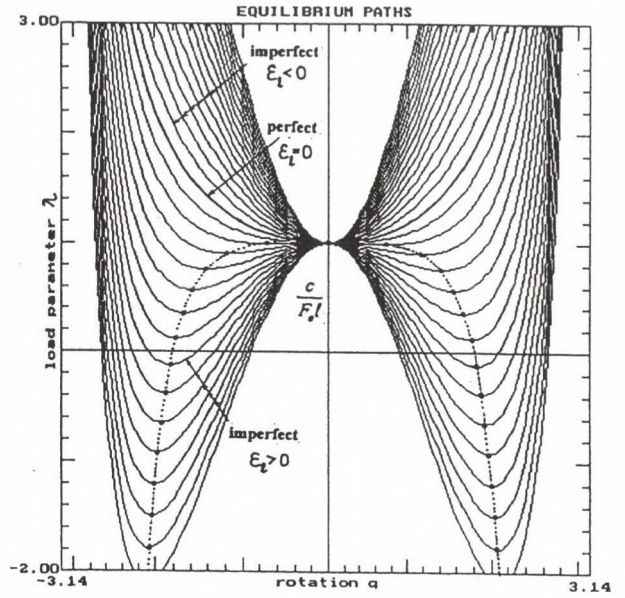
## 5. Conclusion

Geometric, material and loading type imperfection-sensitivities are in the focus of the paper. The classical stable symmetric bifurcation problem modified by bilinear, Hencky-type elastic-plastic material is considered. Global equilibrium paths and the associated imperfection-sensitivity functions are analysed. The presented method is based on the concept of nonsmooth tangent modulus developed by the author for nonlinear materials and loading devices.

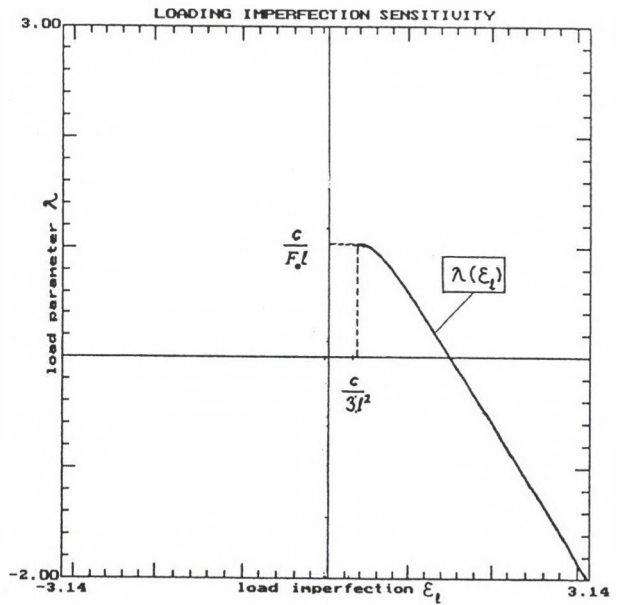
In contrast to the classical geometric *imperfection-sensitivity function* of the elastic structure, the elastic-plastic structure has an *imperfection-sensitivity set* bordered by the imperfection-sensitivity functions of the individual perfectly elastic and perfectly plastic structure.



a.



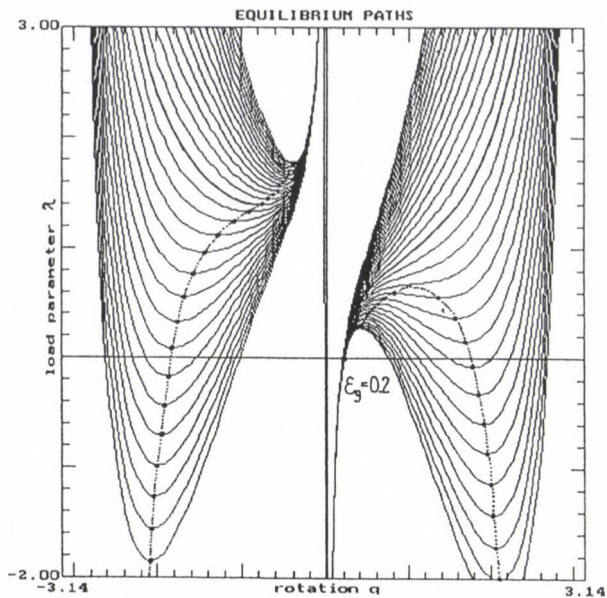
b.



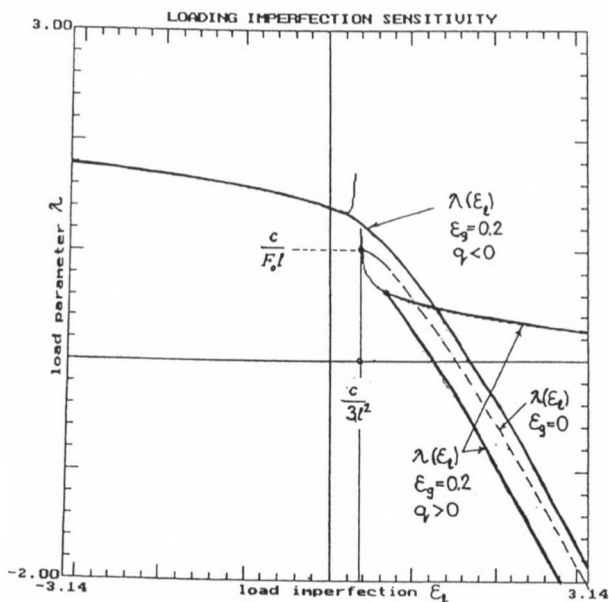
c.

Fig. 11. Loading imperfection-sensitivity





a.



b.

Fig. 12. Loading imperfection-sensitivity with geometric imperfection

Moreover, for imperfections in pair, the domain of possible critical states can be determined which leads to an imperfection-sensitivity set again.

As an important conclusion, evidently, a material softening or a loading hardening have a destabilising effect. On the contrary, material hardening or loading softening has a stabilising effect. This question is analysed in a generally way in [6] based on the modification of the structural tangent stiffness matrix.

#### ACKNOWLEDGEMENT

The present study was supported by the Hungarian Scientific Research Fund (OTKA T-015851)

#### REFERENCES

1. Thompson, J.M.T.-Hunt, G.W.: A General Theory of Elastic Stability, Wiley, London, 1973.
2. Bazant, Z.P.-Cedolin, L.: Stability of Structures. Elastic, Inelastic, Fracture and Damage Theories, Oxford University Press, New York, Oxford, 1991.
3. Kurutz, M.: Stability of structures with nonsmooth nonconvex energy functionals. The one dimensional case, *European Journal of Mechanics, A/Solids*, 12 (1993) 3, 347-385
4. Kurutz, M.: Equilibrium paths of polygonally elastic structures, *Journal of Mechanics of Structures and Machines*, 22 (1994) 2, 181-210
5. Kurutz, M.: Equilibrium paths of polygonally damaging structures, Part I. Part II. *International Journal of Damage Mechanics*, 5 (1996) 1, 16-41. 42-67
6. Kurutz, M.: Modification of the structural tangent stiffness due to nonlinear configuration-dependent conservative loading, *Computer Assisted Mechanics and Engineering Sciences*, 3 (1996) 4, 367-388
7. Kurutz, M.: Postbifurcation equilibrium paths due to nonlinear configuration-dependent conservative loading by using nonsmooth analysis, *Journal of Mechanics of Structures and Machines* (accepted for publication).

## MECHANICAL MODELS OF THE HUMAN SKULL-BRAIN SYSTEM

LOVAS, A.\*

(Received 25 September 1995)

The paper describes the development of a three-dimensional mechanical model of the human head presenting the modeling technique of a living structure from mechanical point of view. Finite Element Method (FEM) is probably the most powerful method for determining the stress patterns within the skull bone. Boundary Element Method (BEM) is a newer technique that can simulate the behavior of the brain. This paper suggests that probably one of the best way to reduce the number of degree of freedom of the complex skull-brain structure required to solve the problem is to use mixed finite element - boundary element solution.

The presented numerical study illustrates the 3D static-dynamic FEM solution of the empty skull and 2D static BEM solution of the linked skull and brain system, in this moment separately. A parametric study was subsequently conducted to identify the model response where material properties of the cranial bone, and stiffness coefficients of the springs, as well as the impact site and the distribution of load were varied.

### 1. Introduction

Biomechanics is the study of the motion and effect forces on mechanical and functional behaviors of biological bodies in order to help the medical physicians diagnose diseases and make remedial decisions. In practical point of view we have to use approximate numerical solutions to get the response of complex biological systems due to static-dynamic loads. Finite Element Method and Boundary Element Method are good tools to take into account the complicated geometric shape and material properties of the living systems.

The human head is an extremely complex living biological system, that has attracted enormous attention because of its susceptibility to injury. Head injury constitutes approximately 50% of all injuries sustained in automobile collisions, in pilot ejections, in sport accidents and other human activities. When the load is a direct impact, energy is transferred to the brain from the skull, a pressure gradient is formed in the brain with positive pressure at the impact site and negative pressure at the contre-coup site. The brain tissues can be lacerated by the compressive or tensile stresses resulting cerebral contusions or intracranial haematoma [3].

Experimental head injury models include those using subhuman primates, human cadavers, manikins and other physical models.

Following the early mathematical models, simple 2D plane strain and layered axisymmetric models (spherical and oval), simulating skull and brain were developed. Ward and Thompson's model [9] was one of the first finite element models that approximated the three dimensional anatomy of the brain, however, the skull was assumed to be rigid. Shugar's model [8] closely simulated the human head geometry, but the model was bilaterally symmetric. Hosey

---

\* Lovas, Antal, H-1029 Budapest, Bercsényi u. 5, Hungary



and Liu's model [4] was a very comprehensive one because it included many elements of the skull-brain-cervical-spine complex including dura folds and cord. Since it has only 922 nodes, it was not feasible to perform a detailed parametric study at that time. The model was run for only 6ms, on the other hand the impact duration was 4ms. Ruan, Khalil and King's finite element model [7] represented the skull, the facial bones and the brain as well, but the spinal cord and neck were not represented.

All models were linear, using mainly elastic homogeneous and isotropic materials. Mechanical properties of cranial-bone have been identified with a reasonable degree of confidence, but by contrast, properties of the living brain are not well known [5]. Usually, the brain was assumed to be nearly incompressible with a bulk modulus. Large discrepancies in skull strains and brain pressures of the different models were not surprising because of the many approximations and assumptions made.

The modeling procedure and selection of material constants our model are described in following.

## 2. Mechanical models of the skull-brain system

Application of closed-form theories is limited to structures of relatively simple geometrical and material properties. FEM is probably the most powerful method for determining the stress patterns within the skull. A FEM model describes the four relevant aspects of the a structure (loading conditions, geometry, material properties, boundary/interface conditions) in discrete, numerical form [10]. From an engineering mechanics point of view, the human skull is a nearly closed shell structure. Therefore, during the finite element discretisation process, it is mathematically divided into a set of layered (represent the outer table, diploe and inner table) or simply nonlayered thin shell elements.

We obtain a final system of equations of the skull model that can be expressed in matrix form:

$$\mathbf{K}^s \mathbf{u}^s = \mathbf{t}^s \quad (1)$$

where  $s$  superscript marks the skull system,  $\mathbf{K}^s$  is the stiffness matrix,  $\mathbf{u}^s$ ,  $\mathbf{t}^s$  are the nodal displacements, the total given consistent nodal forces vectors, respectively.

FEM needs the discretization of the whole brain into a series of block-like elements. Using twenty-node or minimum eight-node hexahedron elements, the finite element model of the brain has some thousands new points in addition of the skull nodes.

The basic idea underlying the BEM is that the governing differential equation of the domain of interest is transformed into an integral equation on the boundary of the domain [1]. Reduction of the dimension of the problem is one of the main features of the BEM. This reduction means that only the boundary and not the whole domain needs to be discretized. The boundary element models are simple to create even in 3D cases, as the mesh is defined only on the external surfaces.

Modeling the skull by FEM means that the geometry data of the connected surface are ready, because neglecting the half thickness of the cranial bone, the finite element nodes of the thin shell problem are covering the closed brain. In this case there are advantages in combining finite and boundary element solution.

Let us treat the boundary element region as one finite element. Consider the two regions of the head where the brain region  $R^b$  expressed in terms of boundary solutions (where  $b$  superscript marks the brain system), and the skull region  $R^s$  discretized into finite elements.

Let us suppose that  $S_1$  and  $S_2$  are the external boundaries of  $R^b$  and  $R^s$ , respectively. The two regions are joined together at the interface  $S_i$ , where  $i$  subscript marks the interface boundary.

The boundary element matrices for  $R^b$  can be written as

$$\mathbf{H}^b \mathbf{u}^b = \mathbf{G}^b \mathbf{p}^b + \mathbf{b}^b, \quad (2)$$

where  $\mathbf{H}^b$  and  $\mathbf{G}^b$  are the so-called influence matrices,  $\mathbf{u}^b$  and  $\mathbf{p}^b$  are the displacements and surface tractions vector, respectively and  $\mathbf{b}^b$  is the term due to the body forces of the brain.

There are known functions, called fundamental solutions, which describe the displacement and traction fields ( $u^*$  and  $t^*$ ). The integrals of the fundamental solutions  $u^*$  and  $t^*$ , multiplied by the interpolation functions give the series of coefficients  $g_{ij}^b$  and  $h_{ij}^b$ , respectively, which will be the elements of  $\mathbf{G}^b$  and  $\mathbf{H}^b$  square matrices. The element in the main diagonal of  $\mathbf{H}^b$  have been increasing by a constant, which is 0.5 on a smooth boundary.

Notice that the values of the tractions  $\mathbf{p}^b$  in Eq. (2) are their actual values at the nodes. In finite elements these values are weighted and concentrated at the nodes. These values are represented by the vector  $\mathbf{t}^s$  of Eq. (1).

Hence we can always find a distribution matrix  $\mathbf{N}^b$  such that  $\mathbf{t}^b$  generally can be written

$$\mathbf{t}^b = \mathbf{N}^b \mathbf{p}^b. \quad (3)$$

In order to combine (1) and (2) one can deduce a matrix that can be easily implemented in the finite element codes. We start by transforming Eq. (2) by inverting  $\mathbf{G}^b$ , i.e.

$$\mathbf{G}^{b-1} \mathbf{H}^b \mathbf{u}^b = \mathbf{p}^b + \mathbf{G}^{b-1} \mathbf{b}^b. \quad (4)$$

Next one can convert the values of tractions at the nodes into an equivalent nodal force vector of the type used in finite elements. So premultiply both sides by the distribution matrix  $\mathbf{N}^b$  equation (4) can now be written as

$$\mathbf{N}^b \mathbf{G}^{b-1} \mathbf{H}^b \mathbf{u}^b = \mathbf{N}^b \mathbf{p}^b + \mathbf{N}^b \mathbf{G}^{b-1} \mathbf{b}^b, \quad (5)$$

where the right-hand side vectors have the same forms as in finite elements.

One can now define

$$\mathbf{K}' = \mathbf{N}^b \mathbf{G}^{b-1} \mathbf{H}^b, \quad (6)$$

$$\mathbf{t}^{b_b} = \mathbf{N}^b \mathbf{G}^{b-1} \mathbf{b}^b. \quad (7)$$

Hence Eq. (5) has the following finite element form

$$\mathbf{K}' \mathbf{u}^b = \mathbf{t}^b + \mathbf{t}^{b_b} = \mathbf{t}^{b_1} + \mathbf{t}^{b_i} + \mathbf{t}^{b_b}, \quad (8)$$

where  $\mathbf{t}^{b_1}$ ,  $\mathbf{t}^{b_i}$ ,  $\mathbf{t}^{b_b}$  are the possible external forces, interface forces and nodal body forces vectors of the brain, respectively and  $\mathbf{K}'$  is a stiffness matrix obtained from the boundary element formulation.

The main disadvantage is the fact that  $\mathbf{K}'$  is generally asymmetric due to the approximations involved in the discretization process and the choice of the assumed solution.



This matrix can be made symmetric by calculating the average of the non symmetric off-diagonal terms, as the asymmetry is usually small especially in this case where there are not corner nodes and extra conditions on the interface as well.

This gives the final matrix of the brain whose coefficients defined by

$$\mathbf{K}^b = 0.5 * (\mathbf{K}' + \mathbf{K}'^T), \quad (9)$$

where  $\mathbf{K}'^T$  is the transpose matrix of  $\mathbf{K}'$  and Eq. (8) can be expressed as

$$\mathbf{K}^b \mathbf{u}^b = \mathbf{t}^b_1 + \mathbf{t}^b_i + \mathbf{t}^b_b. \quad (10)$$

Equation (1) can now be rewritten to the analogue form as

$$\mathbf{K}^s \mathbf{u}^s = \mathbf{t}^s_2 + \mathbf{t}^s_i + \mathbf{t}^s_b, \quad (11)$$

where  $\mathbf{t}^s_2$ ,  $\mathbf{t}^s_i$ ,  $\mathbf{t}^s_b$  are the external forces, interface forces and nodal body forces vectors of the skull, respectively.

In order to join the two regions compatibility and equilibrium on  $S_i$  need to be satisfied, i.e.

$$\mathbf{u}^b_i = \mathbf{u}^s_i \quad \text{on } S_i, \quad (12)$$

$$\mathbf{t}^b_i + \mathbf{t}^s_i = \mathbf{0} \quad \text{on } S_i. \quad (13)$$

One can finally write the global equilibrium equation in matrix form

$$\mathbf{K} \mathbf{u} = \mathbf{t} \quad (14)$$

where

$$\mathbf{K} = \mathbf{K}^s + \mathbf{K}^b, \quad (15)$$

$$\mathbf{u} = \mathbf{u}^s = \mathbf{u}^b, \quad (16)$$

$$\mathbf{t} = \mathbf{t}^e + \mathbf{t}^s_b + \mathbf{t}^b_b, \quad (17)$$

$$\mathbf{t}^e = \mathbf{t}^s_2 + \mathbf{t}^b_1, \quad (18)$$

where  $\mathbf{t}^e$  is the vector of the external nodal forces loaded the head,  $\mathbf{t}^b_1$  is usually a zero vector.

Notices that the total number of degree of freedom of the skull-brain system remains the degree of freedom of the skull, although this method requires the inversion of the nonbanded  $\mathbf{G}^b$  matrix and  $\mathbf{K}$  is nonbanded too. The accuracy of the boundary element solution using quadratic elements is excellent even for the coarse discretisation. Using coarse mesh for the brain, i.e. neglecting the odd nodes of the skull model one can reduce the dimension of  $\mathbf{G}^b$ .

In the case of the head injury the role of the impact load is more important than the statical one, so the main goal is to expand the above method.

Similar matrices obtained for dynamic problem. The dynamic matrix equation of the skull

$$\mathbf{M}^s \ddot{\mathbf{u}}^s + \mathbf{K}^s \mathbf{u}^s = \mathbf{t}^s \quad (19)$$

where  $\mathbf{M}^s$  is the mass matrix of the skull,  $\ddot{\mathbf{u}}^s$  is the acceleration vector.

In the case of the brain using boundary element approximation

$$\mathbf{B}^b \ddot{\mathbf{u}}^b + \mathbf{H}^b \mathbf{u}^b = \mathbf{G}^b \mathbf{p}^b, \quad (20)$$

$$\mathbf{B}^b = \rho^b [ \mathbf{G}^b \mathbf{q}^b - \mathbf{H}^b \mathbf{v}^b ] \mathbf{F}^{b-1}, \quad (21)$$



where  $\rho^b$  is the density of the brain,  $\ddot{u}^b$  is the acceleration vector,  $v^b$  and  $q^b$  are hipermatrices contain the values of the displacements and tractions due to the body forces field, the element of  $F^b$  matrix are the values of the body forces functions at the boundary nodes giving the relationships between the accelerations and the body surface functions [2].

Eq. (20) has the same aspect as the finite element equilibrium equation.

Taking into account Eq. (3) - Eq. (9), can also be written similar as  $K'$

$$M' = N^b G^{b-1} B^b, \quad (22)$$

$$M^b = 0.5 * (M' + M'^T), \quad (23)$$

and Eq. (10) can be expressed as

$$M^b \ddot{u}^b + K^b u^b = t^b_1 + t^b_i. \quad (24)$$

Taking into account Eq. (11) - Eq. (18), the global dynamic matrix equation combining Eq. (19) and Eq. (24) in matrix form

$$M \ddot{u} + K u = t^e, \quad (25)$$

where

$$M = M^s + M^b, \quad (26)$$

$$\ddot{u} = \ddot{u}^s = \ddot{u}^b. \quad (27)$$

Finally, if Rayleigh damping is include Eq. (27) will become

$$M \ddot{u} + (\alpha K + \beta M) u + K u = t^e, \quad (28)$$

where  $\alpha$  and  $\beta$  are the Rayleigh damping coefficients.

### 3. Numerical results

A case study is now presented to illustrate the FEM and BEM solution separately.

#### Example No. 1

In the FEM solution the eight-node thin shell nonlayered elements represent the skull neglecting the facial bones. Data for the coordinates and the thicknesses of the anatomical points were based on 21 CT pictures created in the OTKA project. The empty skull finite element model consists 264 elements and 722 nodes with 229 given thickness data (Fig. 1).

The cranial bone was assumed to be linear elastic, homogeneous and isotrop. The elastic moduli was chosen  $E=4460\text{MPa}$ , Poisson ratio  $\nu=0.21$  [5].

The head-neck junction was presented by 3 translational and 3 rotational spring elements. The stiffness data for the springs were more difficult to obtain [6]. The chosen translational coefficients are  $\rho_{tx}=750\text{N/cm}$ ,  $\rho_{ty}=750\text{N/cm}$  and  $\rho_{tz}=5000\text{N/cm}$ , the chosen rotational coefficients are  $\rho_{rx}=40000\text{Nrad/cm}$ ,  $\rho_{ry}=35000\text{Nrad/cm}$  and  $\rho_{rz}=60000\text{Nrad/cm}$ .

The mass of the skull was calculated using the geometry and the bone density ( $1.41\text{g/mm}^3$ ). The mass of the brain was added to the nodal skull masses.

A parametric study was conducted the material properties of the cranial bone, and the stiffness coefficients of the springs to identify the model response where the impact site (frontal, side, occipital), the distribution of load (concentrated, distributed on 4 elements, distributed on 8 elements) were varied.

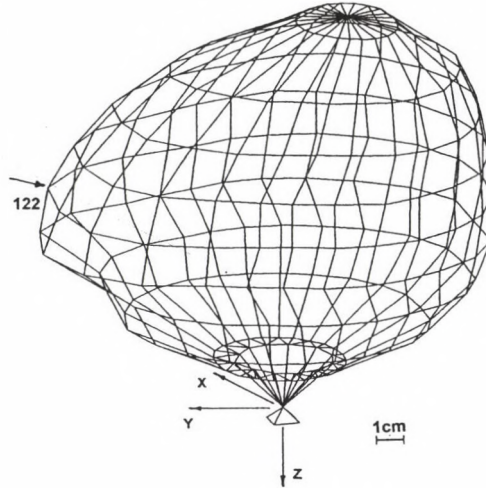


Fig. 1. Skull finite element model

The Young's modulus of the skull bone was varied 1/10 and 10 times of its assumed values (briefly 0.1E, E, 10E). The spring coefficients were varied 10 times and 100 times of there assumed values (briefly r, 10r, 100r). The load-displacement responses were determined. The load was applied on the frontal bone in the mid saggital plane. Fig. 2 shows the effectiveness of these parameters in the resultant displacement [cm] of the node 122 where the load was distributed on 4 elements. We note that the latter is more important because the restraining effect of the neck can vary in a wide range. By keeping the Young's modulus and raising r by a factor of 10 the above displacement will be more than five times smaller.

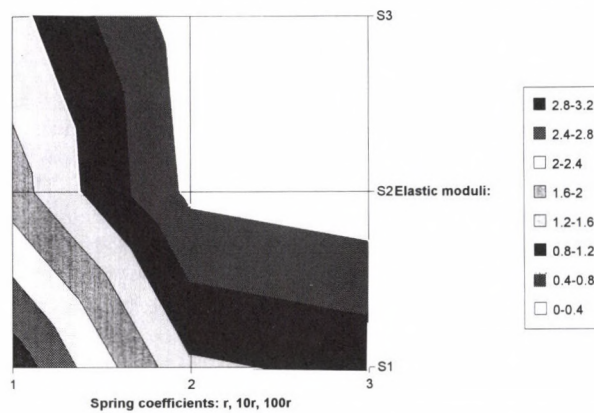
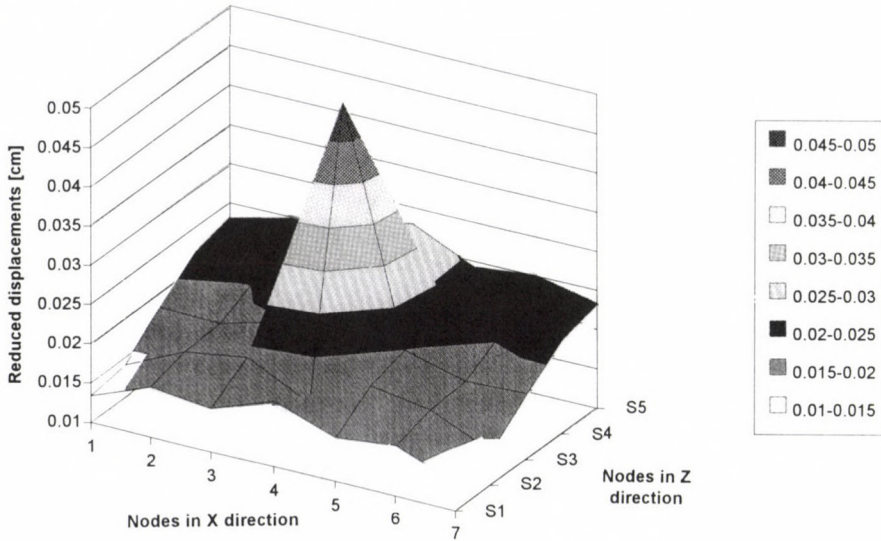
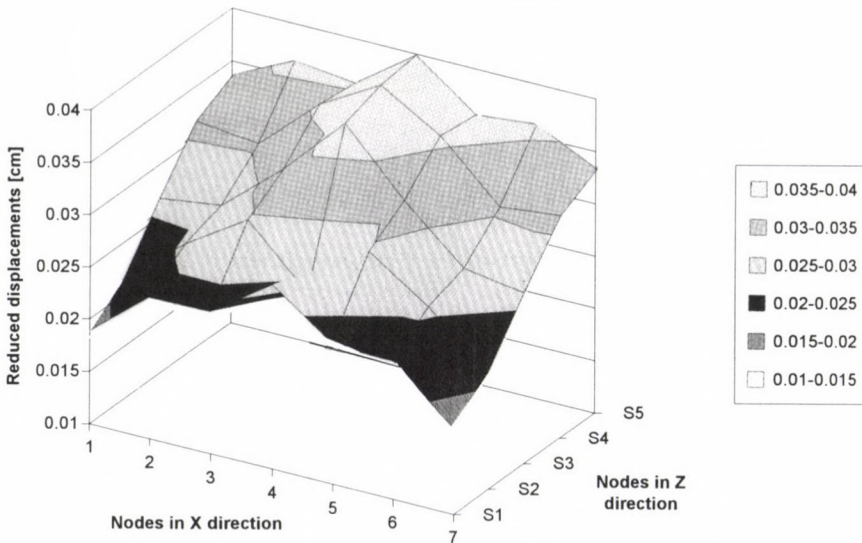


Fig. 2. Resultant displacement of the node 122

Fig 3a and 3b show the displacement fields at the environment of the node 122 (7 and 5 nodes in the directions X and Z, respectively) reduced by the rigid body movements when the load was concentrated and distributed on 4 elements. When the load was concentrated the peak value of the displacement is 30% bigger than the case of the distributed load.



**Fig. 3a.** Displacements due to the concentrated load



**Fig. 3b.** Displacements due to the distributed load



In this study the resonant frequencies from the 1st to the 15th mode were extracted. There were calculated 9 different supposed structures by varying the Young's modulus (0.1E, E and 10E) and the spring coefficients (r, 10r, 100r). Fig. 4 shows the influence of the material properties and spring coefficients on the frequencies where S1: 0.1E, r; S2: 0.1E, 10r; S3: 0.1E, 100r; S4: E, r; S5: E, 10r; S6: E, 100r; S7: 10E, r; S8: 10E, 10r; S9: 10E, 100r. The frequencies from 1 to 6 belong on the rigid body movements, therefore there is a significant gap between the 6th and 7th frequencies. The higher frequencies were not responsive to the spring coefficients. They were varied considerable as Young's modulus varied, but the Young's modulus of the cranial bone is quite well defined.

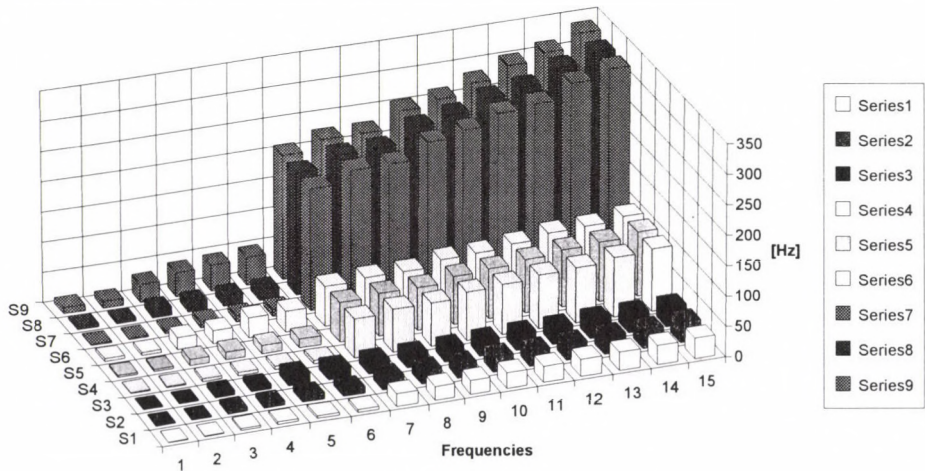


Fig. 4. Frequencies of the FEM model

### Example No. 2

In the first BEM solution 2D plane strain skull model was used on the basis of the material and spring's data of the FEM solution. The BEM model consists 97 nodes and 106 quadratic boundary elements. Fig. 5 shows the BEM model with the supporting and loading systems.

Fig. 6 shows the Mises stress contours at the environment of the applied distributed load.

### Example No. 3

The 2D plane strain BEM model of the brain consists 50 nodes and 49 quadratic elements. The elastic moduli was chosen  $E=0.0667$  MPa. The brain is slightly compressible with a Poisson's ratio ranging from  $\nu=0.48$  to  $\nu=0.499$  [4], [9]. The kinematic load was calculated from the previous solution of the 2D skull model (Fig. 7).

Fig. 8 shows the Mises stress contours in the brain.

The contours of the minimum principal stresses are shown Fig. 9a and Fig. 9b where the Poisson's ratios were 0.49 and 0.48, respectively. The figures show that a quite small difference of the Poisson's ratio affected the brain pressure ranges and the peak value of the stresses more significantly.

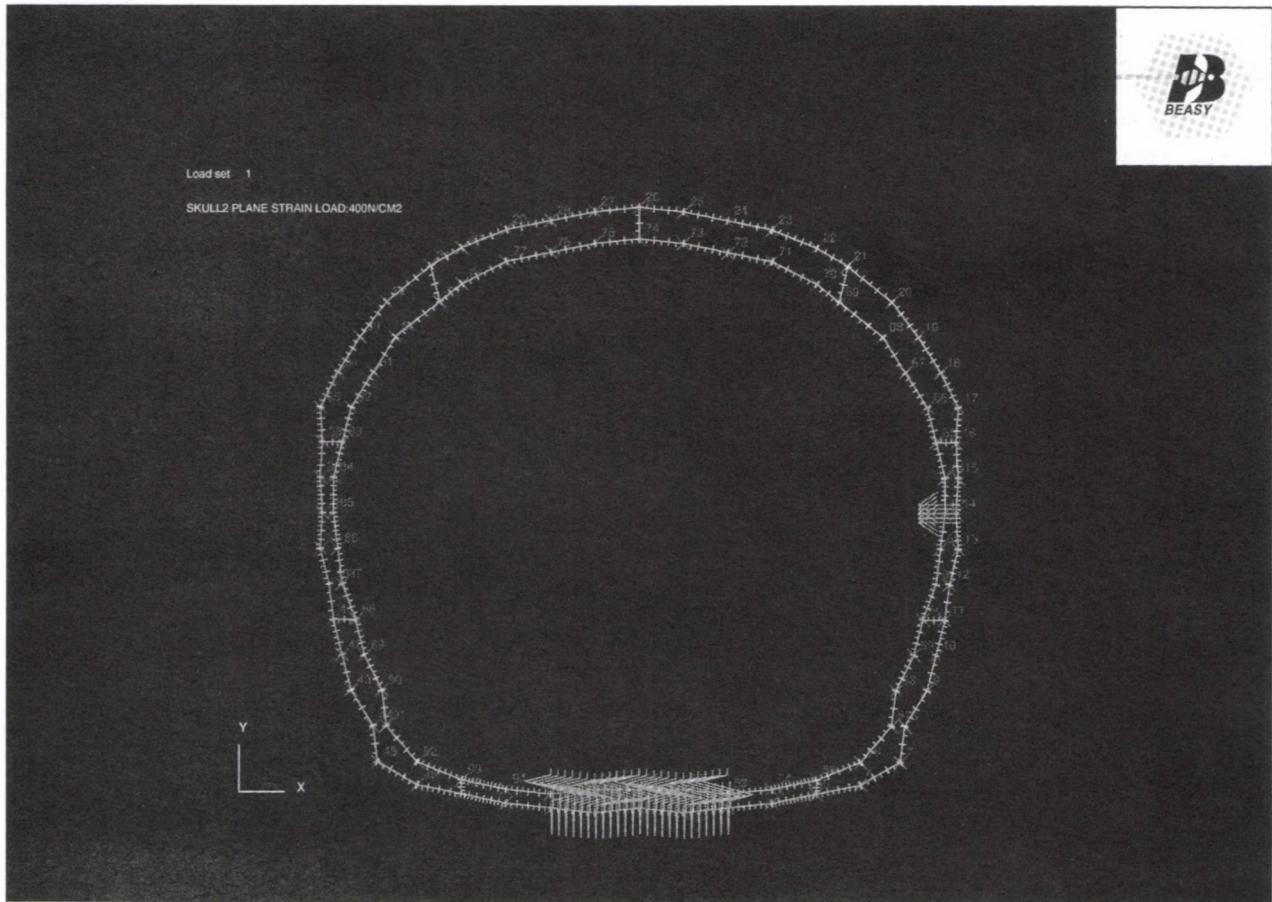


Fig. 5. Skull 2D boundary element model



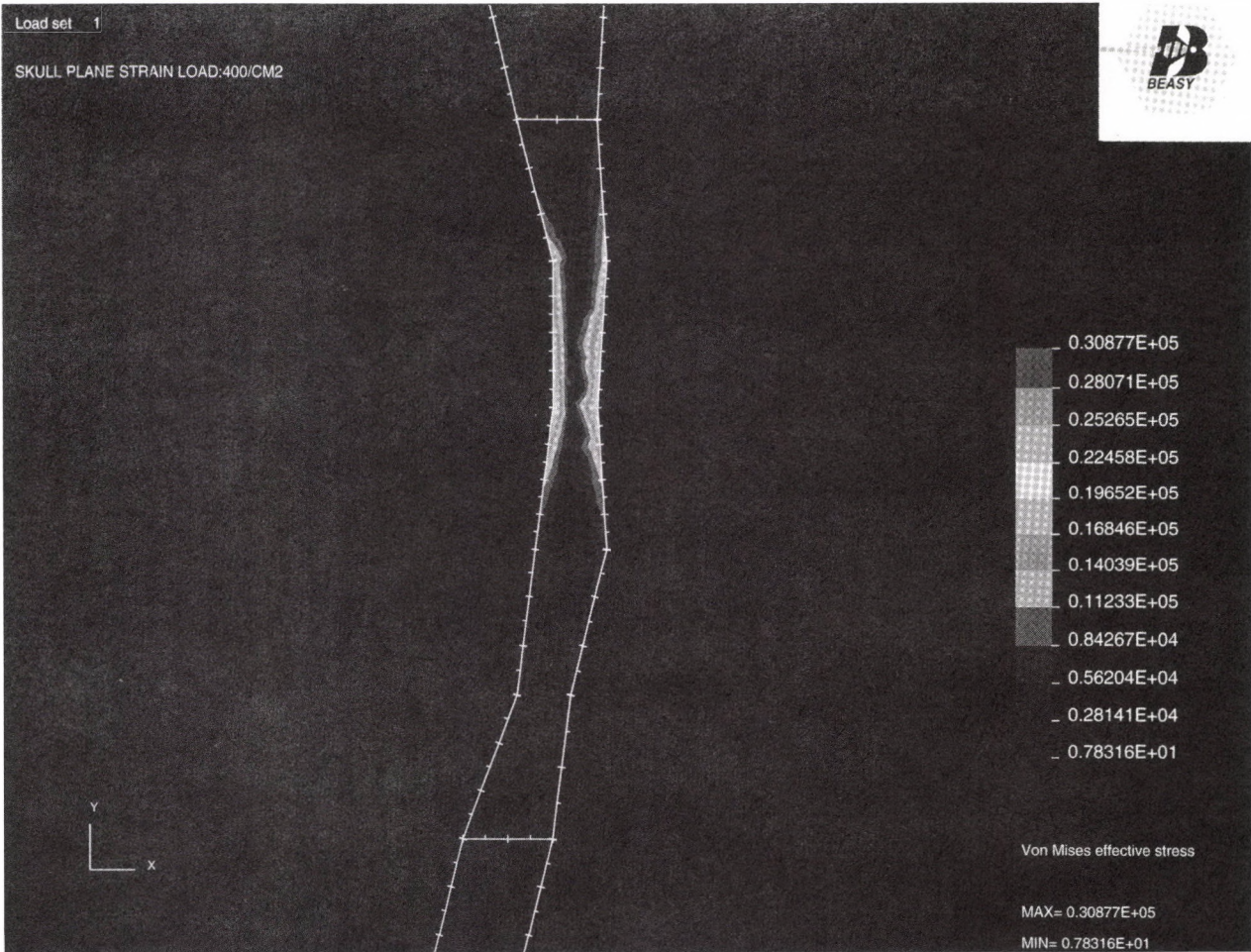


Fig. 6. Mises stress contours



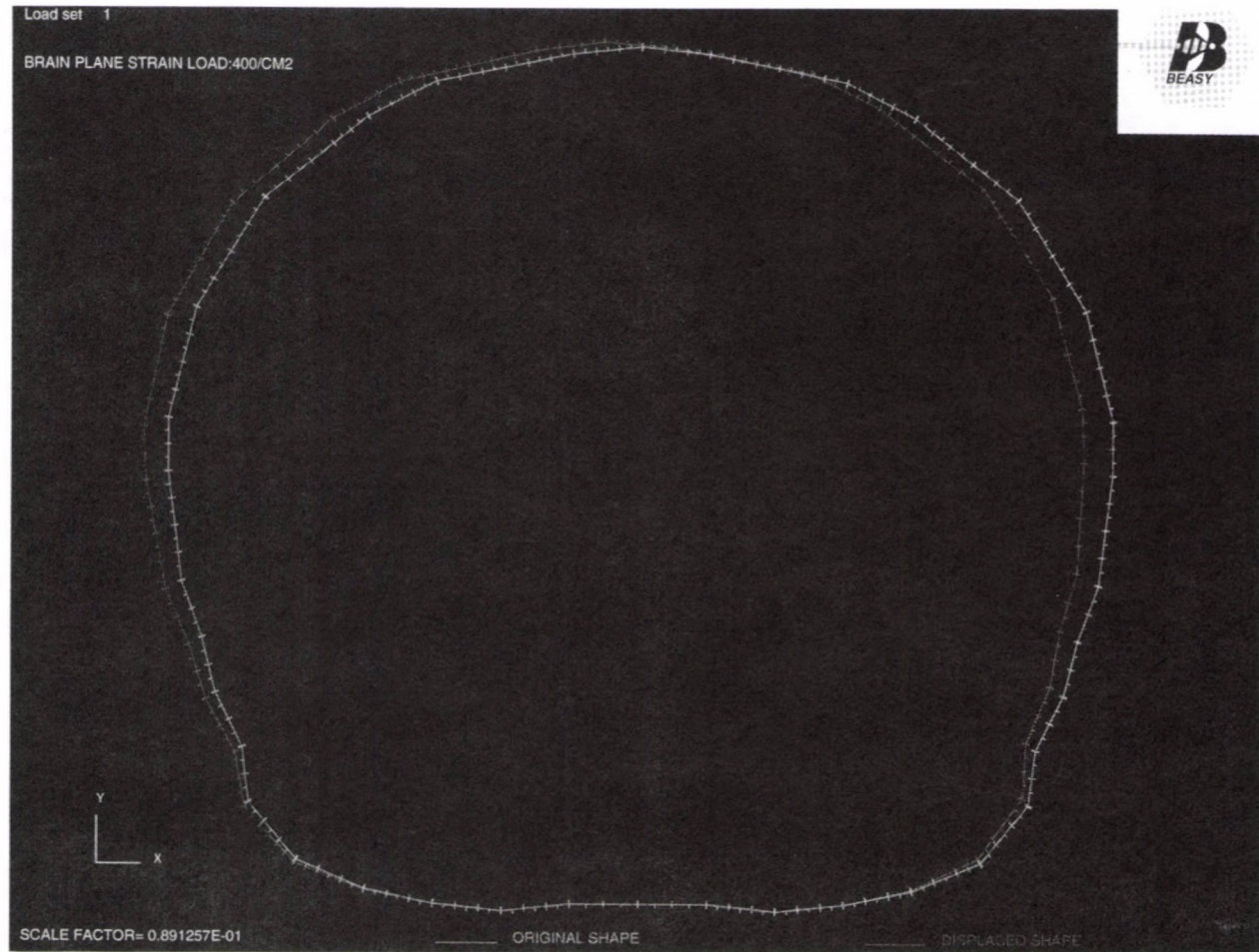


Fig. 7. Brain 2D boundary element model with the initial deformation of the brain

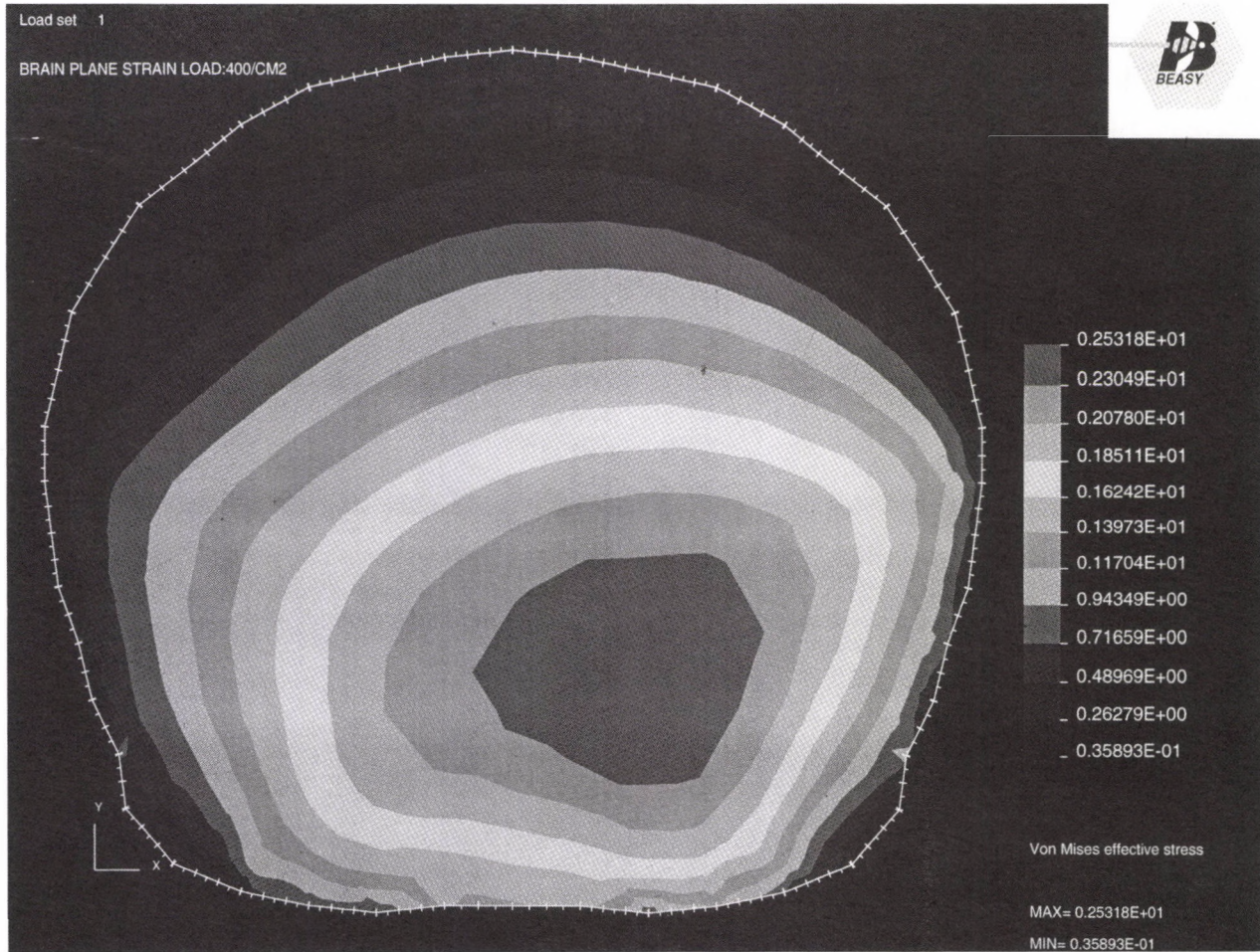


Fig. 8. Mises stress contours in the brain



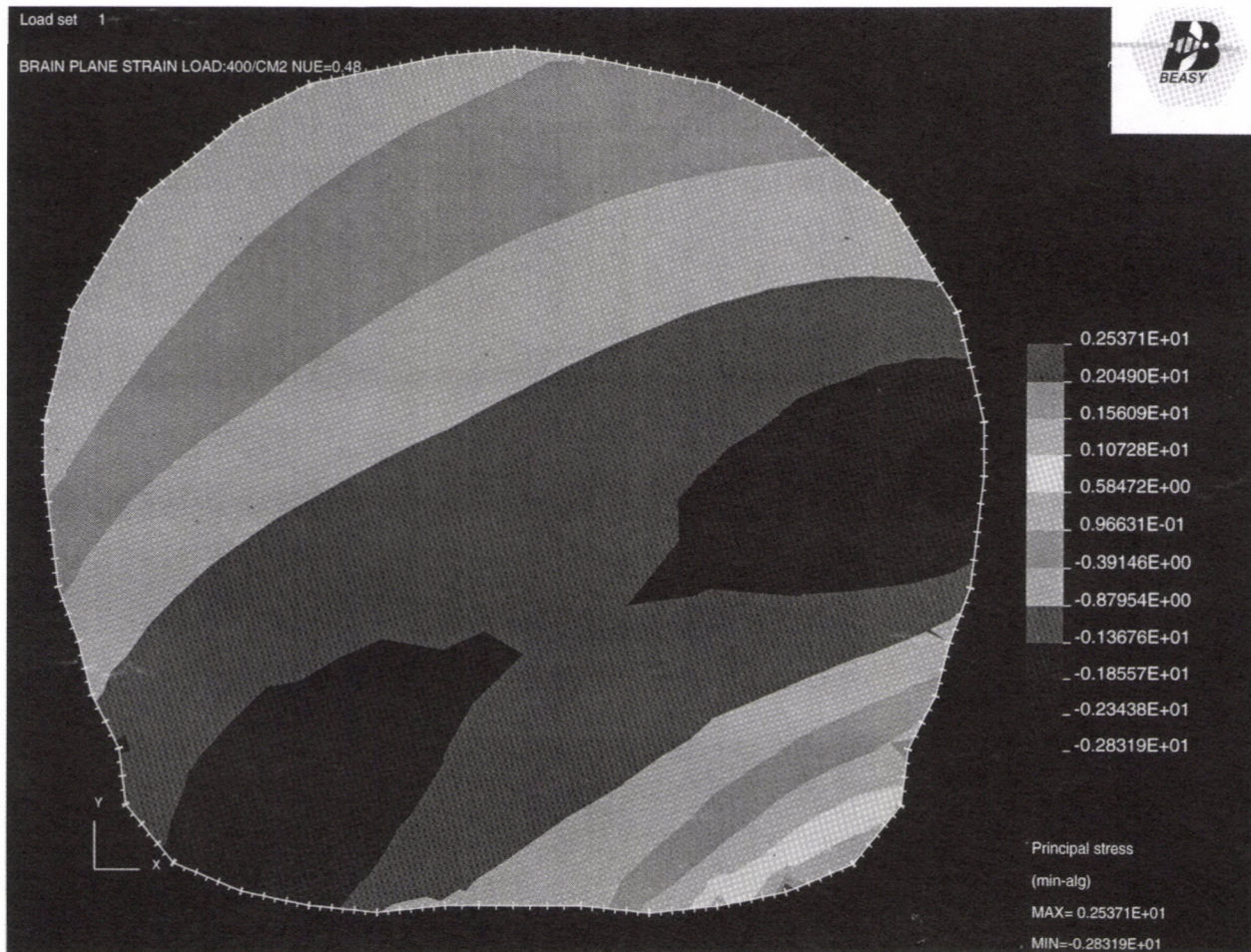


Fig. 9a. Minimum principal stresses due to  $\nu=0.48$



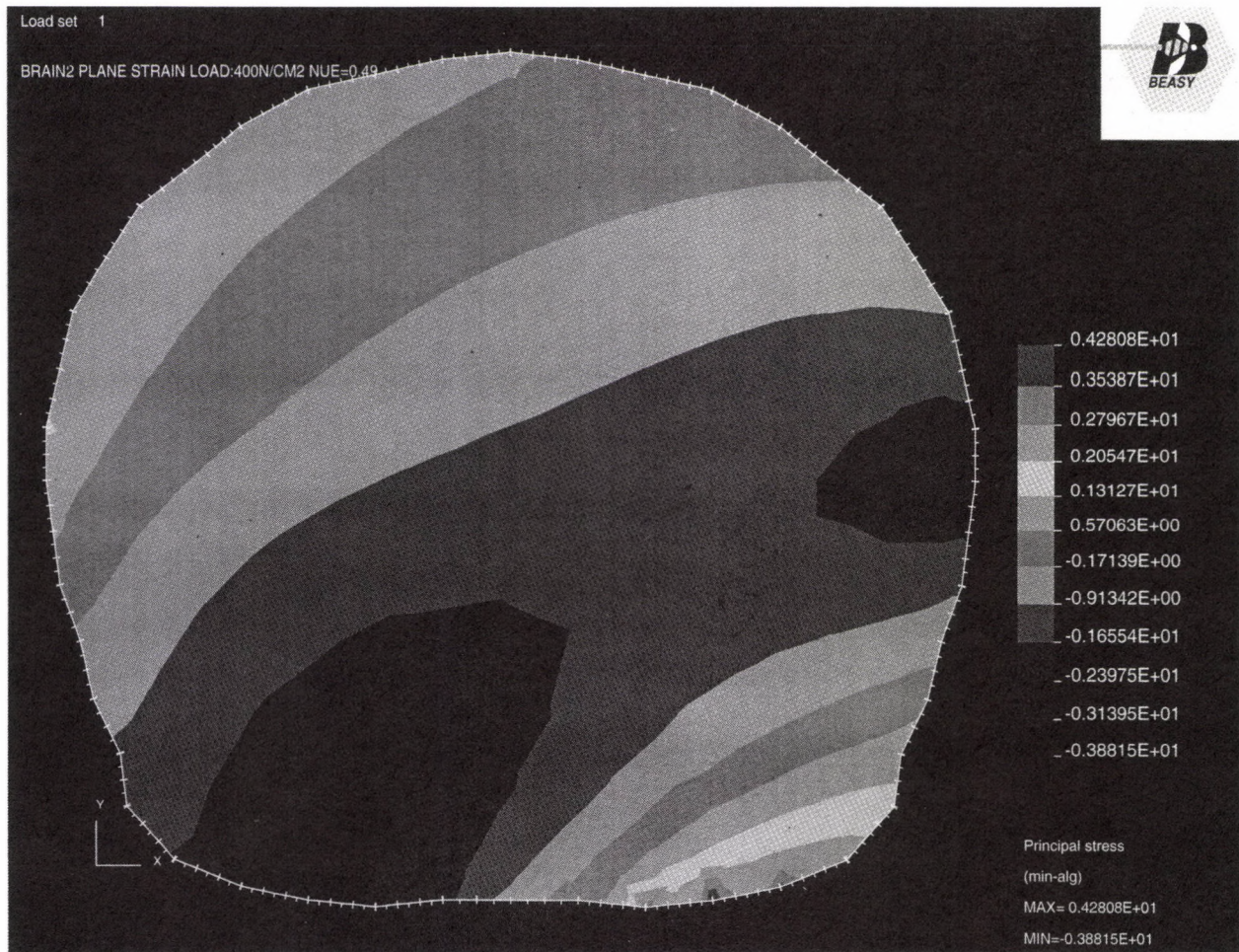


Fig. 9b. Minimum principal stresses due to  $\nu=0.49$

Fig. 10 shows S1: maximum principal stresses, S2: minimum principal stresses, S3: Mises effective stresses variation in the internal points along the axis of the load approaching to the place of the load.

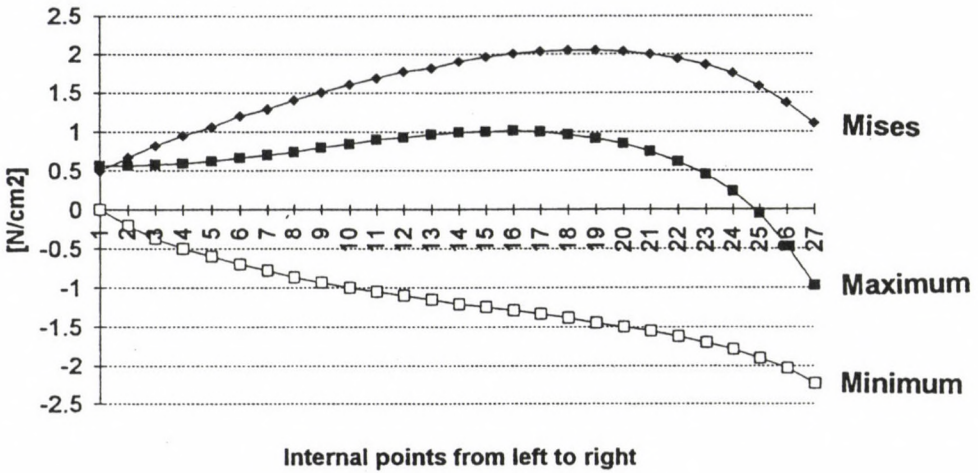


Fig. 10. Brain stresses along the horizontal axis

#### 4. Conclusion

Finally, we note that the results reported here are preliminary and restricted by the assumptions of empty skull using non layered thin shell elements in the case of the FEM solutions and 2D plane strain model was used in the BEM solutions, but we are planning to remove each of these restrictions in the future.

Results of the first example pointed that the spring coefficients are the critical parameters of the input data, especially in the case of dynamical problems. The 2nd and the linked 3rd examples justified the efficiency of the BEM in the biomechanical problems. Due to the 2D plane strain solution the results are limited, but we can expand the analysis by removing this restriction, in particular in the case of the brain model.

#### ACKNOWLEDGEMENTS

This research was supported (in part) by the Hungarian Scientific Research Fund (OTKA I/3-201). The author thanks the colleagues of the research group for the CT pictures to create the geometrical data and for using the developed computer program. This research was also supported (in part) by the Royal Society research grant (UK). The author thanks the colleagues of the Wessex Institute of Technology and BEASY group for using the BEASY computer program.



## REFERENCES

1. Brebbia, C. A.-Telles, J. C. F.-Wrobel, L. C.: *Boundary Element Techniques, Theory and Application in Engineering*. Springer-Verlag, Berlin, New York 1984
2. Dominguez, J.: *Boundary Elements in Dynamics*. Computational Mechanics Publications, Southampton, Boston 1993
3. Goldsmith, W.: *Biomechanics of Head Injury, Biomechanics - Its Foundation and Objectives* (ed. by Fung, Y. C.-Perrone, N.-Anliker, M.). Prentice-Hall Inc., Englewood Cliffs, New York (1972) 585-634
4. Hosey, R. R.-Liu, Y. K.: *A Homeomorphic Finite Element Model of the Human Head and Neck, Finite Elements in Biomechanics* (ed. by Gallagher, R. H.-Simon, B. R.-Johnson, P. C.-Gross, J. F.). John Wiley, (1982) 379-400
5. McElhaney, J. H.-Melvin, J. W.-Roberts, V. L.-Portnoy, H.: *Dynamic Characteristics of the Tissue of the Head, Perspectives in Biomedical Engineering* (ed. by Kenedi, R. M.), MacMillan Press, London (1973) 215-222
6. Merril, T.-Goldsmith, W.-Deng, Y. C.: *Three-Dimensional Response of Lumped Parameter Head-Neck Model Due to Impact and Impulsive Loading, Journal of Biomechanics*, 17 (2), (1984) 81-95
7. Ruan, J. S.-Khalil, T.-King, A. I. (1994): *Dynamic Response of the Human Head to Impact by Three-Dimensional Finite Element Analysis, Journal of Biomechanical Engineering*, 116, (1994) 44-50
8. Shugar, T. A.-Katona, M. G.: *Development of Finite Element of Head Injury Model, Journal ASCE*, 101, (1975) 223-239
9. Ward, C. C.-Thompson, R. B.: *The Development of a Detailed Finite Element Brain Model. In 19th Stapp Car Crash Conference Proceedings, 17-19 November 1975, San Diego, Calif., Warrendale, Pa.: Soc. Auto. Eng., (1975) 641-674*
10. Zienkiewicz, O. C.: *The Finite Element Method*. 3rd ed., McGraw-Hill, New York 1977



## ELASTIC-PLASTIC ANALYSIS OF FRAMES IN CASE GEOMETRICAL NONLINEARITY

NÉDLI, P. \*

(Received: 29 August 1996)

The one parameter load history analysis of elastic-plastic frames is studied taking into account geometrical nonlinearity. It is shown that using a discrete structural model, this analysis can be formulated as a series of nonlinear programming problems where each problem corresponds to the determination of the next active yield location(s). By using as objective function the work of the base load on the displacements, not only the limit point of the load displacement diagram can be determined but also the postcritical path can be followed. The solution of each nonlinear programming problem is used as the initial feasible solution of the next one. The problems were solved by the MINOS computer code.

### 1. Introduction

Rigid plastic limit analysis usually gives an optimistic estimate of the loadbearing capacity of a frame. Especially in case of sway frames, the displacements which develop during the load history influence unfavorably the loadbearing capacity of the structure. It is also frequent that the limit point of the equilibrium path is reached before a yield mechanism is formed. Therefore it is necessary to take into account the influence of geometrical nonlinearity to obtain a more realistic estimate of the loadbearing capacity of the structure.

In the sequel, the one parameter load history of elastic plastic frames will be studied. A discrete structural model is used for the analysis. The state of the structure is described by the following finite dimensional vectors:  $\mathbf{q}$  (nodal loads),  $\mathbf{t}$  (generalized initial strains),  $\mathbf{v}$  (nodal displacements),  $\mathbf{s}$  (generalized stresses),  $\boldsymbol{\varphi}$  (plastic potentials) and  $\boldsymbol{\lambda}$  (plastic multipliers). Plastic flow is restricted to critical sections which are taken at extremities of the members. For yield condition the classical plastic hinge approach is used (Fig. 1).

$$\begin{bmatrix} \varphi^+ \\ \varphi^- \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} N \\ T \\ M \end{bmatrix} - \begin{bmatrix} M_p^+ \\ M_p^- \end{bmatrix} \leq \mathbf{0}$$

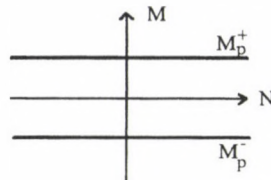


Fig. 1.

For a general  $K$ -th critical section, the yield condition can be written in the following condensed form:  $\boldsymbol{\varphi}_K = \mathbf{N}_K \mathbf{s}_K - \mathbf{k}_K \leq \mathbf{0}$ , where  $\mathbf{N}_K$  denotes the matrix in Figure 1;  $\mathbf{s}_K$  is the vector of internal forces  $N, T, M$ ;  $\mathbf{k}_K$  is the vector of plastic limit moments and index  $K$  refers to the section. Similarly, the yield condition can be written for a member and also for the whole structure in an analogous form:  $\boldsymbol{\varphi} = \mathbf{N} \mathbf{s} - \mathbf{k} \leq \mathbf{0}$ , where matrix  $\mathbf{N}$  has a bloc diagonal structure. The form of the flow rules associated with the given yield condition considering the whole structure is the following:

\*Nédli, Péter, H-1118 Budapest, Homoród utca 20, Hungary

$I = \{i \mid \varphi_i = 0\}$ ,  $\dot{\mathbf{p}} = \mathbf{N}_I^T \dot{\boldsymbol{\lambda}}_I$ ,  $\dot{\boldsymbol{\lambda}}_I \geq \mathbf{0}$ ,  $\dot{\boldsymbol{\phi}}_I^T \dot{\boldsymbol{\lambda}}_I = 0$ ,  $\dot{\boldsymbol{\phi}}_I \leq \mathbf{0}$ . Here  $I$  designates the set of the indices of those elements of vector  $\boldsymbol{\varphi}$  which value is 0 (these are the active yield locations). "I" used as index refers to the appropriate subvector or submatrix. Vector  $\dot{\mathbf{p}}$  stands for the generalized plastic strain rates.

In first order theory i.e. in a geometrically linear elastic-plastic problem, the one parameter load history analysis can be formulated as a parametric linear complementarity problem.

## 2. First order elastic-plastic analysis

Several important algorithmic elements of the geometrically linear elastic-plastic analysis can be used in the geometrically nonlinear case as well, so it seems useful to make a brief review of first order governing relations.

Due to the nature of the elastic-plastic constitutive law at a given state of the load history, the governing relations are formulated for the rate (or increments) of the variables (2.1):

$$\begin{bmatrix} 0 & \mathbf{G} & 0 & 0 \\ \mathbf{G}^T & \mathbf{F} & \mathbf{N}_I^T & 0 \\ 0 & \mathbf{N}_I & 0 & -\mathbf{E}_I \end{bmatrix} \begin{bmatrix} \dot{\mathbf{v}} \\ \dot{\mathbf{s}} \\ \dot{\boldsymbol{\lambda}}_I \\ \dot{\boldsymbol{\phi}}_I \end{bmatrix} + \begin{bmatrix} \dot{\mathbf{q}} \\ \dot{\mathbf{t}} \\ 0 \end{bmatrix} = \mathbf{0}, \quad \dot{\boldsymbol{\phi}}_I \leq \mathbf{0}, \quad \dot{\boldsymbol{\lambda}}_I \geq \mathbf{0}, \quad \dot{\boldsymbol{\phi}}_I^T \dot{\boldsymbol{\lambda}}_I = 0. \quad (2.1)$$

The first equation is the equilibrium equation, the second one is the compatibility equation and the third one concerns the change of the plastic potential.  $\mathbf{G}$  is the equilibrium matrix,  $\mathbf{F}$  is the flexibility matrix of the structure and  $\mathbf{E}_I$  is an appropriate unit matrix. From the mathematical point of view (2.1) is a linear complementarity problem which can be solved by an algorithm based on the simplex method. Due to the discrete structural model and the linearized yield condition, as far as no new yielding happens and the load rate  $(\dot{\mathbf{q}}, \dot{\mathbf{t}})$  is constant which is the case in one parameter loading, relation (2.1) is unchanged and the solution is constant as well. This fact makes possible another equivalent formulation in which instead of the rate of the variables, the values of the variables belonging to the next yield location are used. This formulation has advantages and disadvantages as well. Its main advantage is that it can be extended more easily to the geometrically nonlinear case than the rate formulation as the rates are not constant if geometrical nonlinearity is involved. Its disadvantage is that the size of the governing relations is increased. This formulation can be summarized as follows.

Let us consider a known state where the equilibrium, compatibility and yield conditions are satisfied. The values of the variables belonging to this state are indexed by 'o'. Assuming a one parameter loading, the goal is to determine the values of the variables at the next yield location. This problem is described by the following relations:

$$\begin{aligned} \text{a) } & \begin{bmatrix} 0 & \mathbf{G} & 0 & 0 & \mathbf{q}_b \\ \mathbf{G}^T & \mathbf{F} & \mathbf{N}^T & 0 & \mathbf{t}_b \\ 0 & \mathbf{N} & 0 & -\mathbf{E} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{v} \\ \mathbf{s} \\ \boldsymbol{\lambda} \\ \boldsymbol{\varphi} \\ \mathbf{m} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \mathbf{k} \end{bmatrix}; \\ \text{b) } & I_o: \{i \mid \varphi_{oi} = 0\}, \quad \bar{I}_o: \{i \mid \varphi_{oi} < 0\}; \\ \text{c) } & \boldsymbol{\lambda}_I \geq \boldsymbol{\lambda}_{oI}, \quad \boldsymbol{\lambda}_{\bar{I}} = \boldsymbol{\lambda}_{o\bar{I}}, \\ & \boldsymbol{\varphi} \leq \mathbf{0}, \quad \boldsymbol{\varphi}^T (\boldsymbol{\lambda} - \boldsymbol{\lambda}_o) = 0; \\ \text{d) } & \mathbf{m} = \max! \end{aligned} \quad (2.2)$$



Here  $q_b$ ,  $t_b$  denote the so-called base load, so the actual load on the structure is the base load multiplied by 'm' which is the load parameter. System (2.2) can be solved also by an algorithm based on the simplex method. From the new state, the analysis can be continued. If the load parameter does not increase anymore, the limit state is reached.

### 3. Higher order elastic-plastic analysis

Considering large displacements but only small strains, when geometrical nonlinearity is taken into account in full precision, we speak about third order theory in case of frames. The discrete nature of the problem can be kept in this case, too and the state of the structure can be characterized by vectors of the same size as in the first order theory. For the third order analysis, the approach elaborated in [2] for elastic frames is used in this study. Its main features can be summarized as follows. The local coordinate system of a member is fixed to its start node and moves with it. At the initial configuration, the direction of each local coordinate system coincides with the global coordinate system. The generalized stresses and strains are defined in the local coordinate systems of the members. The generalized stresses are the internal forces at the starting cross-section of the member, the generalized strains are the relative displacements between the end cross-section and the end node of the member. The reason for this definition is that in this case the determination of the deformed shape of a member can be obtained by solving a system of nonlinear first order differential equations, a so called initial value problem. In the general case, this initial value problem can be solved only numerically, for example by a Runge-Kutta type method suitable for systems of differential equations.

The governing relations describing the determination of the next yielding starting from a known state differ from (2.2) in the following. The system of linear equations in 2.2/a is replaced by a system of nonlinear equations. In order to be able to follow the load history on the postcritical path, the objective function is changed from the load parameter to the work of the nodal base load on the nodal displacements:

$$\begin{array}{ll}
 \mathbf{e}(\mathbf{v}, \mathbf{s}, \mathbf{m}) = \mathbf{0}, & \text{b) } I_o: \{i | \varphi_{oi} = 0\}, \bar{I}_o: \{i | \varphi_{oi} < 0\}; \\
 \text{a) } \mathbf{c}(\mathbf{v}, \mathbf{s}, \boldsymbol{\lambda}, \mathbf{m}) = \mathbf{0}, & \text{c) } \lambda_{i_e} \geq \lambda_{o_{i_e}}, \lambda_{i_c} = \lambda_{o_{i_c}}, \\
 \mathbf{f}(\mathbf{v}, \mathbf{s}, \boldsymbol{\lambda}, \varphi) = \mathbf{k}; & \varphi \leq \mathbf{0}, \varphi^T (\boldsymbol{\lambda} - \boldsymbol{\lambda}_o) = 0; \\
 & \text{d) } \mathbf{q}_b^T \mathbf{v} = \max!
 \end{array} \quad (3.1)$$

Here  $\mathbf{e}$ ,  $\mathbf{c}$ ,  $\mathbf{f}$  mean nonlinear vector-vector functions which determine the lefthand side of the equations. The equilibrium equation and the yield condition can be written in a closed form at a given value of the variables ( $\mathbf{v}$ ,  $\mathbf{s}$ ,  $\boldsymbol{\lambda}$ ,  $\varphi$ ,  $\mathbf{m}$ ), but the compatibility equation cannot. It can be evaluated only by solving numerically the initial value problem for each member. The complete system gives a nonlinear programming problem.

It should be mentioned that this formulation has of course its limitations. The possibility of unloading in the plastic hinges is excluded on the path between two consecutive yield configurations. Bifurcation of the equilibrium path cannot be detected by this formulation. Relation 3.1 d) is a heuristic one which can be considered as kind of displacement control as it expresses the experience that the work of the base load on the displacements is monotonically increasing during the load history. It needs further study to determine that in which class of problems can these limitations be justified.

The use of the third order theory is very expensive computationally, because the evaluation of the compatibility equation demands the numerical solution of an initial value problem for each member. When within one member the displacements are small, a good approximation can be obtained by the second order theory. This condition can be satisfied by



subdividing the structural elements into an appropriate number of members. Amongst the different second order approximations, the one used in this study is described in [3] (pp. 95-116) and based on the solution of the boundary value differential equation of a beam loaded axially, too. By using the so-called stability functions which are derived from the solution of the differential equation, the stiffness and the flexibility matrix of a member can be assembled. This way, the equilibrium and compatibility equations as well as the yield condition can be written in the form of matrix equations. The matrices in these equations are not constant as in the case of the first order theory but depend on the unknown variables. The mathematical programming problem to determine the next yield location is given by the following relations:

$$\begin{aligned}
 \text{a) } & \begin{bmatrix} 0 & \mathbf{G}(\mathbf{v}) & 0 & 0 & \mathbf{q}_b \\ \mathbf{G}^T(\mathbf{v}) & \mathbf{F}(\mathbf{s}) & \mathbf{N}^T(\mathbf{v}) & 0 & \mathbf{t}_b \\ 0 & \mathbf{N}(\mathbf{v}) & 0 & -\mathbf{E} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{v} \\ \mathbf{s} \\ \boldsymbol{\lambda} \\ \boldsymbol{\varphi} \\ \mathbf{m} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \mathbf{k} \\ 0 \end{bmatrix}; \\
 \text{b) } & I_o: \{i | \varphi_{oi} = 0\}, \bar{I}_o: \{i | \varphi_{oi} < 0\}; \\
 \text{c) } & \lambda_{I_i} \geq \lambda_{oI_i}, \lambda_{\bar{I}_i} = \lambda_{o\bar{I}_i}, \\
 & \boldsymbol{\varphi} \leq 0, \boldsymbol{\varphi}^T (\boldsymbol{\lambda} - \boldsymbol{\lambda}_o) = 0; \\
 \text{d) } & \mathbf{q}_b^T \mathbf{v} = \max!
 \end{aligned} \tag{3.2}$$

Summarizing, it can be seen that the determination of the next yield location leads to a fully nonlinear programming problem in the case of geometrical nonlinearity while in the case of the first order theory, it leads to a linear complementarity type problem. This difference does not influence the fact, that in both cases the determination of the load history can be broken down into a series of mathematical programming problems. The solution of each of these problems gives the state belonging to the next yield location(s). These solutions can be called as basic solutions of the load history. In the case of the first order theory, the linear interpolation between these states gives the exact value of the variables while in the geometrically nonlinear case, it is only an approximation.

#### 4. Numerical solution and sample problem

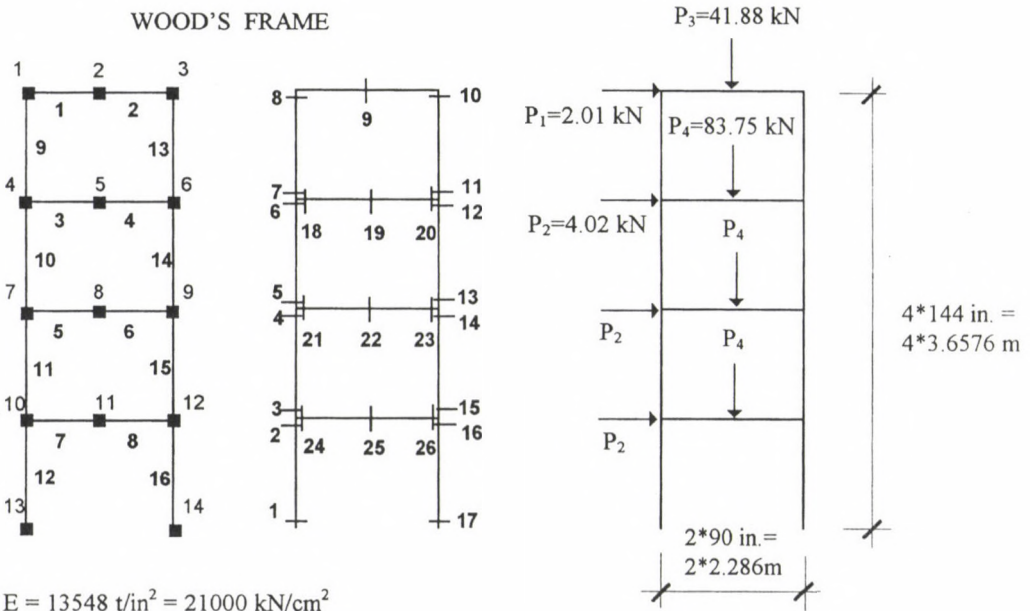
It happens sometimes that an engineering issue leads to a mathematical problem for which there is not yet an available mathematical solution algorithm. An example for this is the finite element method where the engineering solution algorithm preceeded the mathematical one. The contrary is true, too. There are fields of the mathematics which are well elaborated including solution algorithms as well but they are not widely used in engineering. One of them is mathematical programming which was initiated to solve economical problems and presently it has reached such a developed state that commercial mathematical programming packages are available which can handle several thousand unknowns and constraints. In the present study, the MINOS (Modular In Core Optimization System) [4] was used to solve the mathematical programming problems described in points 2 and 3.

This package uses Wolfe's reduced gradient method to solve the nonlinear programming problem. It demands the user to write two FORTRAN subroutines. One for the description of the vector-vector function of the lefthand side of the constraints and its Jacobian, the other one is for the description of the objective function and its gradient. The Jacobian and the gradient need not be given. In this case, MINOS calculates them by numerical differentiation. In this study, this facility has been used.

For illustration, the solution of a problem is shown which was already studied by other authors [5]. The structure is shown in Fig. 2. The frame contains 14 nodes and 16

members. The size of the unknown vector  $[v, s, \lambda, \phi, m]^T$  is:  $14 \cdot 3 + 16 \cdot (3 + 4 + 4) + 1 = 219$ . The number of constraints is:  $14 \cdot 3$  (equilibrium) +  $16 \cdot 3$  (compatibility) +  $16 \cdot 4$  (yield condition) + 1 (normality) = 155.

Figs 3-5 show the comparison of the load-displacement diagrams and load-rotation diagrams in the plastic hinges obtained by the different theories. Fig. 6. shows the displacement and bending moment diagrams belonging to the end of the load history. In the case of the first order theory, the load history ends at the plastic limit state. In the case of the geometrically nonlinear theories, the postcritical part (descending branch) of the equilibrium path can be studied, too. The deformation capacity of the plastic hinges limits the validity of the analysis. This was not examined in this study but the analysis was stopped after a certain number of steps. This means that in the geometrical nonlinear case, the endstate of the history does not have a special meaning for the given example.



MEMBER DATA:

Member	A		I		$M_p$	
	in <sup>2</sup>	cm <sup>2</sup>	in <sup>4</sup>	cm <sup>4</sup>	t.in.	kNm
1,2	5.3	34.19	55.63	2315	244.0	61.98
3-8	7.35	47.42	122.34	5092	428.0	108.71
9,13	5.89	38.00	34.71	1445	205.3	52.15
10,14	7.37	47.55	43.69	1819	259.3	65.86
11,15	8.28	53.42	86.69	3608	393.5	99.95
12,16	10.32	66.58	115.06	4789	502.3	127.58

Fig. 2.

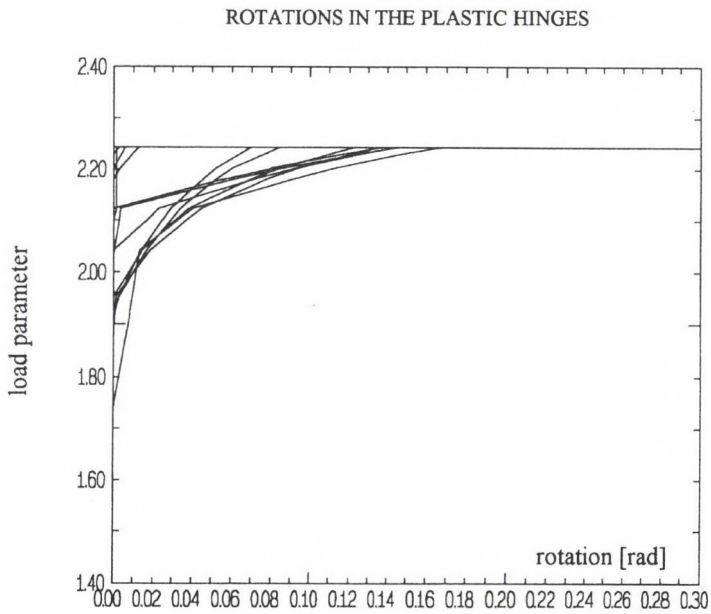
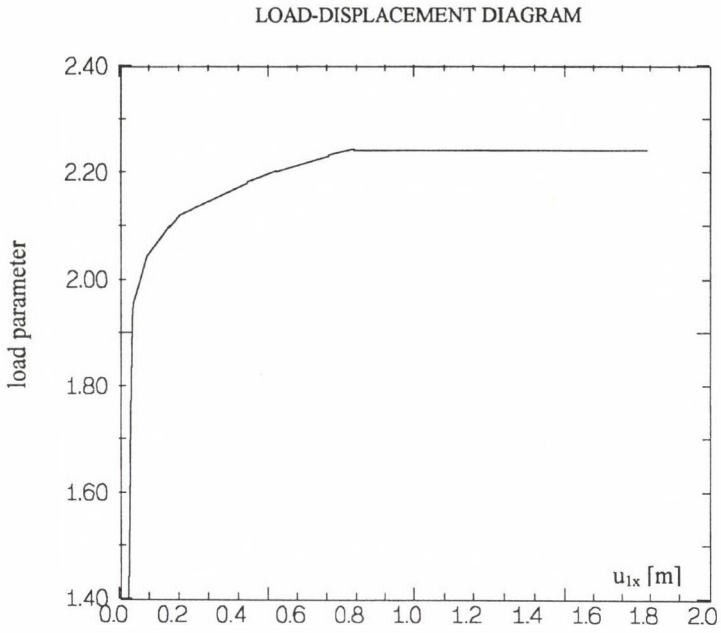
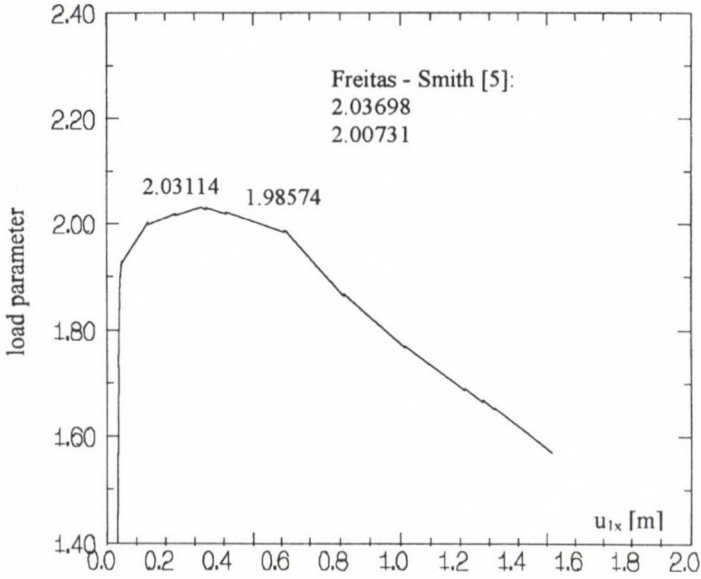


Fig. 3. 1<sup>st</sup> order theory results



LOAD-DISPLACEMENT DIAGRAM



ROTATIONS IN THE PLASTIC HINGES

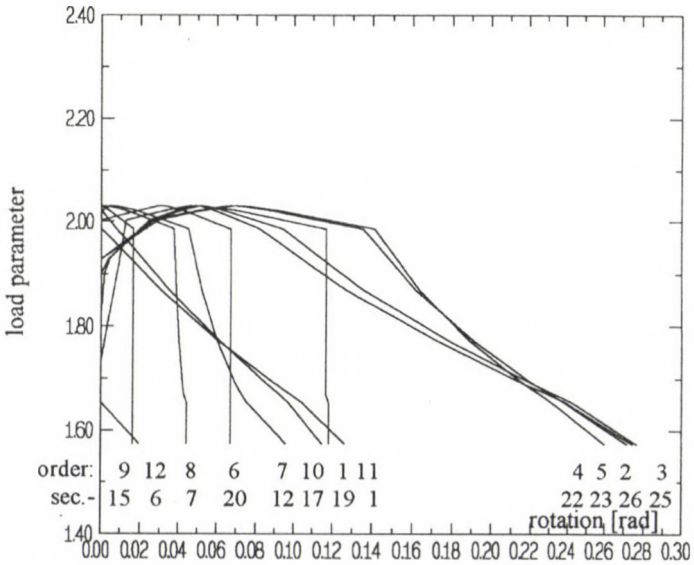
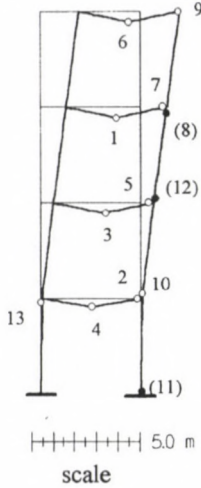


Fig. 4. 2<sup>nd</sup> order theory results



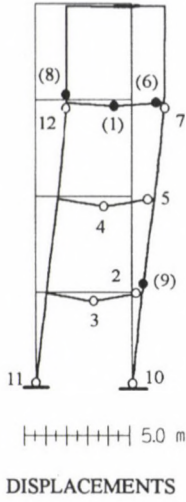
Theory: 1<sup>st</sup> order

Load parameter: 2.24



2<sup>nd</sup> order

1.57



3<sup>rd</sup> order

1.62

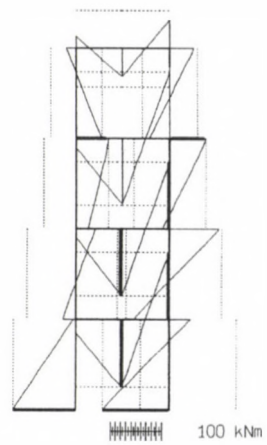
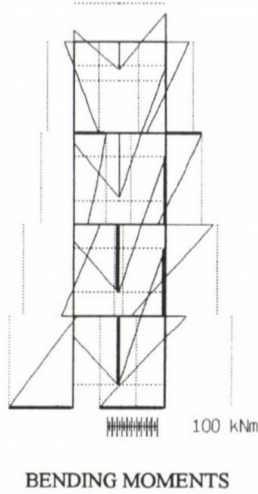
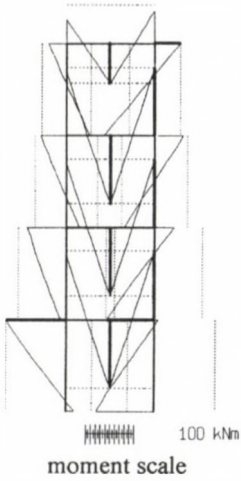
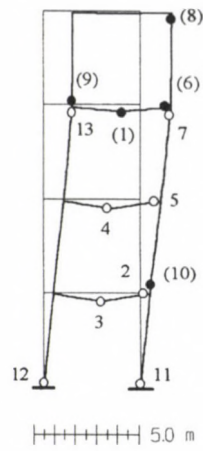


Fig. 6.



## ACKNOWLEDGEMENT

The support of the Hungarian Scientific Research Fund (OTKA) through grants T015851 and T015852 is gratefully acknowledged.

## REFERENCES

1. De Donato, O. - Maier, G.: Historical deformation analysis of elastoplastic structures as a parametric linear complementarity problem. *Meccanica*, **XI** (3) (1976) 166-171
2. Gáspár, Zs.: Large deflection analysis of bar structures. *Acta Technica Hung.*, **87** (1-2) (1978) 49-58
3. Majid, K. I.: *Non-Linear Structures*, Butterworths, London (1972)
4. MINOS User's Guide. University of California, System Optimization Laboratory (1987)
5. De Freitas, J.A.T.- Lloyd Smith, D.: Plastic straining, unstressing and branching in large displacement perturbation analysis. *International Journal for Numerical Methods in Engineering*, **20** (1984) 2077-2092

## GRÖBNER BASIS CALCULATIONS WITH APPLICATIONS TO MECHANICS

POPPER, GY.\* - KÁROLYI, GY.\*\*

(Received: 12 August 1996)

Gröbner basis calculations provide a transformation of systems of polynomial equations to a more suitable form which enables exact statements on the solutions and helps finding them with arbitrary precision. Here most important features of the method are summarized, and its application is presented for a simple engineering mechanical problem, for the stability of the propped cantilever. The resulting global equilibrium paths of the structure together with the bifurcations are shown as results of the analytical calculations.

## 1. Introduction

In mechanics there are problems which can be formulated as to solve systems of multivariate polynomial equations [1]. Consider the system of polynomial equations in  $n$  variables

$$p_1(x_1, \dots, x_n) = 0, \dots, p_k(x_1, \dots, x_n) = 0. \quad (1)$$

If you add to this set of equations a further equation

$$p_{k+1}(x_1, \dots, x_n) = 0,$$

where  $p_{k+1}$  is a linear combination of the polynomials  $p_i(x_1, \dots, x_n)$ ,  $i = 1, \dots, k$ , with polynomial coefficients  $a_i(x_1, \dots, x_n)$ , then you get an *equivalent* set of equations, i.e., one with all of the original common solutions. Similarly, you can omit one equation of them if it is a linear combination of the others.

Consequently, every calculation of solutions must take place in the set

$$\langle p_1, \dots, p_k \rangle = \left\{ \sum_{i=1}^k a_i(x_1, \dots, x_n) p_i(x_1, \dots, x_n) \right\}, \quad (2)$$

which is called *ideal* generated by the set of polynomials  $p_1, \dots, p_k$ ; the set  $\{p_1, \dots, p_k\}$  is said to form a *basis* for this ideal.

The set  $\{p_1, \dots, p_k\}$  is not the only basis for the ideal  $\langle p_1, \dots, p_k \rangle$ . One can always perform a transformation  $\{p_1, \dots, p_k\} \rightarrow \{g_1, \dots, g_l\}$  into an equivalent, so-called *Gröbner basis* for the ideal  $\langle p_1, \dots, p_k \rangle$ . In other words, the original set of equations (1) can be reduced into a

\* Popper, György, H-1016 Budapest, Szirtes u. 28/a, Hungary

\*\* Károlyi, György, H-1126 Budapest, Ugozca u. 11, Hungary

set of equations

$$g_1(x_1, \dots, x_n) = 0, \dots, g_l(x_1, \dots, x_n) = 0 \quad (3)$$

with all of the original common zeros [2-5].

Such basis *always exists*, and it has the important property, that it allows *exact* conclusions on the solutions of sets of polynomial equations, such as

- to decide whether the given set is solvable,
- whether the set has (at most) finitely many solutions,
- to determine the exact number of solutions in case there are finitely many, and
- make easier their actual computation with arbitrary precision.

The concept of Gröbner bases were introduced by Buchberger about 1965. Buchberger also presented an *algorithm* to compute a Gröbner basis. Today, most modern algebra systems (e.g. Maple or Mathematica) include a "Gröbner basis package" based on the implementation of variants of Buchberger's algorithm [6].

## 2. Term orderings

Every (non-zero) polynomial in variables  $x_1, \dots, x_n$  can be written as a sum of terms of the form  $x_1^{i_1} \dots x_n^{i_n}$  multiplied by numbers (coefficients), where  $i_1, \dots, i_n$  are non-negative integers.

*Definition 1.* An admissible *term ordering*  $<$  for the set of terms  $\{x_1^{i_1} \dots x_n^{i_n}\}$  is one which satisfies the following two conditions:

- (I)  $1 \leq t$  for all  $t \in \{x_1^{i_1} \dots x_n^{i_n}\}$ , where  $1 = x_1^0 \dots x_n^0$ ,
- (II)  $s \leq t$  implies  $sv < tv$  for all  $s, t, v \in \{x_1^{i_1} \dots x_n^{i_n}\}$ .

Each of the following is an admissible term ordering on  $\{x_1^{i_1}, \dots, x_n^{i_n}\}$ . Consider the polynomial

$$(x + y)^2 + x + y + 1$$

and suppose that  $y < x$ , that is,  $x$  is more "principal" than  $y$ .

In the (*pure*) *lexicographical term ordering* this polynomial is written as

$$x^2 + 2xy + x + y^2 + y + 1$$

because

$$x^2 >_L xy >_L x >_L y^2 >_L y >_L 1$$

are the terms in decreasing ordering, where the  $>_L$  notation is used for the lexicographical term ordering.

If the total degree of the term (i.e., the sum of the powers of the variables) is the most important property and if we use the lexicographical method to distinguish between terms of the same total degree, then the above polynomial is written as

$$x^2 + 2xy + y^2 + x + y + 1$$

because

$$x^2 >_D xy >_D y^2 >_D x >_D y >_D 1$$

are the terms in decreasing ordering. Here the  $>_D$  notation was used for the *total degree*, then *lexicographical ordering*.

The *total degree*, then *inverse lexicographical ordering* is also often used. In this system our polynomial is written as



$$y^2 + 2xy + x^2 + y + x + 1$$

because

$$y^2 <_D xy <_D x^2, \quad y <_D x,$$

so that the terms in decreasing ordering are:

$$x^2 >_D xy >_D y^2 >_D x >_D y >_D 1.$$

### 3. Solution of systems of polynomial equations

Suppose, that every (non-zero) polynomial is written in decreasing ordering of its terms (according to a fixed term ordering  $<$ ) as  $\sum_{i=1}^s c_i t_i$  with  $c_i \neq 0$  and  $t_i > t_{i+1}$  for every  $i$ . We call  $c_i t_i$  the *leading monomial*, and  $t_1$  the *leading term* of the polynomial.

*Theorem 1.* Let  $G = \{g_1, \dots, g_l\}$  be a Gröbner basis for the ideal  $\langle p_1, \dots, p_k \rangle$ . Then the system of algebraic equations (1)

$$p_1(x_1, \dots, x_n) = 0, \dots, p_k(x_1, \dots, x_n) = 0$$

is *unsolvable* (has no solutions) if and only if  $G = \{g_1, \dots, g_l\}$  contains a *constant* (polynomial).

*Theorem 2.* A system of polynomial equations has (at most) *finitely many* solutions over the complex numbers, if and only if each variable  $x_i$  appears *alone* (such as  $x_i^m$ ) in one of the *leading terms* of the corresponding Gröbner basis.

Therefore, having once obtained the Gröbner basis of a polynomial equation system, the rigorous statements can be made concerning its solvability. If this basis is computed with respect to a *lexicographic ordering*, you can determine all the solutions by the method illustrated by the following simple example.

### 4. Stability of the propped cantilever

Consider the propped cantilever shown in Fig. 1, comprising a rigid link of length  $l = 1(m)$  pinned to a rigid foundation and supported by a linear extensional spring of stiffness  $c = 1(N/m)$ . The spring is assumed to be capable of resisting both tension and compression and retains its horizontal orientation as the system deflects. Assume that the system is perfect in the sense that the spring is unstrained when the link is vertical.

The goal is to find the *equilibrium paths* of the structure loaded by a dead vertical force of magnitude  $F$ .

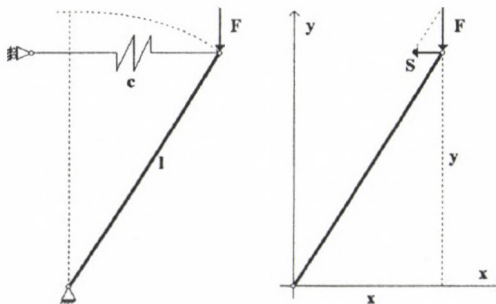


Fig. 1. Layout of the propped cantilever

The geometrical equation  $x^2 + y^2 = 1$  and the equilibrium equation (moments around the origin)  $Sy = Fx$ , with spring force  $S = cx$ , together implies a set of algebraic equations

$$xy - Fx = 0, \quad x^2 + y^2 - 1 = 0, \quad (4)$$

where  $x, y$  are the *indeterminates* (unknown coordinates of the free end-point of cantilever) and  $F$  is a *parameter* (the magnitude of the vertical force).

To compute the reduced Gröbner basis of this set of polynomials, you can apply e.g. the command

$$\text{grobner}[\text{gbasis}]({x * y - F * x, x^2 + y^2 - 1}, [x, y], \text{plex})$$

in the “Gröbner Basis Package” of the symbolic system Maple V, where  $[x, y]$  is the list of indeterminates (not including parameters) which induces the ordering  $x > y$ , and *plex* means that the pure lexicographic term ordering is used. The resulting *reduced Gröbner basis* is

$$[x^2 + y^2 - 1, \quad xy - Fx, \quad y^3 - Fy^2 - y + F] \quad (5)$$

First of all you see that the Gröbner basis does not contain a constant (polynomial) and hence (by *Theorem 1*) the corresponding set of algebraic equations is *solvable*.

If  $F$  is considered to be a parameter (only  $x$  and  $y$  are indeterminates) then  $x^2$  and  $y^3$  are the *leading terms* of the Gröbner basis (with respect not only to *plex*, but to any admissible term ordering) which contain the variables  $x$  and  $y$  alone. Consequently (by *Theorem 2*) the corresponding system of polynomial equations has a *finite number of solutions*, i.e., for any prescribed load there are finitely many equilibrium positions.

The *algorithm* for solving the Gröbner basis system (5) starts by that polynomial which contains only one indeterminate, the  $y$ . There are three solutions:  $y = F$  and  $y = \pm 1$ . For each of them you must solve the other equations.

1. For  $y = F$ , the other equations become  $x^2 + F^2 - 1$  and  $0$ , thus the two solutions for  $x$  are  $x = \pm\sqrt{1 - F^2}$ .
2. For  $y = \pm 1$  the other equations become  $x^2$  and  $(F \pm 1)x$ , and now there is only one solution for  $x$ :  $x = 0$ .

The basic idea of the solving algorithm says that, for each variable  $x_k$ , it is both necessary and sufficient to consider the polynomial of lowest degree in that variable, such that its leading coefficient does not vanish for the values of  $x_{k+1}, \dots, x_n$  being considered.

In our example, when  $y = F$ , the leading coefficient of  $(y - F)x$  vanishes, and you have to take  $x^2 + y^2 - 1$ . However, for  $y = \pm 1$ , the leading coefficient does not vanish, and it is sufficient to take this polynomial, and to ignore  $x^2 + y^2 - 1$ .

If  $F$  is considered to be indeterminate, too, then the variable  $F$  does not appear alone in one of the leading terms of the corresponding Gröbner basis, hence (by *Theorem 2*) there are *infinitely many solutions*.

To prepare the given system of algebraic equations for solving it is more advantageous to use the following command in the “Gröbner Basis Package”:

$$\text{grobner}[\text{gsolve}]({x * y - F * x, x^2 + y^2 - 1}, [x, y, F]).$$

The result is the following list of reduced subsystems whose roots are those of the original system, but whose variables have been successively eliminated and separated as far as possible:

$$[[x^2 + F^2 - 1, y - F], [x, y - 1], [x, y + 1]].$$

Fig. 2 illustrates the solution (in the interval  $-2 \leq F \leq 2$ ), and its bifurcations at  $(x, y, F) = (0, 1, 1)$  and at  $(x, y, F) = (0, -1, -1)$ . The  $x = 0$  solutions correspond to the equilibrium states of the structure with vertical bar for any values of  $F$ . However, for  $-1 \leq F \leq 1$  non-vertical layouts can also occur as equilibrium states.

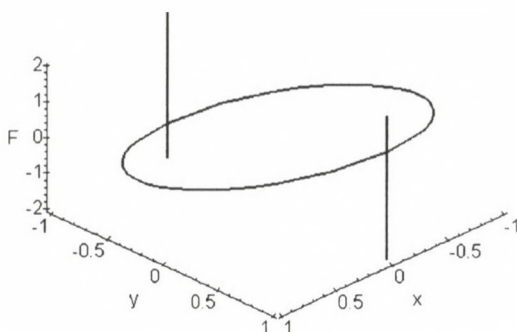


Fig. 2. Equilibrium paths of the propped cantilever

## 5. Conclusions

Main features of Gröbner basis calculations were summarized together with an application for a simple example. Compared to other traditional, known methods based on successive elimination of variables [7], this method is advantageous for three reasons: the resulting new set of equations after the transformation has exactly the same solutions as the original set, this new set is very easy to solve and the unknown parameters of the polynomials do not have to be substituted by numbers before the calculations. The latter feature allows us to earn the results of the original equation system depending on several parameters. As in many other problems in mechanics, it was important in the example, where the goal of the investigation was to find the equilibrium state of a simple structure depending on a load parameter. Knowing the explicit solutions of the corresponding equations expressed in terms of the parameter allowed the easy calculation of the global equilibrium paths. This very important property makes Gröbner basis calculations very useful in such problems.

We must note, however, that the number of steps in calculating the Gröbner basis is increasing exponentially with the number of unknowns. This can be a strong limit in its applications. However, for simpler problems one may expect interesting exact results due to the achieved analytical solutions.

## Acknowledgement

This work was supported by the Hungarian Scientific Research Foundation under Grants No. OTKA T014287 and T015860.



## REFERENCES

1. Thompson, J. M. T.-Hunt, G. W.: *A General Theory of Elastic Stability*. John Wiley & Sons, 1973
2. Davenport, J. H.-Siret, Y.-Tournier, E.: *Systems and algorithms for algebraic computation*. Computer Algebra. Academic Press, 1993
3. Becker, T.-Weispfenning, V.: *Gröbner Bases: a Computational Approach to Commutative Algebra*. Springer-Verlag, 1993
4. Geddes, K. O.-Czapor, S. R.-Labahn, G.: *Algorithms for Computer Algebra*. Kluwer Academic Publishers, 1992
5. Popper, Gy.: *Solving systems of polynomial equations using Gröbner basis calculations with applications to mechanics*. Accepted for publication by *Comp. Ass. Mech. Eng. Sci.*, 1997
6. Char, B. W.-Geddes, K. O.-Gonnet, G. H.-Leong, B. L.-Monagan, M. B.-Watt, S. M.: *Maple V, Library Reference Manual*. Springer-Verlag, 1991
7. Moses, J.: *Solution of a System of Polynomial Equations by Elimination*. *Comm. ACM* 9 (1966) 634-637

## ON THE QUASI-INVERSE OF THE MATRIX OF A SYSTEM OF LINEAR EQUATIONS

SZABÓ, J.\* - TARNAI, T.\*\*

(Received: 26 June 1996)

In this paper we will define the term *quasi-inverse* of the coefficient matrix of a system of linear equations, and will show that by its use any ill-conditioned system of linear equations can be handled easily. Its basic properties will be illustrated by a numerical example.

### 1. Introduction

In the analysis of the theoretical bases of determination of the erection shape of cable nets by means of the equilibrium and compatibility equations of bar structures in a general form [3, 4], a need has arisen for a generally valid solution to the equilibrium equation formulated by a partial change in role of the usual independent and dependent variables [2]. Investigating stability problems Tarnai [6] established that, in the neighbourhood of a critical state and then in the subsequent post-critical state, a finite mechanism-like behaviour of a bar structure becomes important. Therefore, it became necessary to formulate also a generalised solution of the compatibility differential equation similar to that of the equilibrium equation [5].

In this paper we will define the *quasi-inverse* of the coefficient matrix of a system of linear equations, and show how to obtain the above-mentioned generalised solution of a system of linear equations by the quasi-inverse and what advantages the use of the quasi-inverse has (mainly for equations like the equilibrium and compatibility equations of simultaneously statically and kinematically indeterminate bar structures).

### 2. Meaning and use of the quasi-inverse

According to the elementary rules of the matrix arithmetic, by the linear expression

$$A \cdot x + y = 0 \quad (1)$$

due to an arbitrary  $r \times c$  non-zero matrix  $A$ , to a  $c \times 1$  non-zero matrix  $x$  (vector  $x$ ) there

\* Szabó, János, H-1111 Budapest, Budafoki út 57/a, Hungary

\*\* Tarnai, Tibor, H-1037 Budapest, Kolostor u. 17, Hungary

uniquely corresponds a  $r \times 1$  matrix  $\mathbf{y}$  (vector  $\mathbf{y}$ ) so that the sum of vectors  $\mathbf{A} \cdot \mathbf{x}$  and  $\mathbf{y}$  is equal to zero. Expression (1) requires that strict rules for the sizes of matrices  $\mathbf{A}$ ,  $\mathbf{x}$  and  $\mathbf{y}$  - with respect to multiplication and addition - should be fulfilled. Otherwise, to an arbitrary vector  $\mathbf{x}$  there uniquely corresponds a vector  $\mathbf{y}$  by Eq.(1).

At the applications of the theory of sets of linear equations, with respect to the possible solutions of equation (1), very often two questions arise:

- (a) How many elements of vector  $\mathbf{y}$  can (and must) be prescribed arbitrarily?
- (b) In the case of smooth change of matrix  $\mathbf{A}$  - at the change of the rank of  $\mathbf{A}$ , and also in a neighbourhood of change of rank - how can we make the relationship between the independent and dependent variables smooth?

By introducing the term *quasi-inverse* we can answer these two questions so that to equation (1) we associate the equation

$$\mathbf{Q} \cdot \mathbf{v} + \mathbf{u} = 0$$

where matrix  $\mathbf{Q}$ , which we call *quasi-inverse*, apart from rearrangement of rows and columns, is uniquely associated to matrix  $\mathbf{A}$ ; the rank of  $\mathbf{Q}$ , defined in accordance with computational interests, determines the number of the elements of vector  $\mathbf{y}$  whose values should be prescribed, and these elements of  $\mathbf{y}$  appear in vector  $\mathbf{v}$  (the rest of elements of  $\mathbf{v}$  comes from  $\mathbf{x}$ ). In the case of a smooth change of matrix  $\mathbf{A}$ , the matrix  $\mathbf{Q}$  changes smoothly, and so the relationship between the independent  $\mathbf{v}$  and the dependent  $\mathbf{u}$  is smooth as well. These statements will be proved in the course of producing the quasi-inverse.

The *first part* of the procedure is the well-known dyadic decomposition of matrix  $\mathbf{A}$  working with complete pivoting with the following addition: if after the separation of the  $i$ th dyad the absolute value of the selected  $(i+1)$ st pivot is less than a prescribed value  $\varepsilon_Q$ , then the dyadic decomposition is considered finished, and the rank of matrix  $\mathbf{A}$  will be  $\rho = \rho(\mathbf{A}) = i$ . The elements  $a_{hk}$  ( $h = \rho+1, \rho+2, \dots, r$ ;  $k = \rho+1, \rho+2, \dots, c$ ) of the non-decomposed submatrix of  $\mathbf{A}$  will be the elements of a matrix denoted by  $\Delta$ . If the matrix  $\mathbf{A}$  is normed [1] - as in general the equilibrium matrix of bar-and-joint structures - then, to avoid the loss of digits in the computation, it is worth prescribing the value  $\varepsilon_Q = 10^{-3}$ . (If  $\mathbf{A}$  is not normed, then it can be normed.) It should be emphasized that the value of  $\varepsilon_Q$  is determined exclusively by computational aspects, and this fact does not influence the validity of the relationships.

In course of the dyadic decomposition based on pivoting, the order of columns and/or rows of matrix  $\mathbf{A}$  in general changes, and at the same time the order of the elements of vectors  $\mathbf{x}$  and  $\mathbf{y}$  also changes. If on a copy of the original matrix  $\mathbf{A}$ , the rearrangements of rows and columns are executed step by step, and the rearrangements of elements in the vectors  $\mathbf{x}$  and  $\mathbf{y}$  are also done, then a matrix  $\mathbf{C}$  instead of  $\mathbf{A}$ , a vector  $\mathbf{s}$  instead of  $\mathbf{x}$ , and a vector  $\mathbf{t}$  instead of  $\mathbf{y}$  is obtained. The equation

$$\mathbf{C} \cdot \mathbf{s} + \mathbf{t} = 0 \tag{2}$$

replacing equation (1), considering its content, is identical to (1). The elements of vectors  $\mathbf{s}$  and  $\mathbf{t}$  are:

$$\mathbf{s}^T = [x_{a1} \quad x_{a2} \quad \dots \quad x_{ar}], \quad \mathbf{t}^T = [y_{\beta 1} \quad y_{\beta 2} \quad \dots \quad y_{\beta r}].$$



The dyadic decomposition of matrix  $C$  is

$$C = \begin{bmatrix} L_1 \\ L_2 \end{bmatrix} \bullet [U_1 \mid U_2] + \begin{bmatrix} 0 & 0 \\ 0 & \Delta \end{bmatrix} \quad (3)$$

where  $L_1$  and  $U_1$  are lower and upper non-singular triangular matrices of rank  $\rho$ , and  $\Delta$  is a not decomposed submatrix of  $C$  such that the absolute value of each element of  $\Delta$  is less than  $\varepsilon_Q$ . It is obvious that expression (3) produces the submatrices

$$\begin{aligned} C_{11} &= L_1 \bullet U_1, & C_{12} &= L_1 \bullet U_2, \\ C_{21} &= L_2 \bullet U_1, & C_{22} &= L_2 \bullet U_2 + \Delta \end{aligned}$$

of matrix  $C$ :

$$C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}.$$

$C_{11}$  is a non-singular submatrix of rank  $\rho$  whose inverse, in the forthcoming investigations, will be denoted by  $Q_{11}$ :

$$Q_{11} = C_{11}^{-1} = U_1^{-1} \bullet L_1^{-1}.$$

In the *second part* of the procedure, taking the partitioned form of equation (2)

$$C_{11} \bullet s_1 + C_{12} \bullet s_2 + t_1 = 0 \quad (4a)$$

$$C_{21} \bullet s_1 + C_{22} \bullet s_2 + t_2 = 0 \quad (4b)$$

with simple transformation we obtain the expression of the quasi-inverse  $Q$ . We will execute a transformation of the system of equations (4a, b) in the following two steps. First, we multiply (4a) from the left by  $Q_{11}$  and we obtain:

$$Q_{11} \bullet t_1 + Q_{11} \bullet C_{12} \bullet s_2 + s_1 = 0, \quad (5a)$$

second, we multiply (5a) from the left by  $C_{21}$  and then subtract it from (4b) that yields:

$$-C_{21} \bullet Q_{11} \bullet t_1 + (C_{22} - C_{21} \bullet Q_{11} \bullet C_{12}) \bullet s_2 + t_2 = 0. \quad (5b)$$

Equations (5a) and (5b) can be written in the form:

$$\begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \bullet \begin{bmatrix} t_1 \\ s_2 \end{bmatrix} + \begin{bmatrix} s_1 \\ t_2 \end{bmatrix} = 0 \quad (6)$$

or

$$\mathbf{Q} \bullet \mathbf{v} + \mathbf{u} = 0 \quad (7)$$

where

$$\begin{aligned} \mathbf{Q} &= \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} \mathbf{t}_1 \\ \mathbf{s}_2 \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{t}_2 \end{bmatrix}, \\ \mathbf{Q}_{11} &= \mathbf{C}_{11}^{-1} = \mathbf{U}_1^{-1} \bullet \mathbf{L}_1^{-1}, \quad \mathbf{Q}_{12} = \mathbf{Q}_{11} \bullet \mathbf{C}_{12} = \mathbf{U}_1^{-1} \bullet \mathbf{U}_2, \\ \mathbf{Q}_{21} &= -\mathbf{C}_{21} \bullet \mathbf{Q}_{11} = -\mathbf{L}_2 \bullet \mathbf{L}_1^{-1}, \\ \mathbf{Q}_{22} &= \mathbf{C}_{22} - \mathbf{C}_{21} \bullet \mathbf{Q}_{11} \bullet \mathbf{C}_{12} = \mathbf{L}_2 \bullet \mathbf{U}_2 + \Delta - \mathbf{L}_2 \bullet \mathbf{U}_1 \bullet \mathbf{U}_1^{-1} \bullet \mathbf{L}_1^{-1} \bullet \mathbf{L}_1 \bullet \mathbf{U}_2 = \Delta. \end{aligned}$$

(It should be noted that for exact dyadic decomposition we have  $\mathbf{Q}_{22} = \Delta = 0$ .) The elements of the vectors in equation (6) are:

$$\begin{aligned} \mathbf{s}_1^T &= [x_{a1} \quad x_{a2} \quad \dots \quad x_{ap}], & \mathbf{s}_2^T &= [x_{a\rho+1} \quad x_{a\rho+2} \quad \dots \quad x_{ac}], \\ \mathbf{t}_1^T &= [y_{\beta 1} \quad y_{\beta 2} \quad \dots \quad y_{\beta p}], & \mathbf{t}_2^T &= [y_{\beta\rho+1} \quad y_{\beta\rho+2} \quad \dots \quad y_{\beta r}]. \end{aligned}$$

According to our statement, the values of vectors  $\mathbf{s}_1$ ,  $\mathbf{t}_2$  corresponding to arbitrary vectors  $\mathbf{t}_1$ ,  $\mathbf{s}_2$  are uniquely determined by equation (6):

$$\left. \begin{aligned} \mathbf{s}_1 &= -\mathbf{Q}_{11} \bullet \mathbf{t}_1 - \mathbf{Q}_{12} \bullet \mathbf{s}_2 \\ \mathbf{t}_2 &= -\mathbf{Q}_{21} \bullet \mathbf{t}_1 - \mathbf{Q}_{22} \bullet \mathbf{s}_2 \end{aligned} \right\} \equiv \mathbf{u} = -\mathbf{Q} \bullet \mathbf{v}$$

and vectors

$$\mathbf{s} = \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \end{bmatrix}, \quad \mathbf{t} = \begin{bmatrix} \mathbf{t}_1 \\ \mathbf{t}_2 \end{bmatrix}$$

composed of the given vectors  $\mathbf{t}_1$ ,  $\mathbf{s}_2$  and of the calculated vectors  $\mathbf{s}_1$ ,  $\mathbf{t}_2$  satisfy equation (2), that is, equations (4a, b). One can be convinced about this by substitution:

$$\begin{aligned} \mathbf{C}_{11} \bullet \mathbf{s}_1 + \mathbf{C}_{12} \bullet \mathbf{s}_2 + \mathbf{t}_1 &= -\mathbf{L}_1 \bullet \mathbf{U}_1 \bullet \mathbf{U}_1^{-1} \bullet \mathbf{L}_1^{-1} \bullet \mathbf{t}_1 - \mathbf{L}_1 \bullet \mathbf{U}_1 \bullet \mathbf{U}_1^{-1} \bullet \mathbf{U}_2 \bullet \mathbf{s}_2 + \\ &\quad \mathbf{L}_1 \bullet \mathbf{U}_2 \bullet \mathbf{s}_2 + \mathbf{t}_1 = 0, \end{aligned}$$

$$\begin{aligned} \mathbf{C}_{21} \bullet \mathbf{s}_1 + \mathbf{C}_{22} \bullet \mathbf{s}_2 + \mathbf{t}_2 &= -\mathbf{L}_2 \bullet \mathbf{U}_1 \bullet \mathbf{U}_1^{-1} \bullet \mathbf{L}_1^{-1} \bullet \mathbf{t}_1 - \mathbf{L}_2 \bullet \mathbf{U}_1 \bullet \mathbf{U}_1^{-1} \bullet \mathbf{U}_2 \bullet \mathbf{s}_2 + \\ &\quad \mathbf{L}_2 \bullet \mathbf{U}_2 \bullet \mathbf{s}_2 + \Delta \bullet \mathbf{s}_2 + \mathbf{L}_2 \bullet \mathbf{U}_1 \bullet \mathbf{U}_1^{-1} \bullet \mathbf{L}_1^{-1} \bullet \mathbf{t}_1 - \Delta \bullet \mathbf{s}_2 = 0. \end{aligned}$$

### 3. Numerical example

To exemplify the results that can be obtained when a set of equations with an ill-conditioned matrix is being dealt with, Rózsa [1] has shown the set of equations originated from Ralston:

$$\begin{aligned} 2x_1 + 6x_2 + y_1 &= 0 \\ 2x_1 + 6.00001x_2 + y_2 &= 0 \end{aligned}$$

whose solution for  $x_1 = 1$ ,  $x_2 = 1$  is  $y_1 = -8$ ,  $y_2 = -8.00001$ . Then he has shown the set of equations

$$\begin{aligned} 2x_1 + 6x_2 + y_1 &= 0 \\ 2x_1 + 5.99999x_2 + y_2 &= 0 \end{aligned}$$

whose solution for  $x_1 = 10$ ,  $x_2 = -2$  is  $y_1 = -8$ ,  $y_2 = -8.00002$ .

Now let us investigate how the relationship between the independent and the dependent variables can be described with the quasi-inverse in the case of a little change in a single element of the coefficient matrix. Let us write the above-mentioned sets of equations uniformly in the form  $\mathbf{A} \bullet \mathbf{x} + \mathbf{y} = 0$ :

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \bullet \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = 0$$

that in scalar form is

$$\begin{aligned} 2x_1 + 6x_2 + y_1 &= 0 \\ 2x_1 + (6+w)x_2 + y_2 &= 0. \end{aligned}$$

With respect to the simplicity of this example, we do not apply pivoting, and we accept  $a_{11}$  as a non-singular submatrix. Accordingly we have

$$\begin{aligned} \mathbf{Q}_{11} &= a_{11}^{-1} = 0.5 & \mathbf{Q}_{12} &= \mathbf{Q}_{11} \bullet a_{12} = 3 \\ \mathbf{Q}_{21} &= -a_{21} \bullet \mathbf{Q}_{11} = -1 & \mathbf{Q}_{22} &= w \end{aligned}$$

that is,

$$\mathbf{Q} = \begin{bmatrix} 0.5 & 3 \\ -1 & w \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} x_1 \\ y_2 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} y_1 \\ x_2 \end{bmatrix}.$$

Let

$$\mathbf{v} = \begin{bmatrix} -8 \\ 1 \end{bmatrix} \text{ and so } \mathbf{u} = -\mathbf{Q} \bullet \mathbf{v} = \begin{bmatrix} 1 \\ -(8+w) \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} -8 \\ -(8+w) \end{bmatrix}.$$

Obviously

$$\mathbf{A} \bullet \mathbf{x} + \mathbf{y} = \begin{bmatrix} 2 & 6 \\ 2 & 6+w \end{bmatrix} \bullet \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} -8 \\ -(8+w) \end{bmatrix} = 0.$$





$$\mathbf{B} \cdot d\mathbf{u} + d\mathbf{v} = 0, \quad (8)$$

the corresponding quasi-inverse provides possibilities similar to that for sets of linear equations. About our investigations in this respect we will report elsewhere [7].

3. If (1) is the equilibrium equation of a bar structure and (8) is the compatibility equation of the same bar structure, then

$$\mathbf{B} = \mathbf{A}^T,$$

and the quasi-inverses of  $\mathbf{A}$  and  $\mathbf{B}$  can be determined from each other.

4. With respect to conditioning of sets of equations we have concluded that conditioning of a set of linear equations, that is, its numerical handling, worsens to a greater extent if its coefficient matrix is "close" to a change of rank. "Closeness" is in part a subjective factor, but in part a measure determined by computing. It is not worth increasing the degree of accuracy of the residual term in the decomposition, as in forming the inverse of the coefficient matrix  $\mathbf{A}$ , it worsens the inverse of the non-singular submatrix  $\mathbf{A}_{11}$  having ordinarily normal properties. The quasi-inverse introduced in this paper can be a suitable means for numerical handling of ill-conditioned sets of linear equations.

5. In our current research we want to make clear what advantages the quasi-inverse provides for the investigation of stability and change of post-critical state of spatial bar structures.

#### ACKNOWLEDGEMENTS

We thank Dr. Gy. Popper and Prof. Zs. Gáspár for comments. The research reported here was supported by OTKA Grants I/3 No. 41 and T015860 awarded by the Hungarian Scientific Research Foundation.

#### REFERENCES

1. Rózsa, P.: *Linear algebra and its applications* (in Hungarian). Műszaki Könyvkiadó, Budapest 1974, p. 634
2. Szabó, J. - Kollár, L.: *Structural Design of Cable-Suspended Roofs*. Akadémiai Kiadó, Budapest and Ellis Horwood, Chichester 1984, p. 75. (In Hungarian: Műszaki Könyvkiadó, Budapest 1974, p. 59.)
3. Szabó, J. - Rózsa, P.: Die Matrizengleichung von Stabkonstruktionen (Im Falle kleiner Verschiebungen). *Acta Technica Acad. Sci. Hung.* 71 (1971), 133-148
4. Szabó, J. - Rózsa, P.: Grosse Verschiebungen von Stabkonstruktionen. *Acta Technica Acad. Sci. Hung.* 73 (1972), 53-60
5. Szabó, J. - Tarnai, T.: General theory and numerical analysis of single- and double-layer pin-jointed space grids. In: Parke, G.A.R., Howard, C.M. (eds): *Space Structures 4*, Vol. 1, Thomas Telford, London 1993, pp. 715-722
6. Tarnai, T.: *Kinematically indeterminate structures and structural topology* (in Hungarian). DSc thesis, Budapest 1990
7. Tarnai, T. - Szabó, J.: Analysis of inextensional, kinematically indeterminate assemblies. Manuscript for publication 1997





## PACKING OF EQUAL CIRCLES IN A SQUARE

TARNAI, T.\* - GÁSPÁR, ZS.\*\*

(Received: 5 June 1996)

The problem of the densest packing of  $n$  equal non-overlapping circles in a square is investigated. An improved packing for  $n = 19$  is presented, and the proven and conjectured packings known so far are summarized up to  $n = 20$ . Heuristic upper bounds of the maximum packing density are given, and their numerical values are listed up to  $n = 40$ .

### 1. Introduction

A classical problem of discrete geometry is: To determine the greatest possible minimum distance  $d_n$  between  $n$  points which can be distributed in the unit square; or what is the same, to determine the largest diameter  $d_n$  of  $n$  equal circles which can be packed in a square of side  $1 + d_n$  without overlapping.

Exact solution of this problem is known only for  $2 \leq n \leq 9$  and  $n = 14, 16, 25$  and  $36$  [2]. In 1970, Goldberg [5] published conjectured solutions for  $n \leq 27$  and some sporadic results for  $n > 27$ . Since then many of his results have been improved: for  $n = 10$  by Schlüter [12], for  $n = 11$  and  $13$  by Mollard and Payan [8], for  $n = 14$  by Wengerodt [15], and recently for  $n = 17$  by Melissen and Schuur [7].

Density of packing, denoted by  $D_n$ , is defined as the ratio of the total area of the circles to the area of the square:  $D_n = nd_n^2\pi / [4(1 + d_n)^2]$ .

The aim of this paper is twofold: firstly, to show that for  $n = 19$  Goldberg's packing configuration is not optimal, and to present a new configuration resulting in an improved lower bound on the maximum diameter of the circles, and so an improved lower bound on the maximum density of packing; secondly, to present an upper bound on the maximum packing density.

\* Tarnai, Tibor, H-1037 Budapest, Kolostor u. 17, Hungary

\*\* Gáspár, Zsolt, H-1025 Budapest, Kapy u. 40/b, Hungary

## 2. Improvement of Goldberg's packing

We consider the circle packing version of the problem. The idea of improvement is based on the theory of rigidity of graphs as developed by the authors [13] for packing of circles on a sphere and applied by Connelly [1] for packing of circles in a plane container. To an arrangement of equal non-overlapping circles there corresponds a *graph*. In the case of a general container, the graph of a circle packing defined by Connelly [1] has two different edge lengths. In the case of a square container, however, we can apply a simplification; and we define the graph in the following way: The vertices of the graph are the centres of the circles, and the edges of the graph are straight line segments joining the centres of the touching circles. Thus, all edges of the graph are of equal length, and all edges and vertices of the graph are contained in the closed unit square.

In the case of spherical circle packing Danzer [3] has considered the graph so that the edges can rotate freely around the vertices and the edge lengths can vary freely but simultaneously and in the same proportion. He has defined the graph to be *rigid* if the edge system with the mentioned properties cannot admit motions other than isometries. Danzer's idea is: If the graph is not rigid then it can, in general, be improved.

As shown by Pólya [9], mechanical analogues can be useful in solving mathematical problems. We follow this way, and in our investigation the graph is modelled as a structure consisting of straight bars and frictionless pin joints. Analogously to [13], the bar-and-joint structure is characterized by its geometric or compatibility matrix  $\mathbf{G}$  containing  $b$  rows and  $2j - c$  columns where  $b$  is the number of bars (number of edges of the graph),  $j$  is the number of joints (number of vertices, that is,  $j = n$ ), and  $c$  is the number of supporting forces or reactions (number of kissing points of circles and sides of the square). A reaction force is passing through the point of a circle at which the side of the square is tangent to the circle and through the centre of that circle, so its line of action is perpendicular to the side of the square. (In the case of spherical circle packing  $c$  is constant:  $c = 3$ .) With the help of the geometric matrix of this structure, it is ascertained whether the bar lengths can simultaneously be increased in the same proportion without inner forces. If it is so then it can be done, e.g. due to a uniform increase in the temperature of the bars, until the distance between some of the vertices not connected by an edge will be equal to the increased edge length of the graph where new edges should be introduced. By appearing additional bars the further motions of the structure will be prevented, i.e. the graph will be made rigid in Danzerian sense. If further increase in the temperature would cause a stable state of self-stress where all bars are in compression, then a local optimum is arrived at. The circle packing is, in general, improvable if one of the three properties holds:

$$\begin{aligned} b &= r = 2j - c, \\ b &= r < 2j - c, \\ b &> r, \quad r < 2j - c \end{aligned}$$

in which  $r$  denotes the rank of  $\mathbf{G}$ . The details of this "heating technique" can be found in [13].

Let us consider the arrangement constructed by Goldberg for 19 circles and apply the above-mentioned improving technique. Goldberg's packing and its graph is shown in Fig. 1a. The diameter of the circles in this packing, with seven-digit accuracy, is 0.2895365. The graph contains two isolated points. If we reckon that two edges are needed in order to fix an isolated point, then  $j = 19$ ,  $c = 11$  and  $b = 27$ . Thus,  $2j - c = b$  holds; that is, matrix  $\mathbf{G}$  is quadratic. Due to the symmetry, the graph can have a state of self-stress of one parameter, and so the third



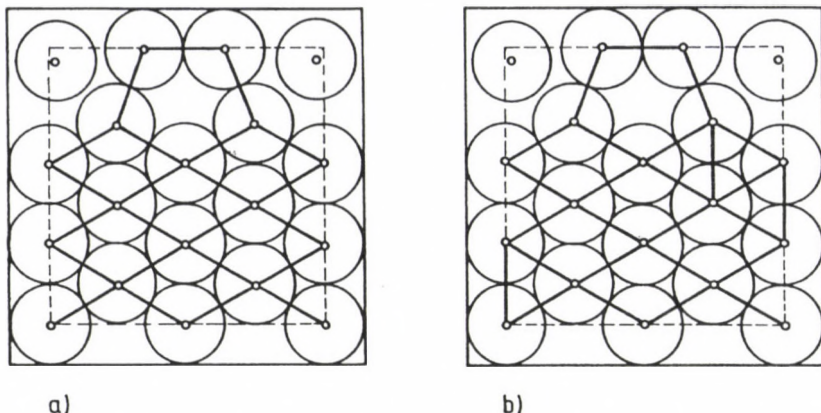


Fig. 1. Packing of 19 circles in a square: (a) circle system and its graph due to Goldberg, (b) the improved configuration and its graph. (The figure composed of dashed lines is the unit square.)

property of the above three holds. In a state of self-stress, all bars can be in compression (or inactive), but the equilibrium is not stable, so the arrangement can be improved.

The graph can be deformed with a simultaneous increase in the edge length, without causing stresses in the edges, until the graph is supplemented by three edges which make the graph rigid. With a further increase in the temperature, the graph will have no edges in tension. Therefore, we have obtained a new packing in which the diameter of the circles is

$$d_{19} = 0.2895419$$

which is a local optimum. The packing and its graph is presented in Fig. 1b.

### 3. Upper bounds on packing density

Let  $D$  denote the maximum density of packing of  $n$  equal circles in a square. Fejes Tóth [4] has shown that the density of packing of at least two equal circles in a convex domain in the plane is always less than  $\pi / \sqrt{12}$ . Therefore,

$$D < \frac{\pi}{\sqrt{12}}. \quad (1)$$

Groemer [6] has sharpened Fejes Tóth's inequality, and his formula has resulted in the following upper estimate better than (1):

$$D \leq \frac{n\pi}{\left[2 - \sqrt{3} + \sqrt{7 - \pi + \sqrt{3}(2n - 6 + \pi)}\right]^2}. \quad (2)$$

In what follows, we give a heuristic reasoning for further sharpening of Groemer's inequality (2), using the special properties of a square.



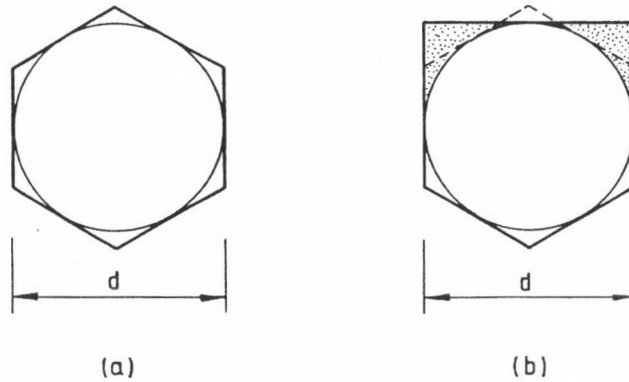


Fig. 2. The Dirichlet cell of a circle (a) in the densest packing in the plane, (b) at the boundary of the square

Consider equal circles of diameter  $d$  packed in a square of side length  $1 + d$ . In this square the *Dirichlet cell* is defined as a domain which consists of all points of the square which are nearer to the centre of a particular circle of the packing than to any other centre. In the densest packing of equal circles in the plane, the circles form a hexagonal arrangement in which the Dirichlet cells are regular hexagons of side length  $d/\sqrt{3}$  (Fig. 2a). The area  $(\sqrt{3}/2)d^2$  of such a hexagon can be considered as *space claim* of a circle. If circles touching a side of the square are in a close arrangement then the Dirichlet cell of such a circle is a pentagon (Fig. 2b) whose area is larger than the space claim of the circle by  $d^2(2-\sqrt{3})/4$ , and this difference appears as the area of an extra interstice corresponding in fact to a semicircle laying in the domain of width  $d/2$  around the unit square. The largest density occurs, if as many circles as possible are touching the boundary of the square, that is, the centres of circles are situated along the sides of the unit square with separation  $d$  between them, and each vertex of the unit square is the centre of a circle (Fig. 3). Along a side of the unit square we cannot put more than  $\text{int}(1/d)$  semicircles, and at a vertex we can put additionally at most a quarter of a circle (a half of a semicircle). [Here the symbol  $\text{int}(x)$  denotes the integer part of the real number  $x$ .] If we add the area of the extra interstices determined above along the boundary of the square, we obtain a lower bound on the real extra interstice area. This bound is even smaller if along the fractional distance  $1 - \text{int}(1/d)d$ , the average extra interstice area is taken into account instead of the actual one. Therefore, a lower bound  $A_e$  of the area of the sum of extra interstices is

$$A_e = \left( d + \frac{d^2}{2} \right) (2 - \sqrt{3}).$$

Let  $A_c$  be the space claim of  $n$  circles:

$$A_c = n \frac{\sqrt{3}}{2} d^2.$$

$A_c + A_e$  cannot be greater than the area of the square of side length  $1 + d$ . Thus, we have the inequality

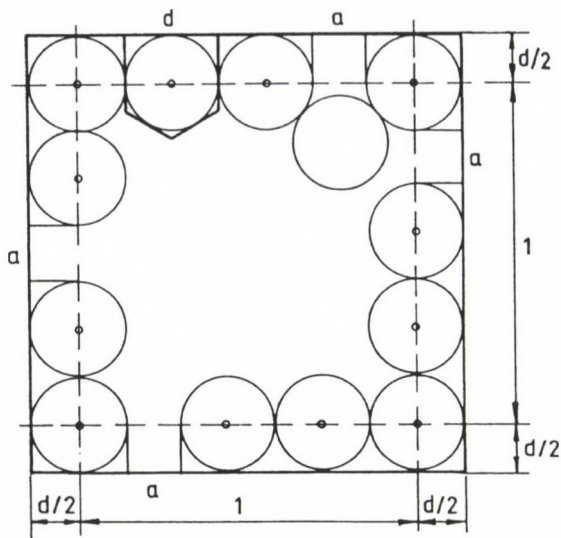


Fig. 3. Arrangement of circles along a side of the square. A gap is at each of the sides

$$n \frac{\sqrt{3}}{2} d^2 + \left( d + \frac{d^2}{2} \right) (2 - \sqrt{3}) \leq (1 + d)^2.$$

From here  $d$  can be expressed, and as  $D = nd^2\pi / [4(1 + d)^2]$ , we obtain an upper bound on the density  $D$ :

$$D \leq \frac{n\pi}{\left[ 2 - \sqrt{3} + \sqrt{3 + 2\sqrt{3}(n-1)} \right]^2} \quad (3)$$

which is exact for  $n = 1$ , and which is better than (2) for every  $n$ .

Let us make the upper bound of packing density (3) sharper by calculating the area of the extra gaps along the boundary exactly. Doing so we have to consider also two additional dense arrangements of circles. Therefore, we have three different possibilities for dense packing of the circles along the boundary.

(a) The above-mentioned arrangement, that is, a circle is packed at each vertex of the square, the other circles are closely packed along the sides and a gap of width  $a = 1 - \text{int}(1/d)d$  appears at each side. Consider the arrowhead-like domain composed of the rectangle  $ABCD$  and the triangle  $EFG$  (Fig. 4). Let its area be denoted by  $A_a$ . Then

$$A_a = a \frac{d}{2} + \frac{a+d}{2} \sqrt{d^2 - \left( \frac{a+d}{2} \right)^2}.$$

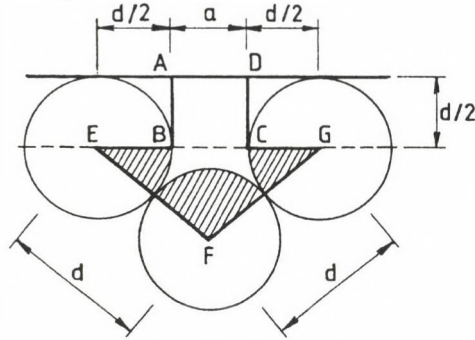


Fig. 4. The gap at the side of the square

The area of the extra gap  $A_g$  is obtained if the space claim of a semicircle (the area of the half of a hexagon) is subtracted from the area of the arrowhead-like domain:  $A_g = A_a - \sqrt{3}d^2/4$ , that is, as a function of  $a$

$$A_g(a) = \frac{ad}{2} + \frac{a+d}{2} \sqrt{d^2 - \left(\frac{a+d}{2}\right)^2} - \frac{\sqrt{3}}{4} d^2. \quad (4)$$

Since on each side of the square there is a gap, the area inequality is obtained for  $n$  equal circles as

$$n \frac{\sqrt{3}}{2} d^2 + \left( \text{int} \left( \frac{1}{d} \right) + \frac{1}{2} \right) d^2 (2 - \sqrt{3}) + 4A_g(a) \leq (1+d)^2, \quad n \geq 5. \quad (5)$$

(b) The gaps appear at two opposite vertices and two adjacent sides such that the gap arrangement is (can be) symmetrical with respect to a diagonal of the square. Let us introduce the circle numbers  $n_1, n_2$  along the sides and distances  $a_1, a_2, c$  in Fig. 5:

$$\begin{aligned} n_1 &= \text{int} \left( \frac{1}{d} \right), \\ a_1 &= 1 - n_1 d, \\ c &= \sqrt{d^2 - a_1^2} - \frac{d}{2}, \\ n_2 &= \text{int} \left( \frac{1 + \frac{d}{2} - c}{d} \right), \\ a_2 &= 1 + \frac{d}{2} - c - n_2 d. \end{aligned}$$



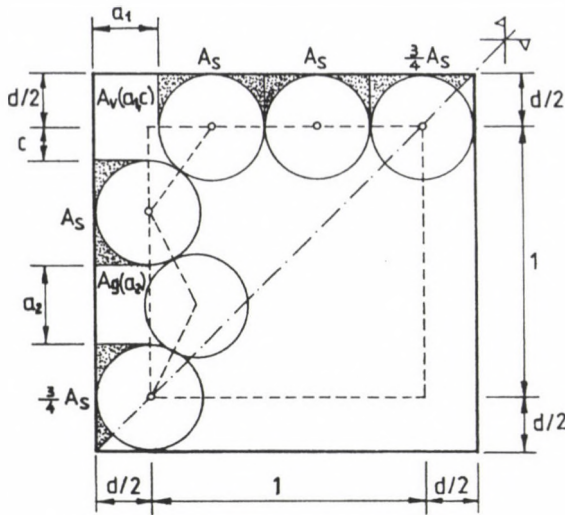


Fig. 5. Arrangement of circles along the sides of the square. Gaps are at two opposite vertices and at two adjacent sides

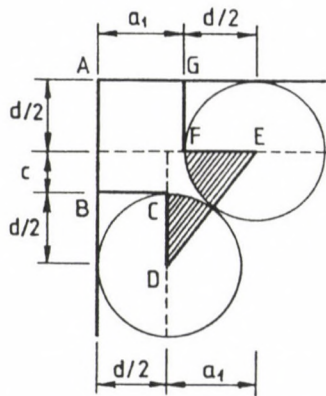


Fig. 6. The gap at a vertex of the square

The extra interstice area corresponding to a semicircle at the boundary, as previously obtained, is

$$A_s = \frac{d^2}{4}(2 - \sqrt{3}).$$

So, the area of the extra interstices corresponding to the shaded area in Fig. 5 is  $(n_1 + n_2 + 1/2)A_s$ . The area of the polygon  $ABCDEF$  in Fig. 6, at the vertex of the square, is  $(a_1 + c)d/2 + (c + d/2)a_1/2$ . Subtracting the space claim of a quarter of a circle from that area we obtain the extra gap area  $A_v$  at a vertex of the square:



$$n \frac{\sqrt{3}}{2} d^2 + \left( 2n_1 + n_2 + n_3 + \frac{3}{2} \right) A_s + 2A_v(a_1, c) + A_g(a_3) + A_v(a_2, c_2) \leq (1+d)^2, \quad n \geq 3. \quad (8)$$

For given value of  $n$ , from (5), (7), (8) we can determine numerically an upper bound  $d_a, d_b, d_c$ , respectively, on the diameter of the circles, whose maximum  $d_m$  is an upper bound on the maximum diameter  $d$ :

$$d_m = \max(d_a, d_b, d_c) \geq d,$$

and we obtain an upper bound of the maximum density  $D$ :

$$D \leq \frac{nd_m^2\pi}{4(1+d_m)^2} \quad (9)$$

which is better than (3). (The upper bound (9) is always less than the upper bound (3) as if circles can be packed along a side without gaps, then the centres of the circles lie on the side of the unit square, but if the other circles are packed in a regular triangular lattice packing then circle centres can never lie on the opposite side of the unit square, because the ratio of the side to the altitude of a regular triangle is an irrational number.) This upper bound is exact for  $n = 2$  (obtained from (7)), for  $n = 3$  (obtained from (8)), and for  $n = 5$  (obtained from (5)).

It remained to be shown that, if the gap on a side is not concentrated at one point on the side of the square but it is divided into two parts at two different points of the side, the upper bound of density cannot decrease, that is, for the function  $A_g(a)$  in (4), the inequality

$$A_g(a_k + a_l) \leq A_g(a_k) + A_g(a_l)$$

is valid if  $a_k + a_l \leq d$ . For this it is enough to show that  $A_g(a) \geq 0$  and the second derivative of  $A_g(a)$  is not positive in the interval  $0 \leq a \leq d$ . Positivity of  $A_g(a)$  is obvious. The second derivative of  $A_g(a)$  is

$$\frac{d^2 A_g}{da^2} = -\frac{3}{4} \frac{a+d}{2} \left[ d^2 - \left( \frac{a+d}{2} \right)^2 \right]^{-1/2} - \frac{1}{4} \left( \frac{a+d}{2} \right)^3 \left[ d^2 - \left( \frac{a+d}{2} \right)^2 \right]^{-3/2} \leq 0$$

since both terms in it are  $\leq 0$ . Therefore, the function  $A_g(a)$  is concave from below in the interval  $0 \leq a \leq d$ , that is, we obtain greater density if the gaps at a side are united into a single gap.

In the cases (b) and (c), there are gaps at vertices of the square, and also at some side there is a gap between circles touching the side. If at such a side, the circles are arranged so that the gap here joins the gap at the vertex, then the area of the extra gap calculated from the united gaps is larger than the sum of the area of the two separate extra gaps. This is so, because in the new position, the circle at the apex of the "arrowhead" corresponding to the gap at the side is at a larger distance from the side in question, since the other circle forming a part of the boundary of the gap at the vertex does not touch the side in question. Therefore, we obtain greater density if the gaps at a side and at the vertex are separated.



#### 4. Conclusions

1. The improving technique applied in this paper for  $n = 19$  can be applied also for some further packings of Goldberg [5]: for instance, for  $n = 22$  and 26 (in the case  $n = 22$  the graph can be decomposed into a set of isolated points, in the case  $n = 26$  the graph is not rigid), and probably for  $n = 21$  (in this case the graph cannot be reconstructed with absolute certainty).

2. A survey of the results of packings of circles in a square last time was given in [2]. Since the list presented there contained minor mistakes and since recently some new results were published it is worth summarizing the results and giving an up-to-date list. The data of the proven and conjectured best packings known so far are collected in Table 1. The graphs of packings are presented in Fig. 8. The survey is given only up to  $n = 20$  since, as mentioned, many of the packings known for  $n > 20$  are not even locally optimal.

3. The upper bounds on the maximum packing density calculated by the formulae (2) and (3) as well as by (9), based on the maxima of the results obtained from (5), (7), (8), are given for up to  $n = 40$  in Table 2. Contrary to the upper bound given by (3) the upper bound given by (9) is not monotonous with  $n$ . As to (9), for most values of  $n$ , from the three inequalities (5), (7), (8), inequality (5) provides an upper bound on the maximum packing density.

Table 1. Proven and conjectured densest packing of  $n$  equal circles in a square of side  $1 + d_n$  (spreading of  $n$  points in the unit square with the greatest minimum separation  $d_n$  possible between them)

$n$	Diameter $d_n$	Density $D_n$	Reference
2	1.4142136	0.5390121*	
3	1.0352762	0.6096448*	
4	1	0.7853981*	
5	0.7071067	0.6737651*	
6	0.6009252	0.6639569*	Graham, cf [10]
7	0.5358983	0.6693108*	Schaer, cf [10]
8	0.5176380	0.7309638*	Schaer and Meir [11]
9	0.5	0.7853981*	Schaer [10]
10	0.4212795	0.6900357	Schlüter [12]
11	0.3982073	0.7007415	Mollard and Payan [8]
12	0.3887301	0.7384682	Goldberg [5]
13	0.3660960	0.7332646	Mollard and Payan [8]
14	0.3489152	0.7356792*	Wengerodt [15]
15	0.3410813	0.7620560	Goldberg [5]
16	0.3333333	0.7853981*	Wengerodt [14]
17	0.3061539	0.7335502	Melissen and Schuur [7]
18	0.3004626	0.7546533	Goldberg [5]
19	0.2895419	0.7523079	this work
20	0.2866116	0.7794936	Goldberg [5]

\* It is proved that the density is a maximum

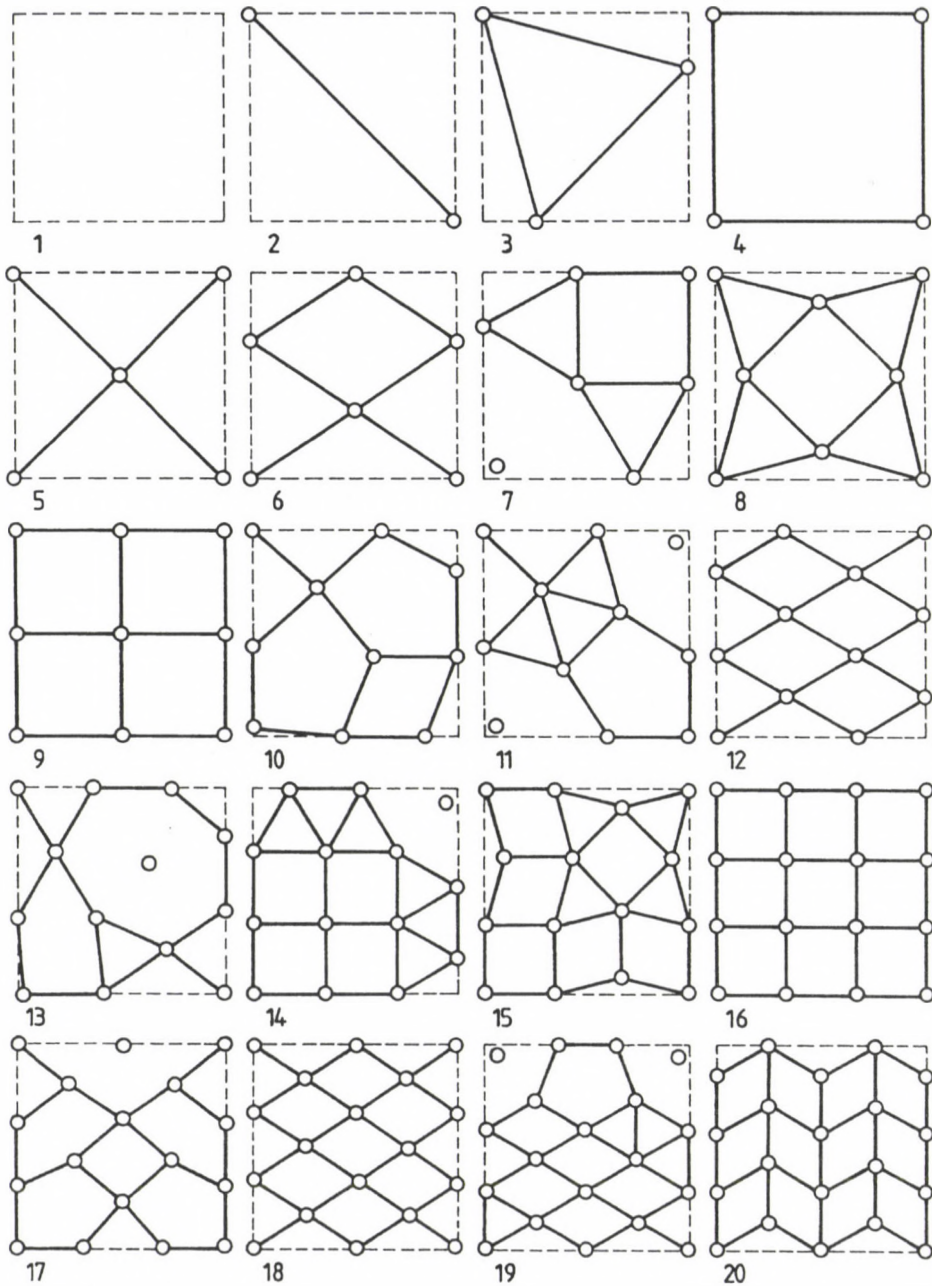


Fig. 8. The graph of the proven and conjectured densest packing of  $n$  equal circles in a square (arrangement of  $n$  points in the unit square such that the minimum distance apart is as great as possible)

Table 2. Upper bounds on the maximum density  $D$  of packing of  $n$  equal circles in a square

$n$	With Groemer's formula (2)	With average extra interstice area, formula (3)	With exact extra gap area, formulae (5), (7), (8)
2	0.8724125	0.7955012	0.5390121*
3	0.8563422	0.8063246	0.6096448**
4	0.8519408	0.8146624	0.7854052*
5	0.8509826	0.8211847	0.6737651
6	0.8513033	0.8264434	0.6701081
7	0.8521455	0.8307966	0.7123433*
8	0.8531989	0.8344782	0.7375840
9	0.8543237	0.8376460	0.7930670
10	0.8554535	0.8404109	0.8131862
11	0.8565563	0.8428525	0.7767688
12	0.8576172	0.8450303	0.7641682
13	0.8586298	0.8469892	0.7639387
14	0.8595926	0.8487642	0.7903124*
15	0.8605062	0.8503828	0.7858901
16	0.8613728	0.8518670	0.8066345
17	0.8621948	0.8532347	0.8369299
18	0.8629750	0.8545007	0.8393169
19	0.8637163	0.8556772	0.8229551
20	0.8644212	0.8567744	0.8132350
21	0.8650924	0.8578009	0.8083586
22	0.8657323	0.8587642	0.8072574
23	0.8663430	0.8596706	0.8172430*
24	0.8669266	0.8605255	0.8185761*
25	0.8674850	0.8613338	0.8221864
26	0.8680198	0.8620995	0.8336828
27	0.8685326	0.8628264	0.8508564
28	0.8690249	0.8635176	0.8571610
29	0.8694979	0.8641761	0.8479190
30	0.8699529	0.8648043	0.8411183
31	0.8703910	0.8654046	0.8363717
32	0.8708132	0.8659790	0.8333982
33	0.8712203	0.8665293	0.8320005
34	0.8716134	0.8670572	0.8320529
35	0.8719932	0.8675642	0.8439693*
36	0.8723604	0.8680516	0.8389198*
37	0.8727157	0.8685207	0.8407982
38	0.8730597	0.8689727	0.8471696
39	0.8733930	0.8694085	0.8564647
40	0.8737162	0.8698291	0.8689143

\* Density is due to (7).

\*\* Density is due to (8).

All other densities not marked in the last column are due to (5).



### 5. Remark

The main body of this paper was prepared for the Constructive Geometry Conference held in Balatonföldvár, in September 1993. Soon after this conference there appeared a nice paper by Peikert [16], in which, for  $n = 19$ , exactly the same solution was presented as we found independently. That time we thought that in that situation this paper is not worth publishing as it would be only a confirmation of Peikert's result. The reason why eventually we present it here is that in the meantime it became a problem to provide a sharp upper bound on the maximum density of packing of equal circles in particular containers in the plane (square, circle) and on a hemisphere, and in this paper, after revising the 1993 manuscript, we give upper heuristic estimates on the packing density in a square, which can be of interest.

### ACKNOWLEDGEMENT

This work was partially supported by OTKA Grants I/3 No. 41 and T014285 awarded by the Hungarian Scientific Research Foundation.

### REFERENCES

1. Connelly, R.: Rigid circle and sphere packings. Part I: Finite packings. *Structural Topology* 14 (1988), 43-60
2. Croft, H.T. - Falconer, K.J. - Guy, R.K.: *Unsolved Problems in Geometry*. Springer, New York 1991
3. Danzer, L.: Finite point-sets on  $S^2$  with minimum distance as large as possible. *Discrete Math.* 60 (1986), 3-66. [Endliche Punktmengen auf der 2-Sphäre mit möglichst grossem Minimalabstand. *Habilitationschrift*, Göttingen 1963]
4. Fejes Tóth, L.: *Lagerungen in der Ebene, auf der Kugel und im Raum*. Zweite Auflage. Springer, Berlin 1972
5. Goldberg, M.: The packing of equal circles in a square. *Math. Mag.* 43 (1970), 24-30
6. Groemer, H.: Über die Einlagerung von Kreisen in einen konvexen Bereich. *Math. Z.* 73 (1960), 285-294
7. Melissen, J.B.M. - Schuur, P.C.: Packing seventeen equal circles into a triangle, a square and a circle. *Philips Research Manuscript MS 16.969* (1992)
8. Mollard, M. - Payan, C.: Some progress in the packing of equal circles in a square. *Discrete Math.* 84 (1990), 303-307
9. Pólya, G.: *Mathematics and Plausible Reasoning*. Vol. I. Induction and Analogy in Mathematics. 2nd edn. Princeton University Press 1968
10. Schaer, J.: The densest packing of 9 circles in a square. *Canad. Math. Bull.* 8 (1965), 273-277
11. Schaer, J. - Meir, A.: On a geometric extremum problem. *Canad. Math. Bull.* 8 (1965), 21-27
12. Schlüter, K.: Kreispackung in Quadraten. *Elem. Math.* 34 (1979), 12-14
13. Tarnai, T. - Gáspár, Zs.: Improved packing of equal circles on a sphere and rigidity of its graph. *Math. Proc. Camb. Phil. Soc.* 93 (1983), 191-218
14. Wengerodt, G.: Die dichteste Packung von 16 Kreisen in einem Quadrat. *Beiträge Algebra Geom.* 16 (1983), 173-190
15. Wengerodt, G.: Die dichteste Packung von 14 Kreisen in einem Quadrat. *Beiträge Algebra Geom.* 25 (1987), 25-46
16. Peikert, R.: Dichteste Packungen von gleichen Kreisen in einem Quadrat. *Elem. Math.* 49 (1994), 16-26



## DYNAMIC ELASTO-PLASTIC ANALYSIS OF STRUCTURES BY MATHEMATICAL PROGRAMMING

VÁSÁRHELYI, A\* - LÓGÓ, J.\*\*

(Received 26 July 1996)

There are a number of methods for the plastic analysis of structures, but not all of them are suitable for the calculation with time dependent loading. Here a new method is presented which is based on nonlinear mathematical programming in the function space  $L^2$ . The dissipative energy and the inertia forces are taken into consideration and the "dissipative forces" are characterized as internal variables. The original model which was created in function space  $L^2$  is transformed into the vector-space during the calculation procedure and the results are transformed back after the computation. The theoretical results are illustrated by numerical example.

### 1. Introduction

Usually the mechanical models are given by different type of differential equations. Their direct solution are obtained after a very difficult calculation or these problems have to solve on approximated way. During the 70-s the mathematical programming as a tool became primary importance in structural plasticity [3]. Dynamic problems usually are solved as an eigenvalue problem of the discretized structures [13]. Mathematically it leads to some type of direct time integration methods and most of them are related to so-called "tangent stiffness" method in which the stiffness matrix of system is modified at every loading step [7,9].

Plasticity as a technical term is defined in mechanics as a certain determination of irreversible deformation of the structures caused by external actions. Due to the plastic deformation some kind of energy is always dissipated. Generally, the theories used in structural mechanics for describing elasto-plastic change with time independent material flow are formulated on the basis of macroscopic observation of the material behaviour [4,8]. The plastic problems are handled mostly as a rate independent one and the time parameter  $t$  measures the order of the events, rather than real time [6]. Dynamical and plastic problems are calculated together not very often and we have not found any method to take the dissipation into consideration.

In this paper a new method is presented which based on nonlinear mathematical programming in the function space  $L^2$ . The dissipative energy and the inertia forces are taken into consideration and the "dissipative forces" are characterized as internal variables. In that work the effect of damping is not considered. The original model which was created in function space  $L^2$  is transformed into the vector-space during the calculation procedure and the results are transformed back after the computation. The theoretical results are illustrated by numerical example.

\* Vásárhelyi, Anna, H-1126 Budapest, Kiss J. alt. u. 34, Hungary

\*\* Lógó, János, H-1118 Budapest, Regős u. 7, Hungary



## 2. Formulation of the structural problem

Ordinarily the variational interpretation of a problem is created according to the type of differential equations that occur (including the initial boundary conditions), and then the approximation functions to be used have an appropriate order which follows from the type of each equation. In the approach presented here the unconstrained optimization problem that derives from the boundary problem is decomposed into a pair of constrained optimization problems. The individual material properties are included as functions of the state characteristics and this fact connects into a system the parts coming from the original equations. Decomposition of the unconstrained optimization problem into a constrained optimization problem is not unambiguous, but the different decompositions all must satisfy the same Euler-Lagrange equations. It is executed in such a way that the primal-dual problems have physical meanings. From the mathematical point of view, the primal and the dual variables are intensive and extensive mechanically and the objective function contains different types of energies.

The state variables (e.g. stresses, strains, etc.) are given in vector space by vector-scalar functions in the case of equilibrium state. For a discretized structure that vector space is supposed to be an 3 dimensional space in a global coordinate system. Every node of the element is defined by a position vector. To each position vector a state vector described in the local coordinate system is attached. The number of the independent components depends on the freedom of the nodes and it equals  $n$  times the dimension of the state vector.

In case of time-dependent problems the state functions are given in both local and global coordinate systems and are vector-vector functions which depend on the time. Within the context of a small displacement theory, the position vectors are time independent.

In this paper the presented models include the theoretical results of the earlier studies [10,11,12]. It is supposed that the Euler and Lagrange descriptions of the structure correspond, that is the position vector does not depend on time. It means that the state characteristics can be given using function subspaces based on the local coordinate axes, and can be expressed via a linear combination of the (unknown) coefficients of basis-functions of these spaces.

In other words, the state variables (e.g.  $\sigma(x,t)$ ) are approximated by:

$$\sigma(x,t) \approx \sum_{i=1}^n \mathbf{x}(t) \cdot N(\xi) \quad (1)$$

where  $N(\xi)$  notes the shape functions depend on the position vector ( $\xi$ ) and the coefficients  $\mathbf{x}(t)$  depend on time.

The state variables are given in both local and global co-ordinate systems as vector-vector functions which elements depend on time. This elements are described in a function space which is determined on the local co-ordinate axes. It means that the state variables can be given as a vector with function elements on the local co-ordinate axes in a function subspace. The state variables ( $\mathbf{x}(t)$ ) can be expressed by the generalized Fourier series according to the basis of the function subspace:

$$\mathbf{x}^\ell(t) = \sum_{j=1}^s \left[ \sum_{i=1}^{\infty} \alpha_{ij} P_i(t) \right] \mathbf{e}_j^\ell, \quad \alpha_{ij} \in \mathfrak{R}, \quad P_i(t) \in L^2, \quad t \in [t_1, t_2] \quad (2)$$

where  $\mathbf{e}_j^\ell$  ( $j=1, \dots, s$ ):  $j$ -th unit vector of the local coordinate system ordered to the  $\ell$ -th node,  $s$ : number of degrees of freedom at the nodes,  $P_i(t)$ :  $i$ -th element of basis of the function subspace (orthonormal polynomial system on  $[t_1, t_2]$ ).

The state variables are described on every node in the following space:  
 $F = L_1^2 \times L_2^2 \times \dots \times L_s^2$  and on the whole structure

$$F^n = (L_2^2 \times L_3^2 \times \dots \times L_s^2)^n \quad (3)$$

where  $n$  is the number of the nodes  $s$  is the number of freedoms.

To describe the elasto-plastic process a nonlinear mathematical programming problem is created in space  $F^n$  where the variables are the Fourier coefficients  $\alpha_{ij}$ . It has been proved elsewhere that the Kuhn-Tucker theorem and Wolfe's duality [1,2] are valid in space  $F^n$ .

Then the nonlinear mathematical programming problem exists in space  $F^n$  in terms of elements ( $\mathbf{x}(t)$ ). The general description of the mathematical model and the verification of the theory can be found in [2].

The solution of the nonlinear programming problem is a stationer curve. The limitation of the presented model is: at least one continuous component has to be assumed, the small displacement theory is valid, stability problems are neglected. For the numerical solution it is necessary to transform the problem into the  $\ell^2$  space where the Kuhn-Tucker theorem is valid [2] and the results are mapped back to the space  $F^n$ .

### 2.1. Nonholonom system of the elasto-plastic state

To facilitate an interpretation for a problem it is necessary to describe what type of the energies are taken into consideration. In that type of computational model the strain energy and the dissipative energy are taken into account. The state variables are defined on the Gaussian points of the finite elements.

The constraints of the **primal problem** - concerning to the intensive state variables - contains the equilibrium equations, the boundary conditions, the plastic yield conditions, the yield conditions of the dissipation capacity and switch equations which are described later. The objective function contains the sum of the complementary and the dissipative energies using their additive property.

The **dual problem** is formed by Wolfe's procedure. The results are: the compatibility equations which are expressed in terms of the extensive variables, the inequalities express the direction of the process, the boundary conditions and the sum of the different types of energies is obtained in the objective function.

In this way there is a possibility to check the primal model. The dual problem has to be clear from a physical point of view.

During the process the plastic nodes are in non-equilibrium state [5] in space  $F^n$ . The state space  $F^n$  is extended by the internal variables acting on the Gaussian points. The course of the dissipation is described by the help of internal variables in the extended space. In extended space using internal variables the non-equilibrium state becomes equilibrium one.

The primal problem is:

$$[\mathbf{B}]^T \sigma(t) \langle \rho \rangle + [\mathbf{B}]^T \mathbf{r}(t) + \mathbf{p}(t) = 0, \quad (4/a)$$

$$\text{Boundary conditions}, \quad (4/b)$$

$$f_i(\sigma_k(t), k = 1, \dots, z) \leq 0, \quad i = 1, \dots, G \quad (4/c)$$



$$\varphi_i(\mathbf{r}_k(t), k = 1, \dots, z) \leq 0, \quad i = 1, \dots, G, \quad (4/d)$$

$$f_i(\sigma_k(t), k = 1, \dots, z)r_{ik}(t) = 0, \quad i = 1, \dots, G, \quad k = 1, \dots, z, \quad (4/e)$$

$$\frac{1}{2} \sigma(t)^* [\mathbf{F}] \sigma(t) \langle \rho \rangle + \frac{1}{2} \mathbf{r}(t)^* [\mathbf{A}] \mathbf{r}(t) \rightarrow \min., \quad \forall t, t \in [t_1, t_2], \quad (4/f)$$

where  $n$  is number of the nodes,  $s$  is the freedom of nodal displacements,  $G$  is total number of the Gaussian points defined on the elements,  $z$  is the freedom of the stresses defined on the Gaussian points,  $[\mathbf{B}]^*$  is the transfer matrix of the structure (dimension is  $n.s, G.z$ ),  $\mathbf{r}(t)$  is the force type internal variable (dimension is  $n.s$ ),  $\langle \rho \rangle$  is the diagonal matrix of the Gaussian weights (dimension is  $G.z, G.z$ ),  $\sigma(t)$  is the vector of the stresses (dimension is  $G.z$ ),  $\mathbf{p}(t)$  is the vector of the external loads acting on the nodes (dimension is  $n.s$ ),  $[\mathbf{F}]$  is the flexibility matrix (dimension is  $G.z, G.z$ ),  $[\mathbf{A}]$  is the matrix of the dissipative property of the phenomena (dimension is  $G.z, G.z$ ).

The mechanical interpretations of the problem (4.a-f) are: (4.a) is the equilibrium equations concern to time-functions of stresses and force type internal variables and external forces in the extended space (the number of the equation is  $n.s$ ). The boundary conditions (4.b) are expressed by equality and/or inequality constraints. The plastic yield conditions on the Gaussian points of the structure are the inequalities (4.c). The inequalities (4.d) yield the size of the force type internal variables on the Gaussian points. The equalities (4.e) take a switch role. The objective function (4.f) is the sum of the complementary strain energy and the complementary dissipation energy.

The model (4.a-f) is "working" on the following way: A point is in elastic state, if the plastic yield condition holds the inequality in a given point. The value of the force type internal variables are zero owing to the switch conditions (4.e). In the equality equations the equal sign is valid in the space  $F^n$ . If the plastic yield condition becomes an equality in a given point the force type internal variable can be appeared in equations (4.a). This comes from the equations (4.e) for every Gaussian point and every stress freedom. The force type internal variables are not zero that is the equations (4.a) become inequalities in the space  $F^n$  (there are no force type internal variables among the usual state variables in mechanics). This expresses the fact that the point being in plastic state is in non-equilibrium state. If the constraint (4.d) is an equality the energy dissipation capacity of the material reaches its maximum in a given Gaussian point. If the constraint (4.d) is an inequality the given Gaussian point of the structure is in either elastic state or plastic state and the material is able to dissipate.

The matrix  $[\mathbf{A}]$  contains material constants which characterize the energy dissipation ability of the phenomena. Handling this type of energy is rather complicate because the lack of the material constants.

The Wolfe's dual problem of (4.a-f) is formed as follows (the dual variables  $\mathbf{u}(t)$ ,  $\lambda(t)$ ,  $\Psi(t)$  and  $\mathbf{x}(t)$  belong to equalities (4.a), inequalities (4.c), (4.d) and equalities (4.e), respectively, and the dual vectors have function elements):

$$[\mathbf{B}]\mathbf{u}(t) + [\mathbf{F}]\sigma(t)\langle \rho \rangle + \lambda(t)^* \frac{\partial f(\sigma(t))}{\partial \sigma(t)} + \mathbf{x}(t)^* \mathbf{r}(t) \frac{\partial f(\sigma(t))}{\partial \sigma(t)} = 0, \quad (5/a)$$

$$[\mathbf{B}]\mathbf{u}(t) + \psi(t)^* \frac{\partial \varphi(\mathbf{r}(t))}{\partial \mathbf{r}(t)} + \mathbf{x}(t)^* \mathbf{f}(\sigma(t)) + \mathbf{r}(t)^* [\mathbf{A}] = 0, \quad (5/b)$$

$$\text{Boundary conditions,} \quad (5/c)$$



$$\lambda(t) \geq 0, \quad \lambda(t)_i f_i(\sigma_k(t), k = 1, \dots, z) = 0, \quad i = 1, \dots, G, \quad (5/d)$$

$$\psi(t) \geq 0, \quad \psi(t)_i \varphi_i(\mathbf{r}_k(t), k = 1, \dots, z) = 0, \quad i = 1, \dots, G, \quad (5/e)$$

$$\begin{aligned} & \frac{1}{2} \langle \rho \rangle \sigma(t)^* [\mathbf{F}] \sigma(t) + \frac{1}{2} \mathbf{r}(t)^* [\mathbf{A}] \mathbf{r}(t) + \mathbf{u}(t)^* [\mathbf{B}] \sigma(t) \langle \rho \rangle + \mathbf{u}(t)^* \mathbf{r}(t) + \\ & + \mathbf{u}(t)^* \mathbf{p}(t) + \lambda(t)^* \mathbf{f}(\sigma(t)) + \psi(t)^* \varphi(\mathbf{r}(t)) + \mathbf{x}(t)^* \mathbf{r}(t) f(\sigma(t)) \rightarrow \max. \\ & \forall t, t \in [t_1, t_2], \end{aligned} \quad (5/f)$$

where  $\mathbf{u}(t)$  is the vector of the displacements,  $\lambda(t)$  and  $\Psi(t)$  are the plastic and dissipative multipliers, respectively.

The mechanical interpretations of the problem (5.a-f) are: (5.a) expresses the compatibility between the displacements and the strains (the number of the equations is  $G \cdot z$ ). (5.b) equations express the compatibility due to the dissipation (the number of the equations is  $G \cdot z$ ). The boundary conditions are saved in original form (5.c). The inequalities (5.d) and (5.e) give the sign constrain of the plastic and the dissipation multiplier, respectively. They show the direction of the process. (5.f) is the objective function.

Let's see the dual problem in details: Taking out the gradient vector from (5.a) the following equation system is obtained:

$$[\mathbf{B}] \mathbf{u}(t) + [\mathbf{F}] \sigma(t) \langle \rho \rangle + (\lambda(t)^* + \mathbf{x}(t)^* \mathbf{r}(t)) \frac{\partial \mathcal{F}(\sigma(t))}{\partial \sigma(t)} = 0$$

The 1st. member means the nodal displacements of the element, the 2nd. member is the displacements due to the elastic stresses, the 3rd. member gives the displacements due to the plastic strains in the direction of the gradient of the plastic yield conditions. The measure of these displacements origins from two parts: one is the plastic multiplier and the other is the dissipation in the plastic zone.

In elastic state the equation system (5.b) does not exist since the force type internal variables are zero. In plastic state the equation system (5.b) consists of the following parts: 1st. member means the nodal displacements of the element, 2nd. member contains the displacements on the direction of the gradients of the dissipation conditions, 3rd. member is zero in plastic state since the function value of the plastic yield conditions become zero, 4th. member express the displacement type internal variables.

The interpretation of the objective function can be obtained after the following rearranging: multiplying the equations (5.a) and (5.b) by their dual variables and substituting into the objective function and rewriting the maximum problem into a minimum the following function

$$\begin{aligned} \text{is obtained:} \quad & \frac{1}{2} \langle \rho \rangle \sigma(t)^* [\mathbf{F}] \sigma(t) + \frac{1}{2} \mathbf{r}(t)^* [\mathbf{A}] \mathbf{r}(t) - \mathbf{u}(t)^* \mathbf{p}(t) - \\ & - \lambda(t)^* f(\sigma(t)) - \psi(t)^* \varphi(\mathbf{r}(t)) - \mathbf{x}(t)^* \mathbf{r}(t) f(\sigma(t)) + \mathbf{x}(t)^* \mathbf{r}(t) f(\sigma(t)) + \\ & + \lambda(t)^* \frac{\partial \mathcal{F}(\sigma(t))}{\partial \sigma(t)} \sigma(t) + \mathbf{x}(t)^* \mathbf{r}(t) \frac{\partial \mathcal{F}(\sigma(t))}{\partial \sigma(t)} \sigma(t) + \psi(t)^* \frac{\partial \varphi(\mathbf{r}(t))}{\partial \mathbf{r}(t)} \mathbf{r}(t) \rightarrow \min. \end{aligned}$$

The 1st. member expresses the strain energy, the 2nd. member is the dissipative energy and the 3rd. member gives the energy due to the external loads. The 4th. and 5th. members are zero, if the inequalities (4.c) less then zero the correspondent dual variables are zero due to the constraints (5.d) and (5.e), if they become equalities the value of the yield functions are zero. The sum of the 6th. and 7th. member is zero. The 8th., 9th. and the 10th. member are zero if

functions  $f$  and  $\varphi$  are the homogeneous functions of the stresses and the force type internal variables [21]. The dual of the problem (4.a-f) is after the above-mentioned modifications:

$$[\mathbf{B}]\mathbf{u}(t) + [\mathbf{F}]\boldsymbol{\sigma}(t)\langle\rho\rangle + (\boldsymbol{\lambda}(t)^* + \mathbf{x}(t)^* \mathbf{r}(t)) \frac{\partial \mathbf{f}(\boldsymbol{\sigma}(t))}{\partial \boldsymbol{\sigma}(t)} = 0, \quad (6/a)$$

$$[\mathbf{B}]\mathbf{u}(t) + \boldsymbol{\psi}(t)^* \frac{\partial \varphi(\mathbf{r}(t))}{\partial \mathbf{r}(t)} + \mathbf{x}(t)^* \mathbf{f}(\boldsymbol{\sigma}(t)) + \mathbf{r}(t)^* [\mathbf{A}] = 0, \quad (6/b)$$

$$\text{Boundary conditions,} \quad (6/c)$$

$$\boldsymbol{\lambda}(t) \geq 0, \quad \boldsymbol{\lambda}(t)_i f_i(\boldsymbol{\sigma}_k(t), k = 1, \dots, z) = 0, \quad i = 1, \dots, G, \quad (6/d)$$

$$\boldsymbol{\psi}(t) \geq 0, \quad \boldsymbol{\psi}(t)_i \varphi_i(\mathbf{r}_k(t), k = 1, \dots, z) = 0, \quad i = 1, \dots, G, \quad (6/e)$$

$$-\frac{1}{2}\langle\rho\rangle \boldsymbol{\sigma}(t)^* [\mathbf{F}]\boldsymbol{\sigma}(t) - \frac{1}{2}\mathbf{r}(t)^* [\mathbf{A}]\mathbf{r}(t) + \mathbf{u}(t)^* \mathbf{p}(t) \rightarrow \min. \quad (6/f)$$

The objective function in problem (6) expresses the principle of energy conservation.

## 2.2. Computational model of the dynamic problem without damping

It is supposed that the mass and the flexibility of the structures are time independent and the structure has no motion at the time  $t_1 = 0$ . The external loads are decomposed into two parts: the first one ( $\mathbf{p}_e(t)$ ) makes the deformations and the second one ( $\mathbf{p}_d(t)$ ) moves the structure. The rate of  $\mathbf{p}_e(t)$  and  $\mathbf{p}_d(t)$  is unknown and it is changed in time. The internal forces are divided in a similar way. The dissipative forces as internal variables are handled as it happened in previous models.

The primal problem of the dynamic analysis can be written on the following way:

$$\mathbf{p}_e(t) + \mathbf{p}_d(t) - \mathbf{p}(t) = 0, \quad (7/a)$$

$$[\mathbf{B}]^* \boldsymbol{\sigma}_e(t)\langle\rho\rangle + [\mathbf{B}]^* \mathbf{r}(t) + \mathbf{p}_e(t) = 0, \quad (7/b)$$

$$[\boldsymbol{\Psi}]^* \boldsymbol{\sigma}_d(t)\langle\rho\rangle + \mathbf{p}_d(t) = 0, \quad (7/c)$$

$$\text{Boundary conditions,} \quad (7/d)$$

$$f_i(\boldsymbol{\sigma}_{ek}(t), k = 1, \dots, z) \leq 0, \quad i = 1, \dots, G \quad (7/e)$$

$$\varphi_i(\mathbf{r}_k(t), k = 1, \dots, z) \leq 0, \quad i = 1, \dots, G, \quad (7/f)$$

$$f_i(\boldsymbol{\sigma}_{ek}(t), k = 1, \dots, z) r_{ik}(t) = 0, \quad i = 1, \dots, G, \quad k = 1, \dots, z, \quad (7/g)$$

$$\frac{1}{2} \boldsymbol{\sigma}_e(t)^* [\mathbf{F}]\boldsymbol{\sigma}_e(t)\langle\rho\rangle + \frac{1}{2} \mathbf{r}(t)^* [\mathbf{A}]\mathbf{r}(t) + \frac{1}{2} \int_0^t \boldsymbol{\sigma}(t)^* dt [\mathbf{M}]^{-1} \int_0^t \boldsymbol{\sigma}(t) dt \langle\rho\rangle \rightarrow \min., \quad \forall t, t \in [t_1, t_2], \quad (7/h)$$

where  $n$  is number of the nodes,  $s$  is the freedom of nodal displacements,  $G$  is total number of the Gaussian points defined on the elements,  $z$  is the freedom of the stresses defined on the Gaussian points,  $[\mathbf{B}]^*$  is the transfer matrix of the structure (dimension is  $n.s$ ,  $G.z$ ),  $\mathbf{r}(t)$  is the force type internal variable (dimension is  $n.s$ ),  $[\boldsymbol{\Psi}]^*$  is the transfer matrix of



the structure for dynamic effect (dimension is n.s, G.z),  $\langle \rho \rangle$  is the diagonal matrix of the Gaussian weights (dimension is G.z, G.z),  $\sigma_e(t)$  is the vector of the stresses which are related with the deformations (dimension is G.z),  $\sigma_D(t)$  is the vector of the stresses which are related with the motions (dimension is G.z),  $\mathbf{p}(t)$  is the vector of the external loads acting on the nodes (dimension is n.s),  $[\mathbf{F}]$  is the flexibility matrix (dimension is G.z, G.z),  $[\mathbf{A}]$  is the matrix of the dissipative property of the phenomena (dimension is G.z, G.z),  $[\mathbf{M}]^{-1}$  is the inverse of the mass matrix (dimension is G.z, G.z).

The unknowns are:  $\mathbf{p}_e(t)$ ,  $\sigma_e(t)$ ,  $\sigma_D(t)$  and  $r(t)$ .

The mechanical interpretations of the problem (7.a-h) are: (7.a) is the decomposition of the external forces, (7.b) and (7.c) are the equilibrium equations concern to time-functions of stresses and force type internal variables and external forces in the extended space, respectively (the number of the equation is n.s). The boundary conditions (7.d) are expressed by equality and/or inequality constraints. The plastic yield conditions on the Gaussian points of the structure are the inequalities (7.e). The inequalities (7.f) yield the size of the force type internal variables on the Gaussian points. The equalities (7.g) take a switch role. The objective function (7.h) is the sum of the complementary strain energy, the complementary dissipation energy and the kinetic energy due to the motion. The above mentioned form of the kinetic energy can be obtained in the following way: According to the Newton II. law it is known that

$$p_d(t) = M \cdot \ddot{u}(t) \quad (8)$$

Integrating both side of eq. (8) one can obtain  $\int_0^t p_d(t) dt = M \cdot \dot{u}(t)$ . The kinetic energy can be

expressed as

$$\frac{1}{2} M \cdot \dot{u}^2(t) = \frac{1}{2} \int_0^t p_d(t) dt \cdot M^{-1} \cdot \int_0^t p_d(t) dt. \quad (9)$$

Using internal forces in eq. (9) the kinetic energy can be written in the following way:

$$\frac{1}{2} \int_0^t \sigma_d(t) dt \cdot M^{-1} \cdot \int_0^t \sigma_d(t) dt. \quad (10)$$

### 3. Numerical example

As an illustration of the proposed computational methods, consider a three-supported beam with time dependent loading. The data of the structure can be seen in Fig. 1. The loads act on the nodes. The loads act on the nodes. The functions of the external loads are given by

$$F(t) = 8 + 7t + 6t^2 + 5t^3.$$

The loads are approximated by the Legendre polynomial system up to four members. The structure is divided into 6 members with 7 nodes. The unknowns are the moments and the shear forces at the Gaussian points of the members.

On the basis of the general form, the statically admissible internal forces are determined by the equilibrium equations and force boundary conditions. The objective function contains the



complementary potential energy, dissipative energy and kinetic energy. The displacement boundary conditions are taken into consideration in the objective function.

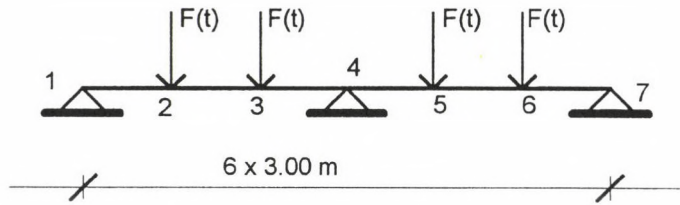


Fig. 1.

The total and the decomposed (elastic and dynamic parts) loads are seen in Fig. 2. in totally elastic case.

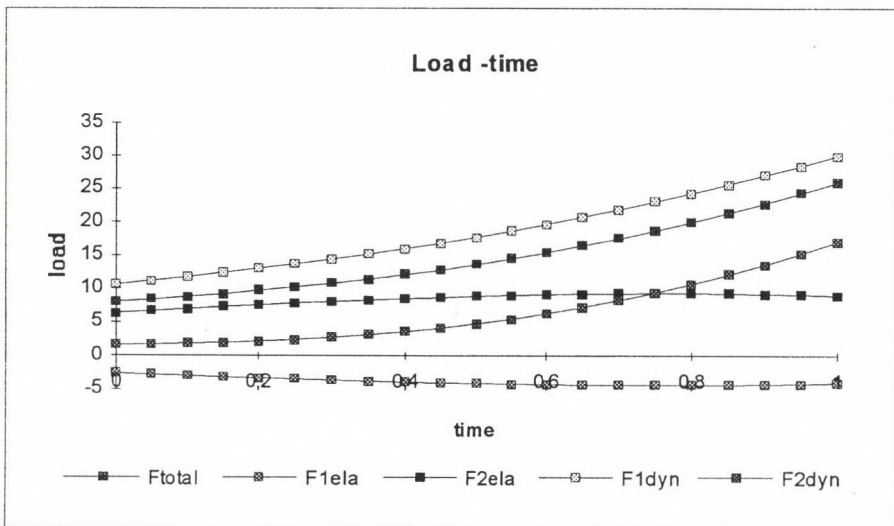


Fig. 2.

To solve this problem we used a nonlinear mathematical programming system. The results can be seen in Figs 3-4. The bending moments (elastic and dynamic parts) above of the middle support can be seen in Fig. 3. and the elastic energy and the dynamic energy of the structure are shown in Fig. 4. One can see that this method and the traditional algorithms give the same result.

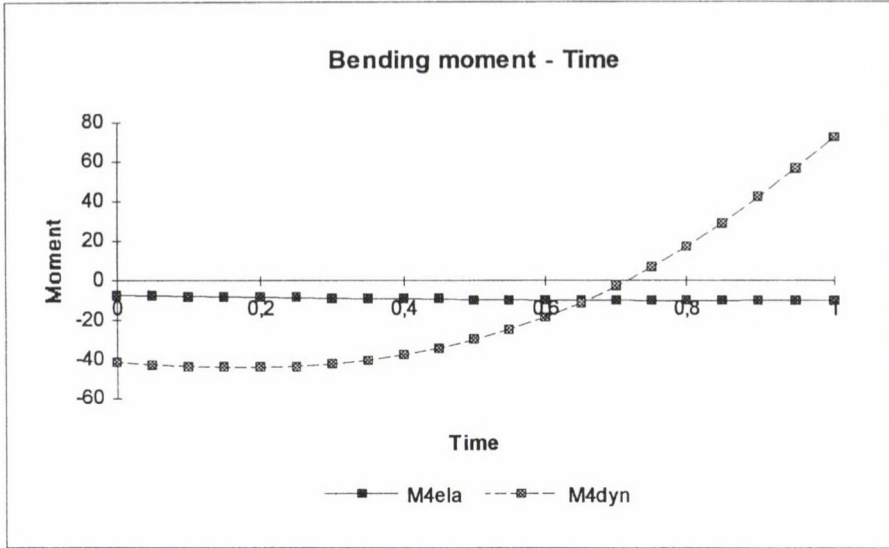


Fig. 3.

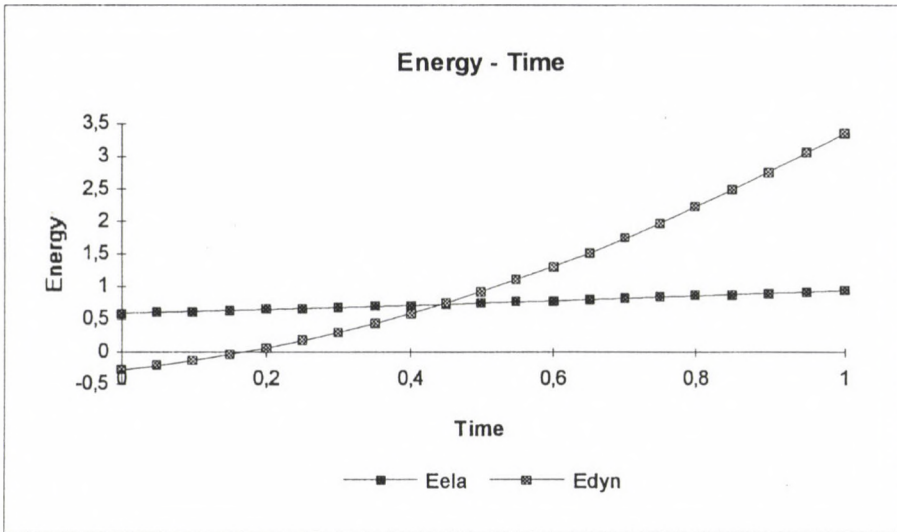


Fig. 4.

If there are some restrictions for the internal forces they give inequalities in the mathematical programming problem. In our example, the Huber-Mises-Hencky yield condition is used. In Figs 5-7 one can follow the results of the case when the plastic hinges during the process make dissipative energy as well.

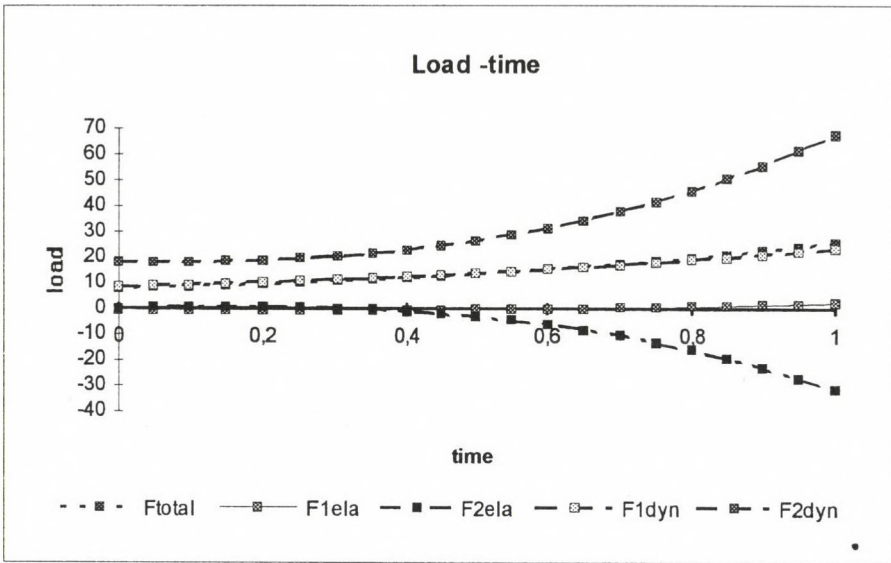


Fig.5.

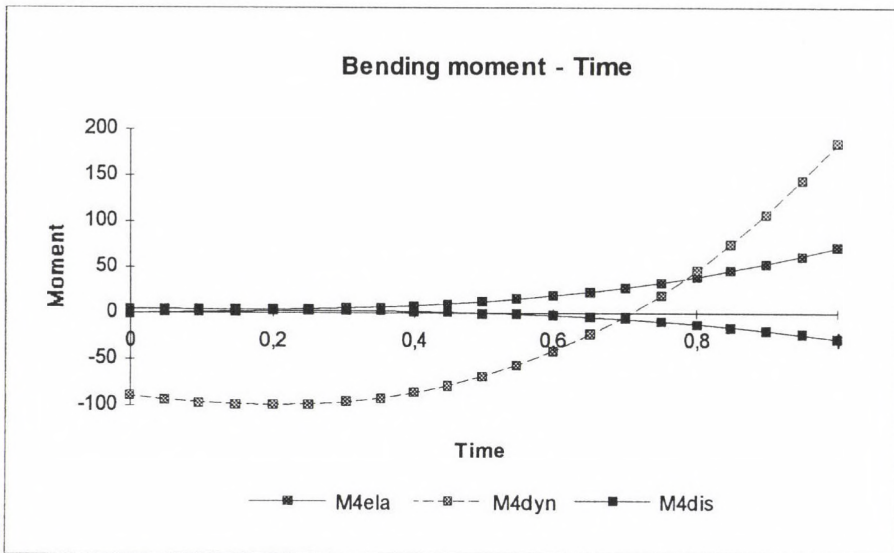


Fig.6.



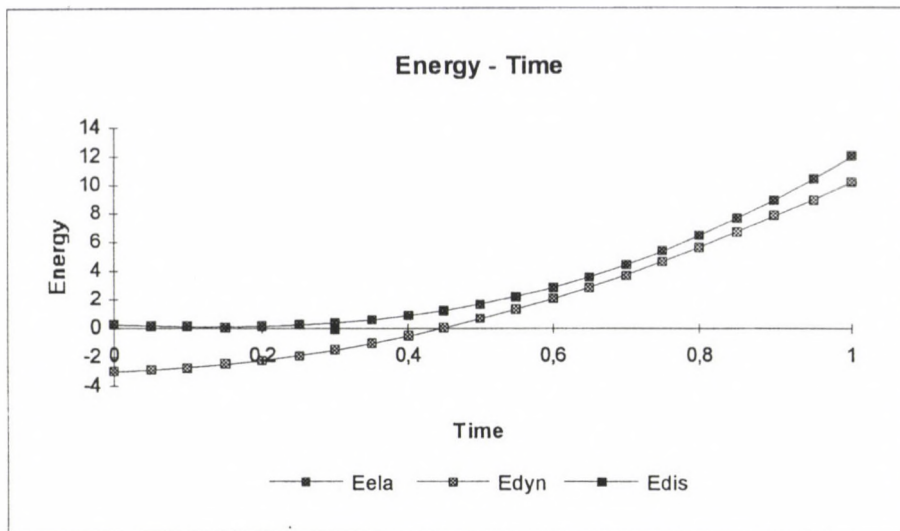


Fig. 7.

#### 4. Conclusion

The elastic, dynamic and dissipative energies are taken into consideration. According to our numerical experiments first the static after the dynamic energies are developed in the structures and finally the material is changed causing dissipation. The rates of the energies depend on the material constants and the mass of the structure.

The theorem mentioned above and algorithms of the nonlinear mathematical programming were used for the numerical solution.

#### Acknowledgements

The present study was supported by The Hungarian Scientific Research Found (OTKA #T014272 and OTKA #F014288).

#### REFERENCES

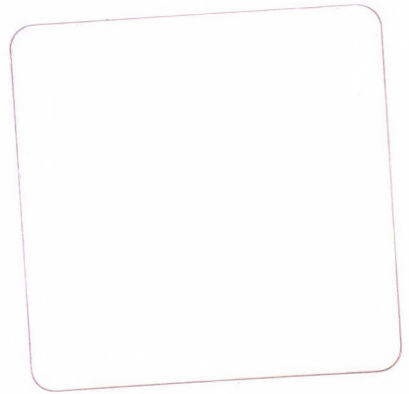
1. Bazarra, M.S.-Shetty, C.M.: Nonlinear Programming, Theory and Algorithms. John Wiley & Sons, New York 1979
2. Bhakta, P.C.-Roychandhuri, S.: Optimization in Banach Spaces, Journal of Math. Analysis and Applications, 134 (1988) 460-470
3. Cohn, M.Z.- Maier, G. (Eds.): Engineering Plasticity by mathematical programming. Pergamon Press. New York 1979
4. Kaliszky, S.: Plasticity, Theory and Engineering Application. Elsevier Science Publishers, Amsterdam 1989
5. Muschik, S.: Non-equilibrium Thermodynamics with Applications to Solids, Springer-Verlag, Wien, New York 1993
6. Roth, C.P.-Bird, W.W.: Internal Variable Formulations for the Plastic Analysis of Plane Frames, Engineering Structures, 17, 3 (1995) 214-220

7. Sayegh, A.H.- Rubinstein, A.F.: Elasto-plastic Analysis by Quadratic Programming, Journ. of Eng. Mech., Div. ASCE, EM6, (1972) 1547-1572
8. Sawczuk, A.: Mechanics and Plasticity of Structures. Ellis Horwood Ltd., Chichester 1989
9. Stricklin, J.A.-Haisler W.E.: Formulation of Solution Procedures for Nonlinear Structural Analysis, Computer & Structures 7, (1977) 125-136
10. Vásárhelyi, A.-Lógó, J.: Analysis of the Structures for Dynamic Effects with Mathematical Programming, Acta Technica Acad.Sci.Hung., 102(3-4), (1989) 379-390
11. Vásárhelyi, A.-Lógó, J.: Mathematical models of processes in mechanics, Structural Optimization, 8, 2-3, (1994) 138-145
12. Vásárhelyi, A.: A Mathematical Model for Processes of Structure Analysis, Computer Assisted Mechanics and Engineering Sciences, 3, (1996) 297-316.
13. Vértés, Gy.: Structural Dynamics. Elsevier. Amsterdam - Oxford 1985



PRINTED IN HUNGARY  
Akadémiai Nyomda, Martonvásár





## NOTICE TO CONTRIBUTORS

Papers in English are accepted on condition that they have not been previously published or accepted for publication.

Only original papers will be published and a copy of the Publishing Agreement will be sent to the authors of papers accepted for publication. Manuscripts will be processed only after receiving the signed copy of the agreement.

Manuscripts in two copies (the original type-written copy plus a clear duplicate one) complete with figures, tables, and references should be sent to

*Acta Technica*  
Nádor u. 7. I. 118  
Budapest, Hungary  
H-1051

Although every effort will be made to guard against loss, it is advised that authors retain copies of all material which they submit. The editorial board reserves the right to make editorial changes.

*Manuscripts* should be typed double-spaced on one side of good quality paper with proper margins and bear the title of the paper and the name(s) of the author(s). The full postal address(es) of the author(s) should be given in a footnote on the first page. An abstract of 50 to 100 words should precede the text of the paper. The approximate locations of the tables and figures should be indicated in the margin. An additional copy of the abstract is needed. Russian words and names should be transliterated into English.

*References.* Only papers closely related to the author's work should be referred to. The citations should include the name of the author and/or the reference number in brackets. A list of numbered references should follow the end of the manuscript.

References to periodicals should mention: (1) name(s) and initial(s) of the author(s); (2) title of the paper; (3) name of the periodical; (4) volume; (5) year of publication in parentheses; (6) numbers of the first and last pages. Thus: 5. Winokur, A.—Gluck, J.: Ultimate strength analysis of coupled shear walls. *American Concrete Institute Journal* 65 (1968) 1029-1035

References to books should include: (1) author(s)' name; (2) title; (3) publisher; (4) place and year of publication. Thus: Timoshenko, S.—Gere, J.: *Theory of Elastic Stability*. McGraw-Hill Company. New York, London 1961

*Illustrations* should be selected carefully and only up to the necessary quantity. Black-and-white photographs should be in the form of glossy prints. The author's name and the title of the paper together with the serial number of the figure should be written on the back of each print. Legends should be brief and attached on a separate sheet. Tables, each bearing a title, should be self-explanatory and numbered consecutively.

Authors will receive proofs which must be sent back by return mail.

Authors will receive 50 reprints free of charge.





301.162

ACTA

TECHNICA

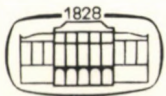
ACADEMIAE SCIENTIARUM HUNGARICAE

EDITOR-IN-CHIEF: P. MICHELBERGER

21.

VOLUME 107  
NUMBERS 3-4

CIVIL ENGINEERING – C/9



AKADÉMIAI KIADÓ, BUDAPEST 1995-96

ACTA TECHN. HUNG. ATSHA8

# ACTA TECHNICA

A JOURNAL OF THE HUNGARIAN ACADEMY OF SCIENCES

---

## CENTRAL EDITORIAL BOARD

T. CZIBERE, K. GÉHER, L. KOLLÁR, P. MICHELBERGER (EDITOR-IN-CHIEF),  
A. LÉVAI, J. PROHÁSZKA, K. REMÉNYI, J. SZABÓ  
GY. CZEGLÉDI (MANAGING EDITOR)

## EDITORIAL COMMITTEE FOR CIVIL ENGINEERING (SERIES C)

A. BÉNYEI, J. FARKAS, ZS. GÁSPÁR, L. KOLLÁR (CHAIRMAN),  
L. SOMLYÓDY

---

*Acta Technica* publishes original papers, preliminary reports and reviews in English, which contribute to the advancement of engineering sciences.

*Acta Technica* is published by

AKADÉMIAI KIADÓ  
H-1117 Budapest, Prielle Kornélia u. 4

### *Subscription information*

Orders should be addressed to

AKADÉMIAI KIADÓ  
H-1519 Budapest, P.O. Box 245

Subscription price for Volume 107 (1995–96) in 4 issues US\$ 98.00, including normal postage, airmail delivery US\$ 20.00.

---

*Acta Technica* is abstracted/indexed in Applied Mechanics Reviews, Current Contents–Engineering, Technology and Applied Sciences, GeoRef Information System, Science Abstracts.

---

© Akadémiai Kiadó, Budapest 1996

## CONTENTS

<i>Bhargava, R. R. – Tyagi, N. K.</i> : Unbounded plate with two circular-arc cracks with coalesced plastic zones closed by variable pressure .....	149
<i>Bojtár, I. – Gálos, M. – Rechterisz, Á.</i> : Determination of fracture mechanical properties on rock's boring core material .....	163
<i>Bódi, I. – Fűzy, J. – Klopka, Z.</i> : Deformation and stress analysis of hyperbolic paraboloid shells taking into account realistic boundary conditions .....	175
<i>Ding Dajun</i> : Semiempirical theory of masonry strength .....	181
<i>Dombay, G. – Piriou, Ph. – Dukan, S. – Kiene, L.</i> : Modeling the influence of environmental parameters and residence time on microbial drinking water quality in distribution systems .....	195
<i>Dulácska, E.</i> : The effect of ductility in combination with the P- $\Delta$ effects, to be taken into consideration in the seismic design of compressed structures .....	219
<i>Hankó, Z.</i> : On the division of the kinetic energy (power) content in turbulent flow. An approximate analytical review .....	249
<i>Keshava, M. K. – Shesha, P. M. N.</i> : Application of a new optimization procedure for the design of a proportional weir .....	265
<i>Klopka, Z. – Ládi, N.</i> : Numerical examination of the stiffening system of tall buildings by using object-oriented programming .....	277
<i>Kozák, I.</i> : Continuum mechanics. Summary of comparison of solid and fluid mechanics .....	303
<i>Kurutz, M.</i> : Tangent modulus for nonsmooth materials .....	339
<i>Nándori, E. – Kovács, M.</i> : Possibilistic estimations of imperfections in structural analysis .....	369
<i>Széll, M.</i> : The underground waterproofing of the University Sport Centre of Budapest .....	395

### BOOK REVIEW

<i>Iványi, A.</i> : Hysteresis Models in Electromagnetic Computation ( <i>G. Fodor</i> ) .....	405
--	-----





# UNBOUNDED PLATE WITH TWO CIRCULAR-ARC CRACKS WITH COALESCED PLASTIC ZONES CLOSED BY VARIABLE PRESSURE

Bhargava, R. R.\* and Tyagi, N. K.\*\*

\**Department of Mathematics, University of Roorkee, Roorkee-247 667, India*

\*\**Govt College, Roorkee-247 667, India*

(Received: 15 October, 1997)

The problem of an infinite plate cut along two circular-arc cracks is investigated in this paper. The two circular cracks are situated along the circumference of a circle. The cracks open in Mode I deformations on account of remotely applied biaxial tension at infinity. The plastic zones are formed ahead of the tips of the cracks and the plastic zones formed at interior tips of the two cracks get coalesced. These plastic zones are closed by normal variable stress distribution acting perpendicular to the rims of the plastic zones. Problem is solved using complex variable technique. Analytical expressions are derived for the load required to close the plastic zones and for crack face opening displacements. A qualitative analysis is carried out to study the behaviour of load required for closure of cracks and crack opening displacement with respect to affecting parameters crack length, plastic zone length and crack radius.

## 1. Introduction

General formulation of two-dimensional circular-arc crack problem is given by Milne-Thomson [1], Muskhelishvili [2] and others. The problem of a circular crack lying along the interface of a circular inclusion embedded in an infinite solid has been investigated by Toya [3] using complex variable formulation. Zhang [4] making use of the basic theorem of the pole point and residue gave the general solution to a circular-arc crack problem in opening, sliding and tearing modes in an infinite plane. The plane problem of two arc cracks placed along the circumference of a circle in an infinite elastic plate is studied theoretically by Gdoutos et al. [5]. Complex potentials are used to solve the curved circular crack problem in two-bonded half-planes made of dissimilar materials by Chen and Hasebe [6]. Dugdale model solution is obtained for a circular-arc crack in an infinite plate by Bhargava and Kumar [7]. Crack opening displacement and stress-intensity factors caused by a concentrated load outside a circular crack have been evaluated by Karapetian and Hanson [8] in terms of elementary functions. Chen [9] obtained the solution of circular-cracked plate problem by using Bueckner's weight function formulation. A circular disk containing a concentric arc crack is studied by Xu [10] using dislocation pile-ups and singular integral equation techniques. Related algo-

rithms for computing Mode I and Mode II stress-intensity factors of the curvilinear crack are obtained in this paper. The interaction between the crack tips of a circular-arc crack is studied by Shiah and Tsai [11] devising a new conformal mapping. The stress concentration around a circular interface crack between two dissimilar elastic planes are analysed by Murase et al. [12].

### 2. Fundamental equations

As is well-known the components of stresses  $P_{ij}(i, j = r, \theta)$  and displacements  $u_i(i = r, \theta)$  may be written in terms of two complex potentials  $\Phi(z)$  and  $\Omega(z)$  as

$$P_{rr} + iP_{r\theta} = \Phi(z) + \Omega(R^2 / \bar{z}) + \bar{z} \left[ \left( \bar{z} / R^2 \right) - (1 / z) \right] \overline{\Psi(z)}, \tag{1}$$

$$2\mu \frac{\partial}{\partial \theta} \{ (u_r + iu_\theta) e^{i\theta} \} = iz \left[ \kappa \phi(z) - \Omega(R^2 / \bar{z}) - \bar{z} \left[ \left( \bar{z} / R^2 \right) - (1 / z) \right] \overline{\Psi(z)} \right], \tag{2}$$

where

$$\Psi(z) = \frac{R^2}{z^2} \overline{\Omega(z)} - \frac{R^2}{z^2} \Phi(z) - \frac{R^2}{z} \Phi'(z), \tag{3}$$

and  $z = x + iy = re^{i\theta}$ . Bar over the function denotes its complex conjugation. Shear modulus is denoted by  $\mu$ ,  $\kappa = (3-4\nu)$  for plane strain case and  $\kappa = (3-\nu)/(1+\nu)$  for generalized plane strain. Poisson’ ratio is denoted by  $\nu$ .

Consider an infinite two-dimensional elastic plate cut along  $n$  arcs  $L_i$  ( $i = 1, 2, \dots, n$ ) with end points  $a_i, b_i$  ( $i = 1, 2, \dots, n$ ). These arcs lie along the circumference of a circle of radius  $R$ . The union of these arcs is denoted by  $L = \bigcup_{i=1}^n L_i$ .

Let the rims of  $L$  be subjected to the stress distribution  $P_{rr}^\pm, P_{r\theta}^\pm$  and  $P_{\theta\theta} = 0$ . The infinite boundary of the plate is subjected to no loads. Using conditions at the rims of the crack and equation (1) following two Hilbert problems are obtained.

$$[\Phi(t) - \Omega(t)]^+ - [\Phi(t) - \Omega(t)]^- = 2q(t);$$

on  $L$

$$[\Phi(t) + \Omega(t)]^- + [\Phi(t) + \Omega(t)]^+ = 2p(t). \tag{4}$$

Under the assumption  $r \lim_{r \rightarrow R} \left\{ e^{-i\theta} \left[ \left( r / R^2 \right) - (1 / r) \right] \Psi(z) \right\} = 0. \tag{5}$



Any point on the rims of the crack is denoted by  $t = Re^{i\theta}$  and

$$\begin{aligned}
 q(t) &= \frac{1}{2} (P_{rr}^+ - P_{rr}^-) + \frac{i}{2} (P_{r\theta}^+ - P_{r\theta}^-); \\
 p(t) &= \frac{1}{2} (P_{rr}^+ + P_{rr}^-) + \frac{i}{2} (P_{r\theta}^+ + P_{r\theta}^-).
 \end{aligned}
 \tag{6}$$

The superscript + denotes the value of the function at any point  $t$  on  $L$  when approached from inner ( $r < R$ ) region and superscript - denotes the value of the function when point  $t$  is approached from outer ( $r > R$ ) region.

In absence of the body forces and no stresses at infinity the solution of equation (4) may be written as

$$\Phi(z) - \Omega(z) = \frac{1}{\pi i} \int_L \frac{q(t)}{(t-z)} dt + D_0 \tag{7}$$

$$\Phi(z) + \Omega(z) = \frac{1}{\pi i X(z)} \int_L \frac{X(t)p(t)}{(t-z)} dt + \frac{1}{X(z)} \left\{ P_n(z) + \frac{D_1}{z} + \frac{D_2}{z^2} \right\} \tag{8}$$

where

$$X(z) = \prod_{i=1}^n (z - a_i)^{1/2} (z - b_i)^{1/2} \tag{9}$$

and

$$P_n(z) = C_0 z^n + C_1 z^{n-1} + C_2 z^{n-2} + \dots + C_n. \tag{10}$$

The constants  $D_i$  ( $i = 0, 1, 2$ ) and  $C_j$  ( $j = 0, 1, 2, \dots, n$ ) are determined using boundary conditions of the problem.

Stress-intensity factor at a tip  $z = a_i$  of a crack  $L_i$  ( $i = 0, 1, 2, \dots, n$ ) may be computed using formula

$$K_1 - iK_2 = 2\sqrt{2\pi} \lim_{z \leftarrow a_i} \left\{ (z - a_i)^{1/2} \Phi(z) \right\}. \tag{11}$$

Introducing

$$V(z) = e^{i\theta} \{ u_r(z) + iu_\theta(z) \}, \tag{12}$$

and substituting the value of  $\Phi(z)$  and  $\Omega(z)$  from equations (7) and (8) into equation (2) one obtains

$$V'(t) = \frac{\kappa + 1}{2\mu} \{ \Phi^-(t) - \Phi^+(t) \}. \tag{13}$$

Integrating which  $V(t)$  is obtained. Consequently the crack face opening displacement  $u_r$  is obtained.

### 3. Statement of problem

An infinite homogeneous, isotropic elastic-perfectly plastic plate is bounded by  $xoy$  plane. The plate is cut along two circular-arc cracks lying along the circumference of a circle of radius  $R$  with centre  $(0,0)$ . These cracks are denoted by  $L_1$  and  $L_2$  and occupy the arcular interval  $c_1: (R,\gamma)$  to  $a_1:(R,\beta)$  and  $b_1:(R,2\pi-\beta)$  to  $d_1:(R,2\pi-\gamma)$ , respectively. Uniform biaxial tension applied at infinite boundary of the plate causes the opening of faces of the crack in Mode I type deformations. The plastic zones develop, on account of this opening of faces of the crack, ahead of the tips of the crack. The two plastic zones developed at two adjacent interior tips of the two cracks get coalesced. Thus, the entire arcular region of inter-crack distance forms a plastic zone denoted by  $\Gamma_1$ : lying from  $(R,2\pi-\gamma)$  to  $(R,\gamma)$ . Other two plastic zones developed at the tips  $a_1$  is denoted by  $\Gamma_2$  and occupies the region  $(R,\beta)$  to  $(R,\alpha)$  and the third plastic zone  $\Gamma_3$  at the tip  $b_1$  occupies the interval  $(R, 2\pi-\alpha)$  to  $(R,2\pi-\beta)$ .

To arrest the crack from further opening each rim of the plastic zone  $\Gamma_i$  ( $i = 1,2,3$ ) is subjected to variable normal stress distribution  $P_{rr} = \sigma_{ye} \sin \theta$ ,  $P_{r\theta} = P_{\theta\theta} = 0$ . Thus, the crack is arrested from further opening. The yield point stress is denoted by  $\sigma_{ye}$  and  $\theta$  is the angle varying along arcular length of the crack. The entire configuration is depicted in Fig. 1.

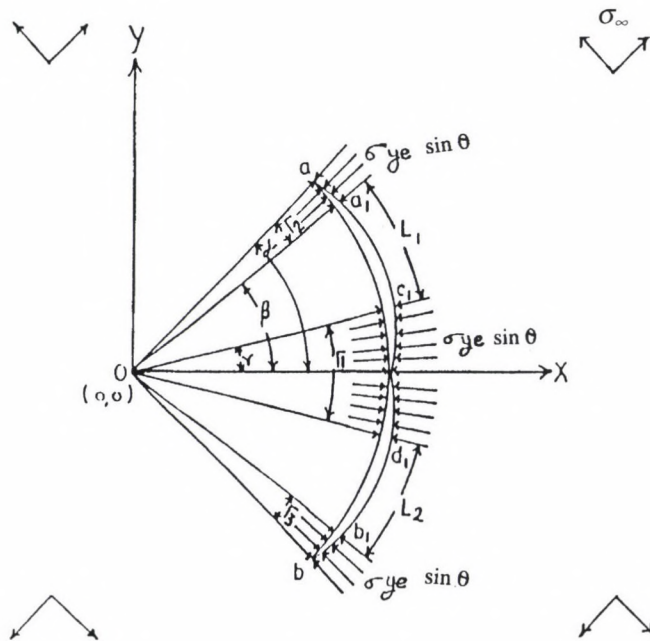


Fig. 1. The configuration

#### 4. Solution of the problem

The solution of above-stated problem is obtained by superposing the solutions of two-component problems contributing toward the stress singularity. These problems are appropriately derived from the above problem and are termed as *Problem I* and *Problem II*. These two problems are discussed below.

##### *Problem I*

An infinite stress-free plate is bounded by  $xoy$  plane. The plate is cut along a circular-arc crack  $S$  lying along the circumference of a circle of radius  $R$  and centre  $(0,0)$ . The crack  $S$  occupies the arcular interval from  $(R, -\alpha)$  to  $(R, \alpha)$ . The crack  $S$  is formed by the union of  $\Gamma_i$  ( $i = 1, 2, 3$ ) and  $L_j$  ( $j = 1, 2$ ) i.e.  $S = \Gamma_3UL_2U\Gamma_1UL_1U\Gamma_2$ . The boundary conditions of the problem are:

- (i) Rims of the crack  $S$  are acted upon by the uniform stresses  $P_{rr} = \sigma_\infty$  and  $P_{r\theta} = P_{\theta\theta} = 0$ .
- (ii) No stresses acting at infinity.
- (iii) Displacements are single valued at the rims of crack.

Solution of above problem is written directly using Muskhelishvili [2] as

$$\Phi_1(z) = \frac{\sigma_\infty}{(3 - \cos \alpha)} \left[ \frac{(z - R \cos \alpha)}{X(z)} - 1 \right], \quad (14)$$

$$\Omega_1(z) = \frac{\sigma_\infty}{(3 - \cos \alpha)} \left[ \frac{(z - R \cos \alpha)}{X(z)} - (2 - \cos \alpha) \right], \quad (15)$$

where

$$X(z) = \left( z^2 - 2R \cos \alpha + R^2 \right)^{1/2}. \quad (16)$$

The opening mode stress-intensity factor at  $(K'_I)_a$  for *Problem I* at the tip  $z = a = Re^{i\alpha}$  is obtained substituting  $\Phi_1(z)$  for  $\Phi(z)$  in equation (11). This may be written as

$$(K'_I)_a = \sigma_\infty \sqrt{(2\pi R \sin \alpha)} / (3 - \cos \alpha). \quad (17)$$

##### *Problem II*

An infinite elastic-perfectly plastic plate is bounded by  $xoy$  plane. The plate is cut along a single arc  $S$  (defined above in *Problem I*) along the circle of radius  $R$  with centre  $(0,0)$  and  $S = \Gamma_3UL_2U\Gamma_1UL_1U\Gamma_2$ . The boundary conditions of the problem are:



- (i) Each rim of  $I_i$  ( $i = 1,2,3$ ) is subjected to normal stress distribution  $P_{rr} = \sigma_{ye} \sin \theta$ ,  $P_{r\theta} = P_{\theta\theta} = 0$ .
- (ii) Each rim of  $L_i$  ( $i=1,2$ ) are stress free.
- (iii) No stresses are prescribed at infinity.
- (iv) Displacements are single valued at the rims of the crack  $S$ .

Using boundary conditions (i) and (ii) and equation (1) following two Hilbert problems are obtained:

$$[\Phi_{II}(t) - \Omega_{II}(t)]^+ - [\Phi_{II}(t) - \Omega_{II}(t)]^- = 0, \tag{18}$$

$$[\Phi_{II}(t) + \Omega_{II}(t)]^+ + [\Phi_{II}(t) + \Omega_{II}(t)]^- = 2\sigma_{ye} \sin \theta, \tag{19}$$

on  $\Gamma = \bigcup_{i=1}^3 \Gamma_i$  and  $t$  is any point on  $\Gamma$ . The subscript II denotes that the function refers to *Problem II*.

The solution of equations (18) and (19) may be written using equations (7) and (8) as

$$\begin{aligned} \Phi_{II}(z) = & \frac{D_0}{2} + \frac{\sigma_{ye}}{4\pi X(z)} \left\{ \left[ \frac{G_1}{2} \left( \frac{b_1}{R} - \cos \alpha \right) - \frac{RG_2}{2} (1 + \cos^2 \alpha) - 2C_1 \right] \right. \\ & + \left( \frac{G_1}{R} - G_2 \cos \alpha + 2C_0 \right) z + \frac{G_2}{R} z^2 - \frac{R^2 G_0}{z} \left. \right] - \frac{i\sigma_{ye}}{2\pi} \left( \frac{z}{R} - \frac{R}{z} \right) \left[ -\tan^{-1} \sqrt{\frac{(a-z)(b_1-b)}{(z-b)(b_1-a)}} \right. \\ & \left. - \tan^{-1} \sqrt{\frac{(a-z)(c_1-b)}{(z-b)(c_1-a)}} - \frac{\pi}{2} + \tan^{-1} \sqrt{\frac{(a-z)(d_1-b)}{(z-b)(d_1-a)}} + \tan^{-1} \sqrt{\frac{(a-z)(a_1-b)}{(z-b)(a_1-a)}} \right] \\ & = \Omega_{II}(z) + D_0, \end{aligned} \tag{20}$$

where  $X(z)$  is the same as defined in equation (16) and

$$D_0 = -\sigma_{ye} \left( \frac{G_1}{2R} + G_7 \right), \tag{21}$$

$$D_1 = 0, \tag{22}$$

$$D_2 = 0, \tag{23}$$

$$C_0 = G_7 \tag{24}$$

$$C_1 = \left[ \frac{G_5 - G_3 G_7}{G_2} \right] \tag{25}$$

The constants  $G_i$  ( $i = 0, 1, 2, \dots, 9$ ) and  $H_i$  ( $i = 1, 2$ ) denote

$$G_0 = 2i \left[ -\tan^{-1} \sqrt{\frac{a(b_1 - b)}{b(a - b_1)}} - \tan^{-1} \sqrt{\frac{a(c_1 - b)}{b(a - c_1)}} - \frac{\pi}{2} + \tan^{-1} \sqrt{\frac{a(d_1 - b)}{b(a - d_1)}} + \tan^{-1} \sqrt{\frac{a(a_1 - b)}{b(a - a_1)}} \right],$$

$$G_1 = 4Ri \left[ H_2 \sin \frac{\gamma}{2} - H_1 \sin \frac{\beta}{2} \right],$$

$$G_2 = -2i \left[ -\tan^{-1} \sqrt{\frac{b_1 - b}{a - b_1}} - \tan^{-1} \sqrt{\frac{c_1 - b}{a - c_1}} - \frac{\pi}{2} + \tan^{-1} \sqrt{\frac{d_1 - b}{a - d_1}} + \tan^{-1} \sqrt{\frac{a_1 - b}{a - a_1}} \right],$$

$$G_3 = (G_1 + G_2 R \cos \alpha),$$

$$G_4 = \frac{1}{2R} \left[ \frac{G_1}{2} \left( \frac{b_1}{R} - \cos \alpha \right) + \frac{R G_2}{2} (1 + \cos^2 \alpha) - R G_0 \cos \alpha + \bar{G}_1 \right],$$

$$G_5 = - \left[ R G_0^2 + \left\{ \frac{G_1}{2} + \left( \frac{b_1}{R} - \cos \alpha \right) - \frac{R G_2}{2} (1 + \cos^2 \alpha) \right\} G_2 + \left( \frac{G_1}{R} - G_2 \cos \alpha \right) G_3 \right. \\ \left. + \frac{G_2}{2R} \left\{ (b_1 + 3R \cos \alpha) G_1 + R^2 (3 \cos^2 \alpha - 1) G_2 \right\} \right],$$

$$G_6 = - \left[ \frac{(G_1 + \bar{G}_1)}{2R} + G_4 + \frac{G_5}{R G_2} \right],$$

$$G_7 = \left[ \frac{R \bar{G}_2 (R \bar{G}_2 G_6 - \bar{G}_3 G_6 - 2 R \bar{G}_2 \bar{G}_2)}{(R^2 G_2 \bar{G}_2 - R G_2 \bar{G}_3 - R \bar{G}_2 G_3 + G_3 \bar{G}_3) - 4 R^2 G_2 \bar{G}_2} \right],$$

$$G_8 = \log_e \left[ \frac{(1 - \cos \alpha e^{-i\beta} + 2H_1 e^{-i\beta/2})(1 - \cos \alpha e^{i\gamma} + 2H_2 e^{i\gamma/2})(1 - \cos \alpha e^{i\alpha}) e^{2i(\beta - \gamma - \alpha)}}{(1 - \cos \alpha e^{-i\gamma} + 2H_2 e^{-i\gamma/2})(1 - \cos \alpha e^{i\beta} + 2H_1 e^{i\beta/2})(1 - \cos \alpha e^{-i\alpha})} \right],$$

$$G_9 = \log_e \left[ - \frac{\left\{ e^{-i\beta/2} (e^{-i\beta/2} + 2H_1) - \cos \alpha \right\} \left\{ e^{i\gamma/2} (e^{i\gamma/2} + 2H_2) - \cos \alpha \right\}}{\left\{ e^{-i\gamma/2} (e^{-i\gamma/2} + 2H_2) - \cos \alpha \right\} \left\{ e^{i\beta/2} (e^{i\beta/2} + 2H_1) - \cos \alpha \right\}} \right], \quad (26)$$

$$H_1 = \sin^{1/2} \left( \frac{\alpha + \beta}{2} \right) \sin^{1/2} \left( \frac{\alpha - \beta}{2} \right), \quad H_2 = \sin^{1/2} \left( \frac{\alpha + \gamma}{2} \right) \sin^{1/2} \left( \frac{\alpha - \gamma}{2} \right). \quad (27)$$

The opening mode stress-intensity factor  $(K_I'')_a$  for this problem substituting  $\Phi_{II}(z)$  from equation (20) for  $\Phi(z)$  into equation (11) at tip  $z = a = Re^{i\alpha}$  as

$$\begin{aligned} (K_I'')_a = & \frac{\sigma_{ye}}{2(\pi R \sin \alpha)^{1/2}} \operatorname{Re} \left[ \left( \frac{1-i}{\sqrt{2}} \right) \left\{ \frac{G_1}{2} \left( \frac{b_1}{R} - \cos \alpha \right) + \frac{RG_2}{2} (1 + \cos^2 \alpha) \right. \right. \\ & \left. \left. + 2C_1 + \left( \frac{G_1}{R} - G_2 \cos \alpha + 2C_0 \right) a + \frac{G_2}{R} a^2 - \frac{R^2 G_0}{a} \right\} \right]. \end{aligned} \quad (28)$$

### 5. Plastic zone length and crack opening displacement

Plastic zone length for the original problem is obtained by superposing appropriately the stress-intensity factors obtained at the tip  $z = a = Re^{i\alpha}$ , using the condition

$$(K_I')_a + (K_I'')_a = 0, \quad (29)$$

$$\begin{aligned} & \sigma_{\infty} \cdot 2^{3/2} \cdot (\pi R \sin \alpha) - \sigma_{ye} (3 - \cos \alpha) \operatorname{Re} \left[ \left( \frac{1-i}{\sqrt{2}} \right) \left\{ \frac{G_1}{2} \left( \frac{b_1}{R} - \cos \alpha \right) \right. \right. \\ & \left. \left. - \frac{RG_2}{2} (1 + \cos^2 \alpha) + 2C_1 + \left( \frac{G_1}{R} - G_2 \cos \alpha + 2C_0 \right) a + \frac{G_2}{R} a^2 - \frac{R^2 G_0}{a} \right\} \right] = 0. \end{aligned} \quad (30)$$

Determining  $\alpha$  from above equation, the plastic zone's arcular length is then calculated evaluating  $|R(\alpha - \beta)|$ .

The Dugdale model crack opening displacement  $u_R$  at the crack face is calculated by using

$$u_R(z) = \frac{\sigma_{\infty}}{\sigma_{ye}} \cdot u_R'(z) - u_R''(z), \quad (31)$$

$u_R'$  COD for *Problem I*, obtained using equation (13) substituting value of  $\Phi_I(z)$  for  $\Phi(z)$  from equation (14)

$$u_R' = \operatorname{Re} \left[ -R(z-a)^{1/2} (z-b)^{1/2} / \{z(3 - \cos \alpha)\} \right]. \quad (32)$$



Similarly  $u_R''$  COD for *Problem II*, substituting  $\Phi_{II}(z)$  for  $\Phi(z)$  from equation (20) into equation (13) and integrating one obtains

$$\begin{aligned}
 u_R'' = \operatorname{Re} \left[ \left( \frac{1}{4\pi} \right) \left\{ G_8 R \log_e \left( \frac{2R^2 - 2zR \cos \alpha + 2RX(z)}{z} \right) + \right. \right. \\
 + \left\{ \frac{G_1}{2} \left( \frac{b_1}{R} - \cos \alpha \right) - \frac{RG_9}{2} (1 + \cos^2 \alpha) + 2C_1 \right\} \log_e \{ 2X(z) + 2z - 2R \cos \alpha \} + \\
 + \left( \frac{G_1}{R} - G_9 \cos \alpha + 2C_0 \right) \left\{ X(z) + R \cos \alpha \log_e \{ 2X(z) + 2z - 2R \cos \alpha \} \right\} + \\
 + \frac{G_9}{R} \left\{ \left( \frac{z + 3R \cos \alpha}{2} \right) X(z) + \frac{R^2}{2} (3 \cos^2 \alpha - 1) \log_e \{ 2X(z) - 2z - \right. \\
 \left. \left. - 2R \cos \alpha \} \right\} \right] \left( \frac{R}{z} \right) \quad (33)
 \end{aligned}$$

Now calculating crack tip opening displacement at the actual crack tips  $z = a_1 = Re^{i\beta}$  and  $z = c_1 = Re^{i\gamma}$  is trivially and obtained substituting  $u_R'$  and  $u_R''$  into equation (31) for prescribed  $\sigma_\infty/\sigma_{ye}$ ,  $\beta$ ,  $R$  and  $\gamma$ .

## 6. Results

The behaviour of load required to close the plastic zones with respect to the increase in crack angle and crack radius is plotted in Figs 2 and 3, respectively. It may be observed from Fig. 2 as the crack length is increased, keeping the plastic zone size fixed than less load is required for arresting the crack. Also it is seen that further less load is required if the distance between the two cracks is increased keeping crack and plastic zone lengths fixed. Figure 3 shows the variation of load ratio (load applied at infinity/yield point stress) with respect to crack radius. As the radius of the circle on which the cracks lie is increased the effect of the crack on each other diminishes.

The variation of crack opening displacement (COD) at the interior tip of the crack versus the interior plastic zone is shown in Fig. 4. The crack opens more if the plastic zone size is increased, as expected. Figure 5 shows the variation of COD at the exterior tip of the crack versus length of the exterior plastic zone. Here can also be that for bigger plastic zone size the crack opens more.

In Fig. 6 COD at interior tip is plotted against increasing crack length. It is interesting to note that if the cracks are moved apart increasing the crack radius, the crack opening variations smoothens out.

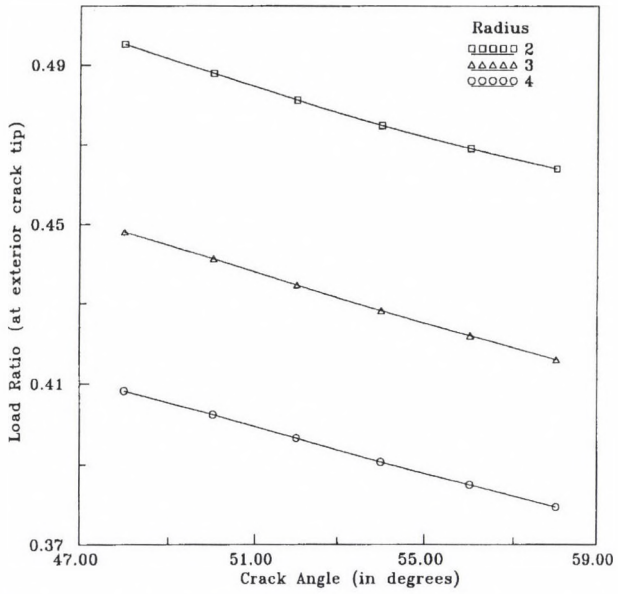


Fig. 2. Variation of load ratio at the exterior tip of the crack versus crack radius

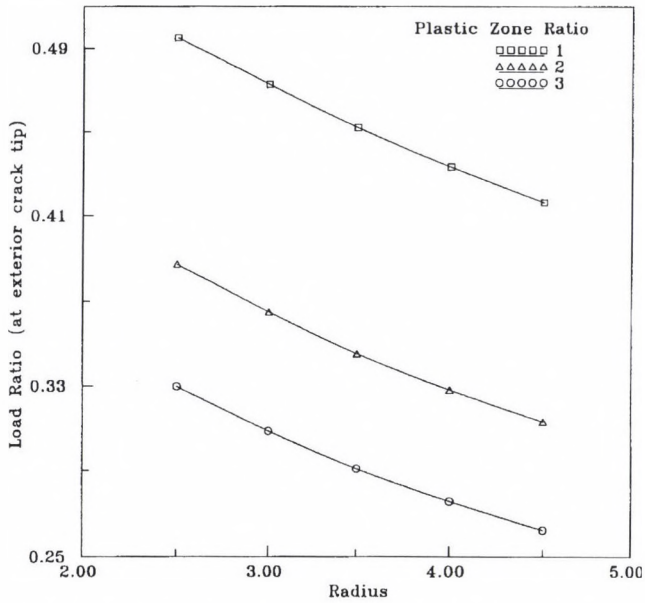


Fig. 3. Variation of load ratio at the exterior tip of the crack versus crack radius

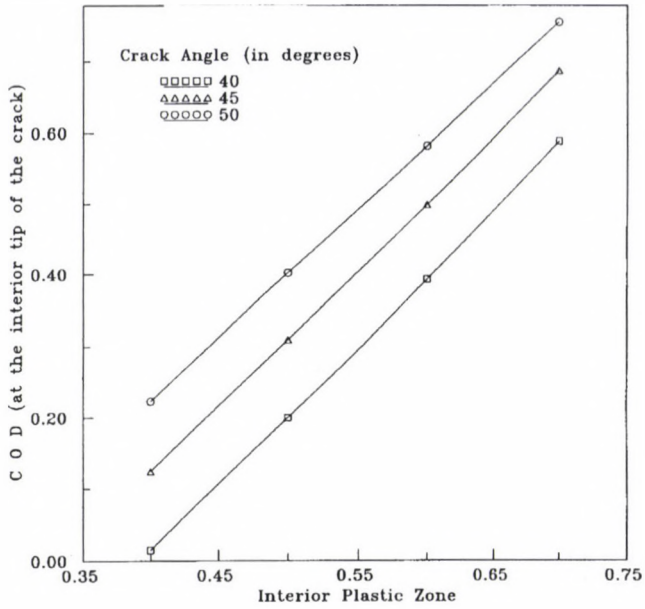


Fig. 4. Variation of crack opening displacement at the interior tip of the crack versus interior plastic zone length

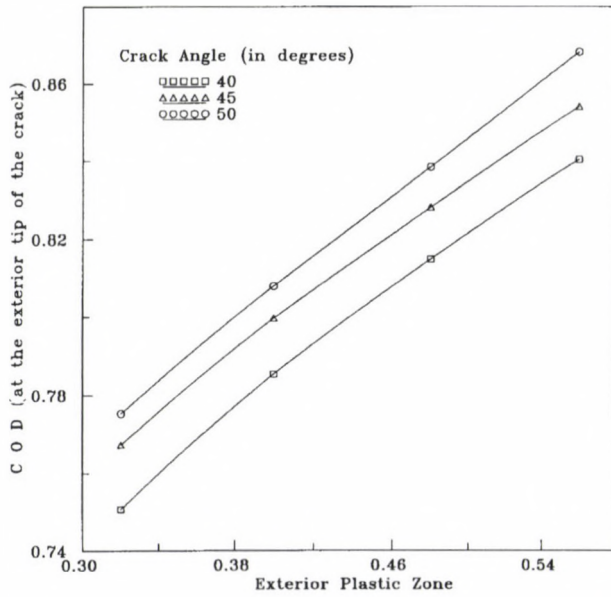


Fig. 5. Variation of crack opening displacement at the exterior tip of the crack versus exterior plastic zone length



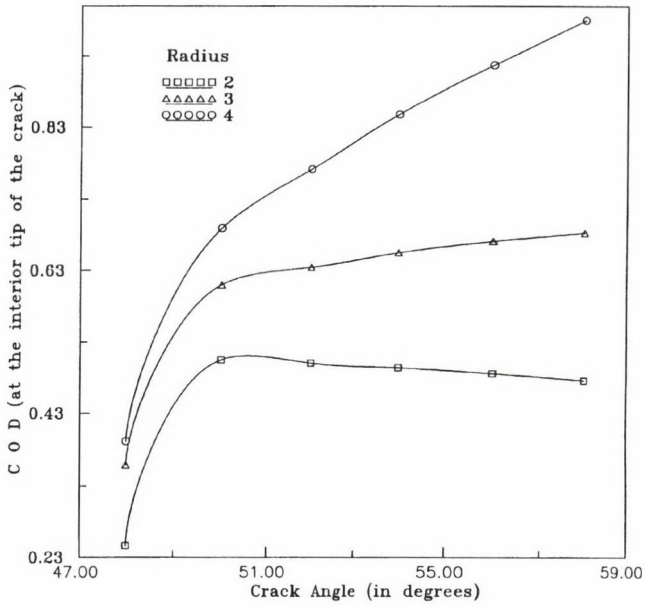


Fig. 6. Variation of crack opening displacement at the interior tip of the crack versus crack angle

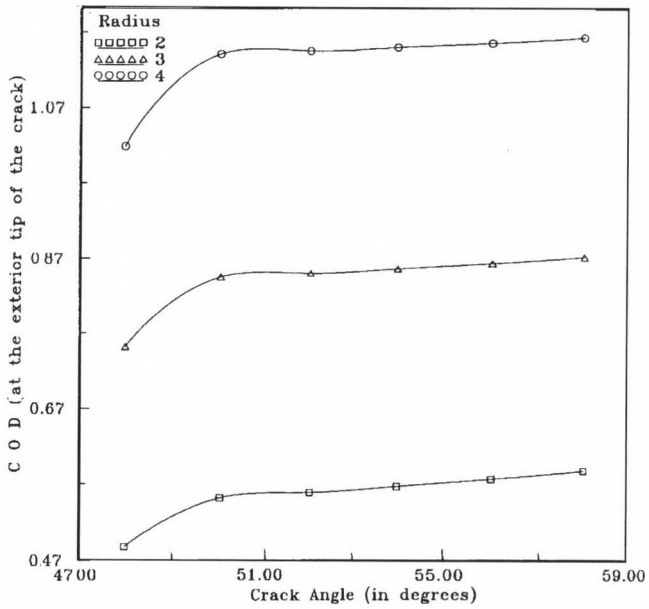


Fig. 7. Variation of crack opening displacement at the exterior tip of the crack versus crack angle

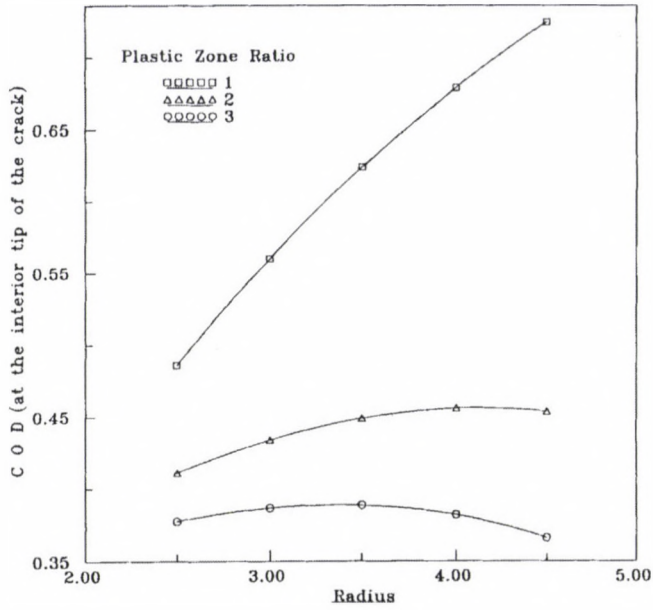


Fig. 8. Variation of crack opening displacement at the interior tip of the crack versus crack radius

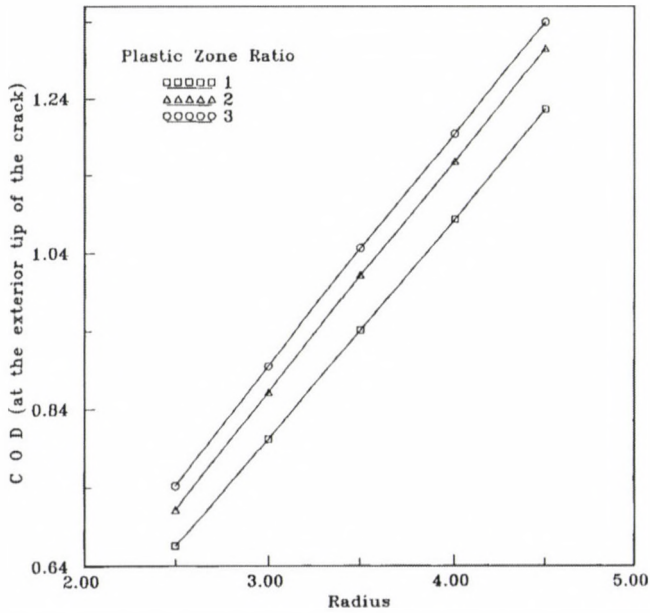


Fig. 9. Variation of crack opening displacement at the exterior tip of the crack versus crack radius

The effect of increasing crack length on COD at the exterior tip is plotted in Fig. 7. It may be observed that at the exterior tip, since the effect of the existing crack in neighbourhood is less, the COD variation is not as turbulent as in case of interior tip of the crack.

Figures 8 and 9 show the behaviour of COD at interior and exterior tips as the crack radius (the radius of the circle on circumference of which the cracks lie) is increased. In Fig. 8 the COD pattern is such that the COD at two interior tips of the two cracks tend to coalesce. The different plastic zone ratio, for which the graphs are drawn, is the ratio between exterior plastic zone to interior plastic zone. Figure 9 shows the same variation at exterior tip of the crack. The COD at the exterior tip increases as the size of crack radius is increased.

### Acknowledgement

Authors are grateful to Professor R. D. Bhargava [Senior Professor (retd.), Indian Institute of Technology, Bombay] for his invaluable advice during the course of this work.

### References

1. Milne-Thomson, L. M.: *Plane Elastic System*. 2nd ed. Springer-Verlag, 1960
2. Muskhelishvili, N. I.: *Some Basic Problems of Mathematical Theory of Elasticity*. 3rd ed., English Translation by Radok, J. R., P. Noordhoff, Groningen, 1953
3. Toya, M.: *Int. J. Fracture* 9 (1973), 463–470
4. Zhang, X. S.: *Engg. Fracture Mech.* 34 (1989), 275–282
5. Gdoutos, E. E.–Kattis, M. A.–Argyrokorta, V. K.–Koutsougeras, T. J.–Papanelopoulou, M. C.: *Engg. Fracture Mech.* 39 (1991), 671–681
6. Chen, Y. Z.–Hasebe, N.: *Theoretical and Applied Fracture Mechanics* 17 (1992), 189–196
7. Bhargava, R. R.–Kumar R.: *Engg. Fracture Mech.* 46 (1993), 265–273
8. Karapetian, E. N.–Hanson, M. T.: *International Journal of Solids and Structures* 31 (1994) 2035–2052
9. Chen, Y. Z.: *Engg. Fracture Mech* 52 (1995), 173–178
10. Xu, Yong Li: *Int. J. of Solids and Structures* 32 (1995), 2023–2040
11. Shiah, Y. C.–Tsai, Y. M.: *Proceedings of the 11th Conference on Engineering Mechanics*. 1996, Part 1 (of 2)
12. Murase Y.–Nakagawa, K.–Hachiya, T.–Duan, S. J.: *Engg. Fracture Mech.* 53 (1996), 661–673



# DETERMINATION OF FRACTURE MECHANICAL PROPERTIES ON ROCK'S BORING CORE MATERIAL

Bojtár, I.\*–Gálos, M.\*\*–Rechtorisz, Á.\*

\**Technical University of Budapest, Department of Structural Mechanics, Műegyetem rkp. 3, H-1111 Budapest, Hungary*

\*\**Technical University of Budapest, Department of Engineering Geology, Műegyetem rkp. 3, H-1111 Budapest, Hungary*

(Received: 11 March 1997)

The determination of rock fracture toughness is one of the most significant problems in fracture mechanics. This paper suggests a method for determining the stress intensity factor ( $K_{Ic}$ ) of limestone beams under three point bending.

This method is based on a model with new geometry. The regulations valid at present oblige us to use rectangular specimens, though these specimens are not always easy to be produced at rocks because of the material structure, and the number of failed beams is relatively high during creating. It seems to be more advantageous to choose an experimental form which is anyway available in the course of the boring sampling: the half-cylindrical specimen, simply made from cylindrical boring core with cutting it into two and with an artificial notch, is easier to be produced and has less loss than the rectangular one, and the three point bending subjection can be practicable in the same way.

This paper presents the laboratory and finite element investigations for the rectangular and half-cylindrical specimens, and finally suggests analytical formula for the calculation of the stress intensity factor based on the new geometry.

## 1. Introduction

A significant field of rock mechanics' researches is the determination of material properties, which is necessary for the study of discipline. The results of considerations connected with laboratory investigations and research-technology will appear in everyday practice in a short time helping in measuring and structure planning and supporting the assessmental and qualification work respectively. More and more laboratory researches turn their attention to examinations connected with fracture and failure. The determination of stress intensity factor, which is the most useful for qualification inquiry among the fracture mechanical characteristics, and its technological problems arise as research exercise.

It's very important that what kind of specimens we use or can use for laboratory experiments beside the questions: what, with what and how. The available sample is determinant from the point of research-technology. That's why one can only apply such characteristics for the assessment of material properties that contain the correction factors coming from shape and size effects.

The most frequent sample of the laboratory rock mechanics' investigations is the core originating from borings. The deepening borings provide high-level knowledge of rock environment of engineering constructions. Large number of rock variations has to be assessed and qualified with the available core material.

The boring core material provides the economical application of the standard cylindrical specimens in research-technological respect. Therefore – and not only from theoretical considerations – in rock mechanical testing works we use so-called standard cylindrical specimens, which can be made simply from the boring core with cutting into two. If necessary cylindrical specimens can be produced from different samples got in other ways (e.g. from rock mass) with using laboratory boring machine.

For determining the stress intensity factor a research has been standardized, that makes the inclusion of the  $K_{Ic}$  possible with bending of the notched boring core (Ouchterlony, 1989). We also have possibility to produce so-called “modified disc specimens” with notched disks cut from the boring core and to take down this significant material property with the help of these samples. Czoboly and his co-authors (Czoboly, 1986) gave suggestion for using these boring cores.

At geological research-borings for determining the petrographic and settlemental properties the cores often have to be cut into two along their longitudinal axis. These properties can be more easy to analyze at the cut surface, and one of the core's half can be applied for testing and the other for retaining. The notched sample half applied for testing is suitable for the determination of the critical stress intensity factor.

Naturally, the new shape of specimen – contrary to the traditional bending model – raises a large number of technical questions. This paper will give answers for them with demonstrating the results of the theoretical and experimental research. The purpose of our research work was to make the new specimens – which are available in large quantities and can be produced easily – suitable for determining fracture mechanical properties and thus they can be fitted into the system of the rock mechanics' laboratory research.

## 2. Laboratory experiments

Three point bending tests were done on different cross-sectional specimens in the laboratory at the Department of Engineering Geology at Technical University of Budapest. The specimens were made from sandstone, dense and coarse limestone. At each stone first *rectangular*, then *half-cylindrical* specimens were broken. For examination series such rocks were chosen, that was suitable for the homogeneous and isotropic conditions in examination respect. The examined rocks were:

- *dense limestone (I)* (Budapest, Pesthidegkút, Hungary) with carbonate tissue and containing calcite- and clay-veins in some places,
- *coarse limestone* (Vraca, Bulgaria), fine-grained, carbonatic rock with beet pot-ashed bonding,
- *sandstone* (Maulbronn, Germany), auburn, fine-grained rock with glued tissue and silicate bonding,



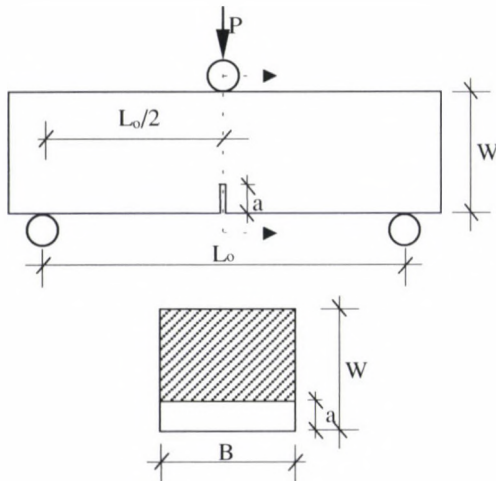
– *dense limestone (II)* (Vác, Hungary) fine-grained, compact rock with micro-crystal tissue and containing calcite-knots.

The artificial notch causing crack and “collecting” tensile stress concentration was the same at all types of material and geometrical forms. The measurements were done on laboratory *air-dried* rock physical specimens.

The geometrical characteristics, data and the type of loading can be seen in Fig. 1 at *rectangular* and Fig. 2 at *half-cylindrical* specimens. It is important to mention that rectangular specimens were made from the same cylindrical cores as we used for the production of the half-cylindrical samples with cutting them into two along their longitudinal axis, therefore the geometrical relationship between the two different cross-sectional models can be directly and simply established. In every case the examination apparatus was a universal compress machine type SZFI (Anyagvizsgáló Készülékek Gyára, Budapest).

The basic strength properties (Young’s modulus, Poisson’s ratio) were also gained from the laboratory experiments of applied rocks. Table 1 shows these values for the four materials.

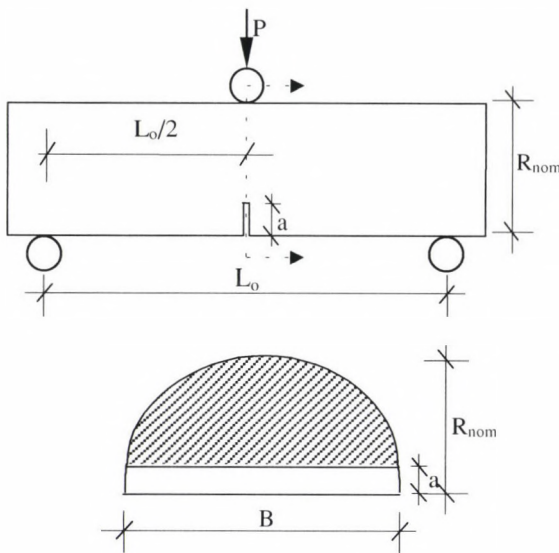
In the course of laboratory measurements we determined the critical load needed for bending failure. To ensure the quasi-static nature of the loading procedure, load-increase was done in such a way that stress-increase speed remained below 1 Mpa/sec, a fixed value in the standard concerning bending. The gained critical loads of rectangular and half-cylindrical specimens can be seen in Table 2.



Sample	$L_0$ [mm]	W [mm]	B [mm]	a [mm]
PI-1	100	33.5	36.7	4.36
PI-2	100	39.6	45.4	6.61
PI-3	100	38.7	39.4	3.63
PI-4	100	39.6	38.9	4.12
PI-5	100	39.8	39.2	4.90
PI-6	100	39.9	39.2	1.54

Fig. 1. The geometrical characteristics of the rectangular specimen and the type of loading





Sample	$L_0$ [mm]	$R_{nom}$ [mm]	B [mm]	a [mm]
CI-1	100	43.3	83.8	2.42
CI-2	100	37.8	83.3	1.48
CI-3	100	38.0	83.2	2.79
CI-4	100	27.3	55.1	1.87
CI-5	100	25.3	54.6	2.06
CI-6	100	26.9	58.9	1.64
CI-7	100	29.9	59.3	2.24
CII-1	100	37.5	63.4	3.80
CII-2	100	30.8	63.0	2.53
CII-3	100	31.8	63.3	3.58
CII-4	100	32.0	63.5	3.06
CII-5	100	30.0	63.3	3.78
CII-6	100	30.9	62.6	2.93
CII-7	100	34.3	71.5	4.69
CII-8	100	21.3	53.9	4.96
CII-9	100	26.1	55.1	3.99
CII-10	100	21.6	54.2	5.63
CII-11	100	26.5	55.1	4.01

Fig. 2. The geometrical characteristics of the half-cylindrical specimen and the type of loading

Table 1  
Strength properties

Sample	Material	Young's modulus (E) [N/mm <sup>2</sup> ]	Poisson's ratio ( $\nu$ ) [-]
PI-1, PI-2; CI-1 – CI-3	Dense limestone (I)	26340	0.20
PI-3, PI-4; CI-4, CI-5	Coarse limestone	8430	0.25
PI-5, PI-6; CI-6, CI-7	Sandstone	7830	0.25
CII-1 – CII-11	Dense limestone (II)	35870	0.24

### 3. Numerical solutions

The regulations based on fracture mechanics' theoretical methods give possibility for the determination of stress intensity factor limit with the help of critical load determined by laboratory measurements. These relations are based on the Irwin's theory, that modified the Kolosov–Muskhelishvili–Westergaard's disc-solving executed with complex stress functions. These discs had an artificial sharp crack and supposed infinite size. Originally Irwin restricted the solution valid for the whole disc to the immediate surroundings of the crack tip. The stress intensity factor introduced in the course of this re-

duction is used for mathematical description of *singularity* and physical description of *stress concentration* (see the general works listed at References, e.g.: Broek, Hahn or Anderson), and its limit can be considered as a material constant at a given geometrical crack face.

The situation at real structures is much more complicated than as mentioned above: boundary conditions provided by finite size and complicated geometrical form is basically different from the assumptions of the original solution. In practical cases this difficulty is absolved in such a way that size- and shape-dependent correction factors are initiated to modify the results derived from theoretical assumptions. This leads to the consequence that new geometrical correction factors have to be determined for each new case (excellent summary can be found from the stress intensity factors of the different real sized elements and from the geometrical correction functions detailed the factors' solutions in e.g. the volumes of "Stress Intensity Factors Handbook" published in Japan and the U.S.A.).

Table 2  
Critical load values

Sample	Material	Critical load ( $P_c$ ) [N]
PI-1	Dense limestone (I)	2450
PI-2	Dense limestone (I)	5000
PI-3	Coarse limestone	410
PI-4	Coarse limestone	380
PI-5	Sandstone	1150
PI-6	Sandstone	1460
CI-1	Dense limestone (I)	7650
CI-2	Dense limestone (I)	7400
CI-3	Dense limestone (I)	6850
CI-4	Coarse limestone	190
CI-5	Coarse limestone	140
CI-6	Sandstone	760
CI-7	Sandstone	890
CII-1	Dense limestone (II)	2920
CII-2	Dense limestone (II)	3400
CII-3	Dense limestone (II)	2280
CII-4	Dense limestone (II)	3060
CII-5	Dense limestone (II)	2630
CII-6	Dense limestone (II)	3900
CII-7	Dense limestone (II)	4000
CII-8	Dense limestone (II)	900
CII-9	Dense limestone (II)	1720
CII-10	Dense limestone (II)	850
CII-11	Dense limestone (II)	1750

Table 3  
Stress intensity factors of rectangular specimens

Sample	Material	Stress intensity factor ( $K_{Ic}$ ) [N/mm <sup>3/2</sup> ]
PI-1	Dense limestone (I)	58.12
PI-2	Dense limestone (I)	83.86
PI-3	Coarse limestone	6.31
PI-4	Coarse limestone	5.99
PI-5	Sandstone	19.23
PI-6	Sandstone	14.56

$$K_{Ic} = \frac{3 \cdot P_c \cdot L_o}{2 \cdot B \cdot W^2} \cdot \sqrt{\pi \cdot a} \cdot Y(\alpha), \quad \text{where} \quad \alpha = \frac{a}{W},$$

$$Y(\alpha) = \frac{1.99 - \alpha \cdot (1 - \alpha) \cdot (2.15 - 3.93 \cdot \alpha + 2.70 \cdot \alpha^2)}{(1 + 2 \cdot \alpha) \cdot (1 - \alpha)^{3/2}}$$

These effects issued from the geometry of specimens are the reasons for the necessity of the numerical examinations connected to the experiments performed by us, as the present regulations give the geometrical correction factors' value only for the rectangular specimens' three point bending case, so these factors have to be determined separately at

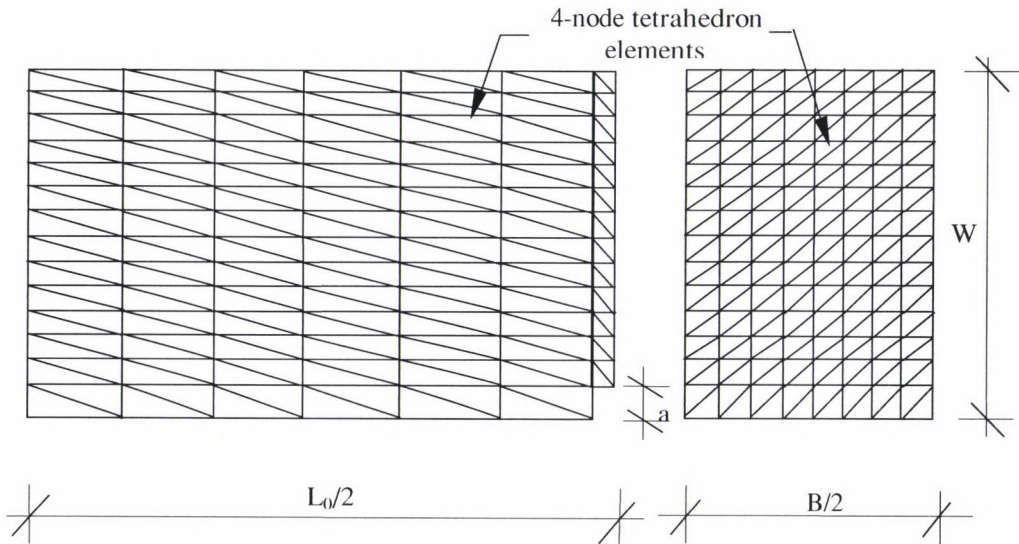


Fig. 3. 3D finite element mesh of rectangular specimens



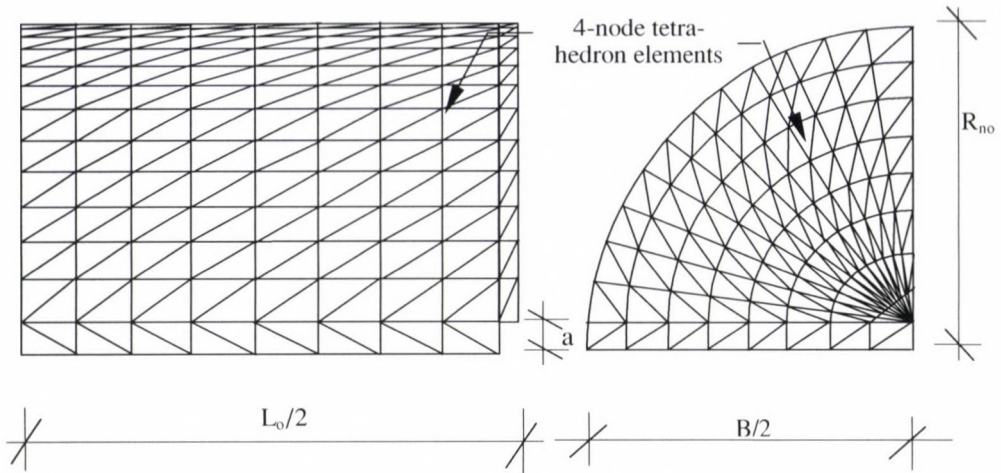


Fig. 4. 3D finite element mesh of half-cylindrical specimens

half-cylindrical specimens. The control of identification is expediently executed with the help of another numerical model for the sake of more unambiguous control. At the present examinations we used fracture mechanical applications provided by *three dimensional* (3D) finite element analysis.

First we examined what kind of stress intensity factors the *analytic formula* (in the case of standard rectangular specimens) defined by the present current competent norms gave for result. Table 3 contains these values showing the applied formula [9].

Next we examined the results of 3D finite element analysis (excellent summary can be found on the fracture mechanical applications of finite element technology in the

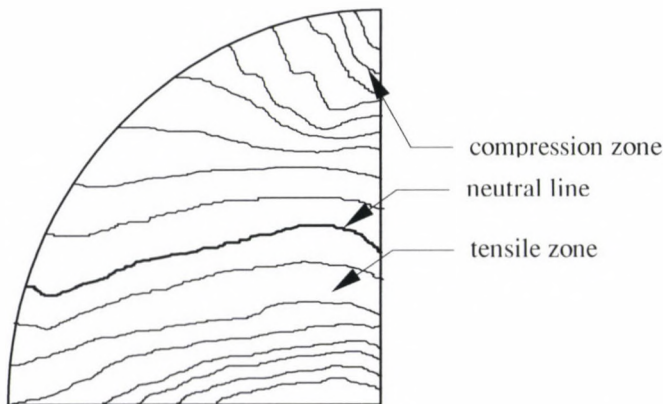


Fig. 5. The contour line picture of stress distribution in axis of rotation direction for half-cylindrical specimen

works of Füzzy, Owen, Fawkes etc.). Figure 3 shows the finite element mesh for rectangular and Fig. 4 for half-cylindrical specimens.

Beside the finite element mesh a typical stress distribution of the vertical segment under the load can be seen for half-cylindrical specimens in Fig. 5.

During fracture mechanical solutions the finite element examinations give estimation for the stress intensity factors with the application of displacement field (with the help of formulas got for displacements from Koloszov–Muszhelisvili’s complex functional solution) at the surroundings of crack tip (in the present case artificial notch). The connection of  $K_{Ic}$  values with displacements can be described with the following simple equation [3]:

$$K_{Ic} = \frac{E}{1-\nu^2} \cdot \sqrt{\frac{\pi}{2 \cdot r_i}} \cdot \frac{v_i}{2},$$

where  $v_i$  is the displacement component in the direction of notch,  $r_i$  is a polar coordinate whose origin is at the crack tip,  $E$  is the Young’s modulus,  $\nu$  is the Poisson’s ratio. Naturally this relation is valid only far from the crack tip, so at the tip, at the singularity place we got the  $K_{Ic}$  value from the extrapolation of the other  $K_{Ici}$  values. Table 4 and Fig. 6 show a sample for this extrapolation based on data of CI-1 specimen.

The stress intensity factors got from the results of 3D finite element solutions can be seen in Table 4 and Table 5.

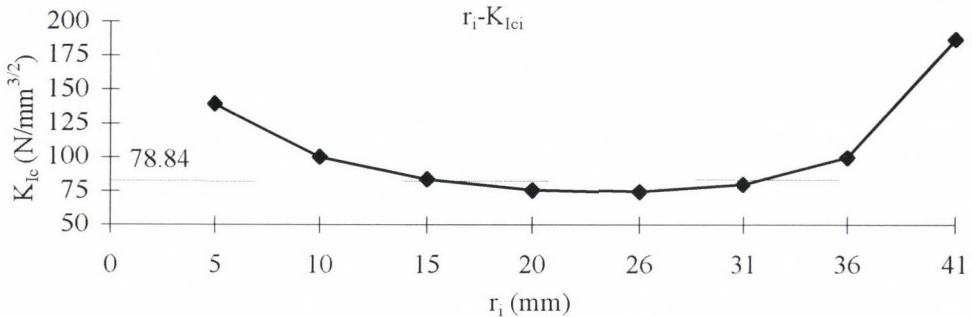


Fig. 6. A sample for the extrapolation of stress intensity factor (CI-1 specimen)

*Table 4*  
Numerical values of an  
extrapolation example (see Fig. 6)

i	$r_i$ [mm]	$K_{Ici}$ [N/mm <sup>3/2</sup> ]
221	5.11	139.19
208	10.22	100.58
195	15.33	84.31
182	20.44	76.44
169	25.55	74.49
156	30.66	80.11
143	35.77	99.76
130	40.88	186.36

$\Rightarrow K_{Ic} = 78.84$  [N/mm<sup>3/2</sup>]

*Table 5*  
The extrapolated value

Sample	Material	$K_{Ic}$ [N/mm <sup>3/2</sup> ]
PI-1	Dense limestone (I)	65.43
PI-2		80.09
CI-1		78.84
CI-2		101.25
CI-3		95.75
PI-3	Coarse limestone	7.53
PI-4		6.77
CI-4		8.18
CI-5		7.56
PI-5		Sandstone
PI-6	24.38	
CI-6	31.98	
CI-7	28.60	
CII-1	Dense limestone (II)	
CII-2		96.86
CII-3		61.84
CII-4		80.01
CII-5		84.10
CII-6		112.36
CII-7		84.57
CII-8		94.96
CII-9		91.58
CII-10		88.64
CII-11		88.85



#### 4. Proposed connection for the solution of stress intensity factor

From the laboratory experiments presented above, or rather from the critical loads given as the results of the experiments, the stress intensity factors could be determined with the help of analytical formulas and finite element calculations. Our next task was to determine a new analytic formula at the half-cylindrical specimen to give the same result for stress intensity factor as the rectangular specimen applied as a good method until nowadays, but being difficult to produce.

After the assessment and analysis of the results we suggest the following relation for the new formula:

$$K_{Ic} = \frac{P_c \cdot L_o}{R_{nom}^3} \cdot \sqrt{4 \cdot \pi \cdot a} \cdot Y(\alpha), \text{ where} \quad \alpha = \frac{a}{R_{nom}}$$

$$Y(\alpha) = 1.52 - 2.20 \cdot \alpha + 7.71 \cdot \alpha^2 - 13.55 \cdot \alpha^3 + 14.25 \cdot \alpha^4$$

$R_{nom}$  is the height of the half-cylindrical specimen,  $L_o$  is the distance between the supports,  $a$  is the crack length (see Fig. 2) and  $P_c$  is the critical load. The stress intensity factors solved from this new formula were compared with the results of the other methods (see Table 6).

Table 6  
The comparison of the numerical and analytical values

Sample	New $K_{Ic}$ (N/mm <sup>3/2</sup> )	$K_{Ic}$ from num. solution (N/mm <sup>3/2</sup> )	Difference %
CI-1	73.73	78.84	6.48
CI-2	85.37	101.25	15.68
CI-3	103.12	95.75	7.15
CI-4	6.34	8.18	22.49
CI-5	6.09	7.56	19.44
CI-6	25.02	31.98	21.76
CI-7	24.61	28.60	13.95
CII-1	52.18	52.47	0.55
CII-2	90.84	96.86	6.22
CII-3	64.34	61.84	3.89
CII-4	79.30	80.01	0.89
CII-5	90.07	84.10	6.63
CII-6	109.91	112.36	2.18
CII-7	101.49	84.57	16.67
CII-8	95.33	94.96	0.39
CII-9	90.64	91.58	1.03
CII-10	91.96	88.64	3.61
CII-11	88.39	88.85	0.52
		Mean difference:	8.31%

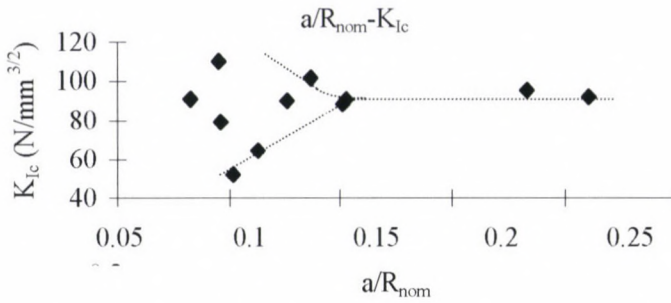


Fig. 7. The effect of the crack length to the stress intensity factor

The low error-percentage difference is shown in Table 6 refers to the serviceability of the formula.

In the last step, besides checking the correction of the new formula, we examined the effect of crack length to the connection suggested by us at half-cylindrical dense limestone specimens (signed CII-i,  $i = 1-11$ ). The results of this examination can be seen in Fig. 7. The value of the stress intensity factor stabilizes with the increase of the relative crack length at 0.15. We did not check this statement above 0.27.

## 5. Conclusions

This paper suggests a new analytic formula for the solution of stress intensity factor at a new and more easy to produce formability specimens of rocks. First critical loads were measured from three point bending tests of sandstone, dense and coarse limestone samples, then stress intensity parameters were determined with the help of presently valid analytic formulas at rectangular specimens, and the results of all examinations were checked with finite element method. The results of the new formula constructed with the help of data assessment and analysis proved its serviceability.

## Acknowledgement

The authors thank OTKA curratorium for the financial assistance of the present research No. T-014270.

## References

1. Anderson, T. L.: Fracture Mechanics. CRC Press, London, 1995
2. Czoboly, E.- Gálos, M.- Havas, L.- Thamm, F.: Appropriate fracture mechanics specimens for testing. Fracture control of Engineering Structure ECF 6 Vol. III. Amsterdam-London EMAS pp. 2.105-2.115, 1986

3. Yaoming, Mi: Three-Dimensional Analysis of Crack Growth. Topics in Engineering Vol. 28. Computational Mechanics Publications Southampton UK and Boston USA, 1996
4. Hahn, H.G.: Bruchmechanik. Treubner Studienbücher. Stuttgart, 1976
5. Kolozov, G.V.: Über die Anwendungen der komplexen Funktionentheorie auf das ebene Problem der mathematischen Elastizitätstheorie. Dissertation at the University of Juriev, 1909
6. Muszhelisvili, N.I.: Some basic problems of mathematical theory of elasticity. Kluwer P. Groningen, 1953
7. Ouchterlony, F.: On the background to formulae and accuracy of rock fracture toughness measurements using ISRM standard core specimens. Int. J. Rock Mech. Min. Sci. & Geomech. Abstr. 26 (1989) 13–23
8. Owen, J. – Fawkes, A.J.: Engineering Fracture Mechanics. Pineridge Press Ltd. Swansea, 1983
9. Stress Intensity Factors Handbook 1987. Committee on Fracture Mechanics. The Society of Materials Science, Japan. Pergamon Press
10. Broek, D.: The Practical Use of Fracture Mechanics. Kurver Academic Publishers, 1988
11. Fűzy, J.: Magasabb szabadságfokú kontinuumok elmélete és építőmérnöki alkalmazásai. Akadémiai Kiadó, 1992



# DEFORMATION AND STRESS ANALYSIS OF HYPERBOLIC PARABOLOID SHELLS TAKING INTO ACCOUNT REALISTIC BOUNDARY CONDITIONS

Bódi, I., Fűzy, J. and Klopka, Z.

*Technical University of Budapest, Department of Reinforced Concrete Structures,  
Bertalan Lajos u. 2, H-1521 Budapest, Hungary*

(Received: 10 April 1998)

When determining the state of stress and deformation of a shell, it is always a basic assumption that the boundaries behave exactly according to the assumed theoretical ones ("rigid" or "semi-rigid" edge beams, tie rods without elongation, etc.). It can be easily seen, however, that this is not strictly correct, not least because of the finite dimensions of the cross-section of the edge beam or tie rod. As a consequence, the deformations of the boundary points will have values other than theoretically assumed. It is this deformational incompatibility that we examine in this work.

The problem was approached numerically, where a large number of problems were analysed using a finite element program. Some design parameters were varied in order to analyse their respective influence on the state of stress and deformation. The dimensions of the edge beam, the connection of adjacent edge beams, the cross-section of the tie rod and the direction of the load were the parameters which were varied. The model whose results matched closest the theoretical solution was chosen as the basic model. The results obtained by changing the design parameters were compared to the basic model and the differences were plotted graphically. Based upon the analysis we have concluded that the deviation from the membrane stress state is caused rather by the displacement of the supports or elongation of the tie rod than by the finite dimensions and the nature of the connection of the edge beams.

## 1. Introduction

The general goal of research was to examine the deformational incompatibility that exists between edge beams and the shells in the membrane stress state. As the method of research we have chosen the finite element examination of a large number of suitably defined particular cases and we have derived theoretical conclusions from them. As this method cannot yield general results, we have chosen several specific problems; in this paper we will present the conclusions obtained for the case of hyperbolic paraboloid shells (Fig. 1).

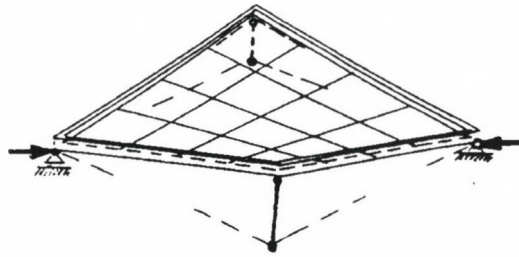


Fig. 1

## 2. Review of the results of analysis

The data were varied in such way, that the effects of changes in the geometry of edge beams and in the support conditions (the tie rod is treated as part of the edge beam system, too) can be traced. Based upon this the following statements could be made.

The most important statement is that the membrane solution can be reached approximately only. With respect to the importance of the above statement, we will describe in detail those assumptions upon which were used in the model.

As the sign of the membrane deformation is opposite to that of the supporting structure (e.g. tie rod), active, external forces having the appropriate signs were employed at the four corners of the structure acting in the direction of the diagonal, replacing the reactions of the horizontal support (tie rod). The tensile stiffness of the edge beam was chosen as infinitely large (infinitely large in the numerical analysis means a very large number) which prevents the compressive deformations of the edge beam on the one hand, and ensure that the longitudinal internal forces of the adjacent shell parts approach zero on the other. The vertical bending stiffness of the edge beam is chosen to be infinitely large, which is to simulate the continuous support; the horizontal bending stiffness (in the plane of the shell) is also set to infinity to preserve the straightness of the

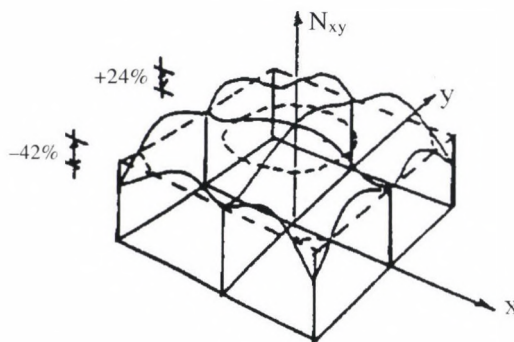


Fig. 2

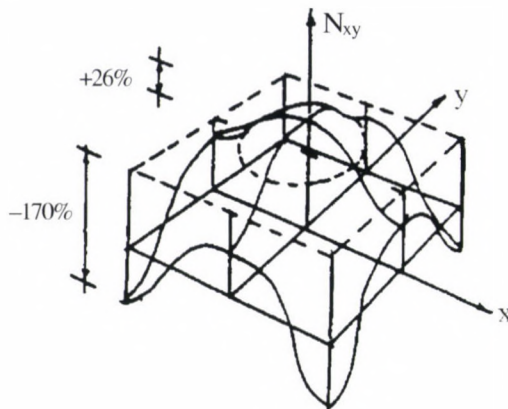


Fig. 3

edge beam. The connections between the edge beams are hinged, enabling the shear deformations in the corners of the shell. In order to minimise the deviation of results due to the approximate assumptions of shallow shells, the load is set to be acting perpendicular to the surface instead of in the vertical direction. The shear stresses  $N_{xy}$  calculated for the above system show a deviation from the nominal stresses (according to the membrane solution) from +24% to -42%, respectively (Fig. 2).

The internal forces  $N_x$  and  $N_y$ , which act along the straight generatrices, can not counterbalance any external load, so the deviation of the calculated stresses  $N_{xy}$  from the nominal ones yields that portion of the load which has to be carried by bending moments (plate action).

If a diagonal tie rod is applied instead of the reaction forces acting as external loads, the nature of the  $N_{xy}$  stress distribution does not change essentially, only the magnitude of the deviations from the membrane stresses increase (Fig. 3). The deviations of the stresses  $N_{xy}$  became smooth again, if the edge beam corner connections are rigidly connected (instead of the hinged connection, see Fig. 4).

In this latter case the rigidly connected edge beams, having a significant bending stiffness, form a closed frame, which eliminates the necessity of the tie rod. In this case a negligible horizontal force arises in the tie rod of the model indeed. The nature of the distribution of shear stresses differ significantly from the previous ones in these cases, and the portion of the load producing plate moments increases (see the dashed line in the figure).

Figure 4 depicts the case where the infinitely rigid bending stiffness of the edge beams is reduced to zero (but the tensile stiffness remains finite). Naturally, in this case the tie rod is necessary but even now the axial force arising in the tie rod approaches 82% to 90% of the theoretical value from depending on whether the connection between the adjacent edge beams are hinged or rigid. This phenomenon refers to that even in this case a frame-like action develops in the edge beams due to the co-operation between the shell and the edge beam.



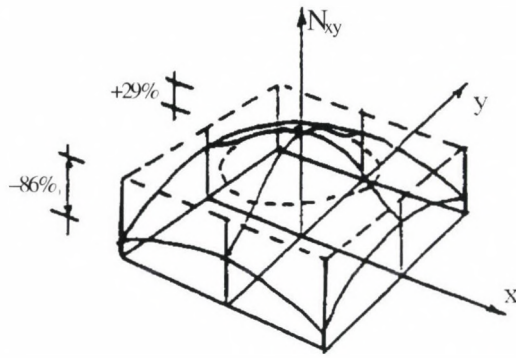


Fig. 4

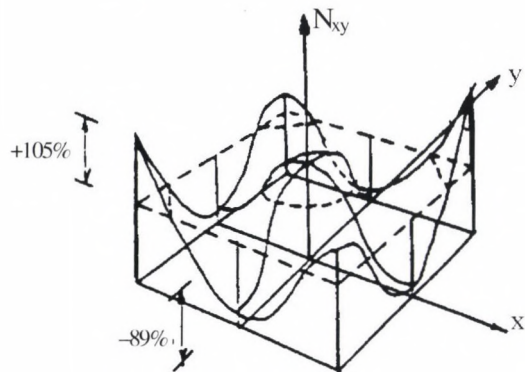
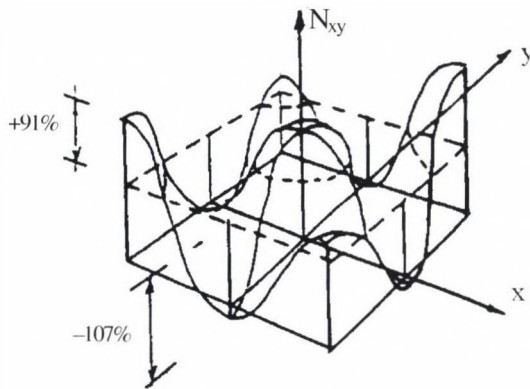


Fig. 5

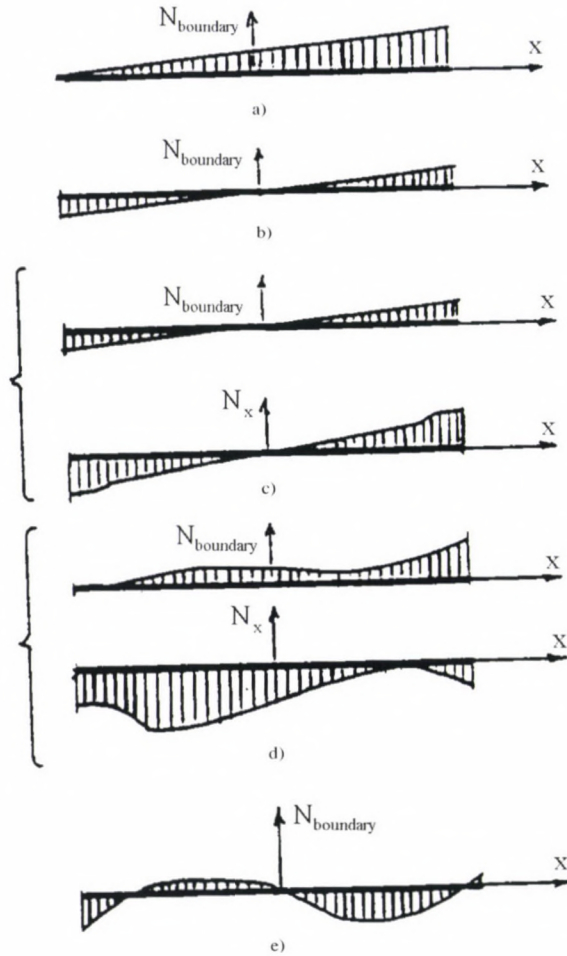


Fig. 6

Furthermore, the distribution of the axial force (normal force -  $N_{\text{boundary}}$ ) arising in the edge beam and the parallel, normal force of the adjoining shell ( $N_x$ ) will be analyzed. The results are summarized in Fig. 6.

In Fig. 6/a the distribution of the normal force is shown in the case when two horizontal reaction forces are acting in the diagonal direction. Figure 6/b depicts the same case with the difference that there are four reaction forces at the corners. The case when the horizontal bending stiffness of the edge beams is quite large, the tensile stiffness is finite and the connection between the edge beams is rigid, is shown in Fig. 6/c. Obviously, this coincides with the case when four reaction forces are applied, a closed frame action develops and the axial force in the tie rod is practically nil. The adjoining

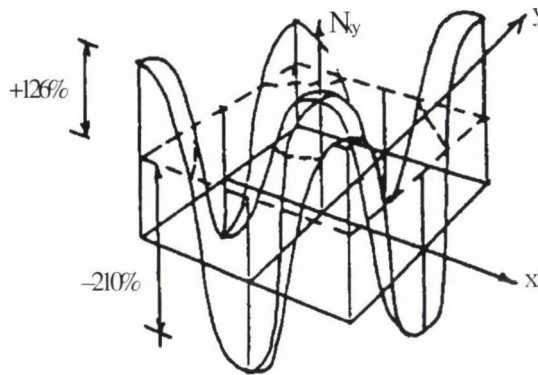


Fig. 7

parallel, normal (membrane) shell force is also shown, which follows the compression of the edge beam having a finite tensile stiffness.

Figure 6/d shows the case, when the horizontal bending stiffness of the edge beam is set to zero. Due to this, the frame effect has to be taken over by the shell. Distortion of this nature can be seen both in distribution of the axial normal force of the edge beam and in the distribution of the adjoining parallel normal shell force. The axial force of the tie rod equals 85% of that of the membrane solution, what means that 15% is taken over by the in-plane frame-like action of the shell.

Finally, Fig. 6/e depicts the case when the dimensions of the edge beam are within the limits of the realisable technical solutions. It can be seen that the distribution of the axial force deviates from the membrane solution to such an extent that it even changes sign several times.

The distribution of the shearing stresses for this case can be seen in Fig. 7.

### 3. Summary

The numerical analysis has shown that the membrane solution in the case of hyperbolic paraboloid shells cannot be achieved even with idealised assumptions. Furthermore, within the limits of realisable technical solutions the results obtained are significantly different from the ones yielded by the membrane solution.



# SEMIEMPIRICAL THEORY OF MASONRY STRENGTH

Ding Dajun

*Nanjing Institute of Technology, NIT, China*

(Received: 5 September 1997)

Basing on a large number of experimental studies conducted in China and referring to Russian materials, author tries to summarize up preliminarily a more complete theory of masonry strength (1st draft) for discussion. He hopes to have the kind help from the vast members of colleagues at home and abroad so as to revise continuously for striving to achieve more complete and more perfect one. Because the studies on brick masonry were the most, hence in this article it will be taken as a dominant to be illustrated.

## 1. Compressive strength of unreinforced masonry

### 1.1. Failure mechanism of masonry under compression

Because of the nonuniformity of mortar, the units in masonry will bear flexural and shearing stresses beside nonuniform compressive stress. Figure 1 shows the sketch of the complicated stress and deformation (exaggerated into 200-folds) states of bricks in masonry given in Russian material [1].

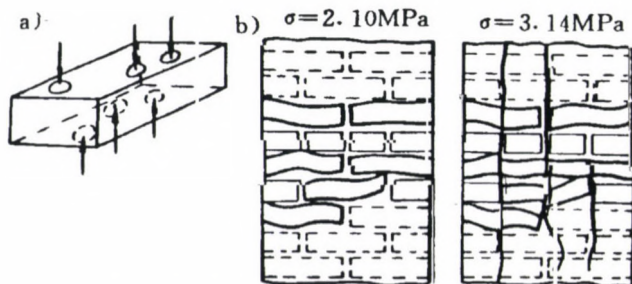


Fig. 1. In masonry, bricks bear nonuniform compressive (a), flexural and shearing stresses, and their deformations (b)

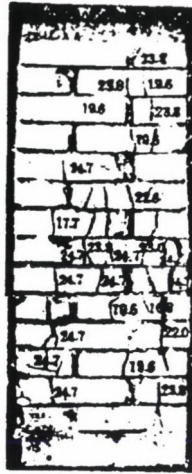


Fig. 2. Failure of brick masonry under compression (experimental photo of NIT)

Owing to bearing nonuniform compressive stress and flexural as well as shearing stresses (the flexural and shearing stresses are caused due to nonuniform compressive stress), masonry strength will be much lower than unit strength. Brick masonry fails under compression is due to the instability of small columns with thickness of  $\frac{1}{2}$  brick, separated by continuous vertical cracks through several courses of brick under flexural and shearing stresses (Fig. 2). [2], i.e. the brick strength has not been fully utilized. The strength of stone masonry with small stone packing pieces in mortar joints is lower than that without these packing pieces [3]. But tests showed the masonry strength still depends on those of its matrices (unit and mortar), because the higher the unit strength is, the higher the strength of units against flexure and shear is also; as the mortar strength is higher, the flexural and shearing stresses caused in unit will be smaller.

Based on a large number of test results obtained in China, to summarize up these has obtained that the average compressive strength of masonry is a combination of those of units and mortar and will be adjusted with the changes of mortar strength, the general formula [4]:

$$f_m = k_1 f_1^\alpha (1 + 0.07 f_2) k_2 \quad (1)$$

where

$f_1, f_2$  – strengths of unit and mortar, respectively, in MPa;

$k_1$  – factor, varying with the kinds of unit in masonry and laying methods, see Table 1, where the kinds of unit mean whether there are holes in units or not, for row-lock cavity wall, the void ratio attains 80%, so  $k_1$  is very low and equal to 0.13; for rubble masonry, because the lap of rubble in masonry is not good, so  $k_1$  is also low and equal to 0.22; as for the masonry of hollow brick with void ratio of

- 25% or so, but it is considered as that of a solid brick due to neglecting the favourable influence of its larger thickness on masonry strength;
- $k_2$  – further adjustment factor of masonry strength due to the change of mortar strength (see Table 1 and below);
- $\alpha$  – utilization factor, being related to the thickness of units, in rowlock cavity wall, the thickness of a course in masonry is increased as it is built on edge, so  $\alpha = 1$ .

Table 1  
Average values of axially compressive strength (MPa)

Sequence	Kinds of masonry	$f_m = k_1 f_1^\alpha (1 + 0.07 f_2) k_2$		
		$k_1$	$\alpha$	$k_2$
1	clay brick, hollow brick, nonclay silicate brick	0.78	0.5	as $f_2 \leq 1$ , $k_2 = 0.6 + 0.4 f_2$
2	1-brick thick row-lock cavity wall	0.13	1.0	as $f_2 = 0$ , $k_2 = 0.8$
3	small hollow concrete block	0.46	0.9	as $f_2 = 0$ , $k_2 = 0.8$
4	medium block	0.47	1.0	as $f_2 > 5$ , $k_2 = 1.15 - 0.03 f_2$
5	crude-worked stone	0.79	0.5	as $f_2 < 1$ , $k_2 = 0.6 + 0.4 f_2$
6	rubble	0.22	0.5	as $f_2 < 2.5$ , $k_2 = 0.4 + 0.24 f_2$

Remark: 1. all  $k_2$  excluded the conditions in Table 1 are equal to 1.

2.  $f_1$  – average value of compressive strength of units (brick, stone, block),  $f_2$  – average value of mortar, both in MPa.

From Table 1, it can be seen that for the masonry built with thicker units,  $\alpha$  will be greater as (0.9–1.0), for the others,  $\alpha = 0.5$ .

$(1 - 0.07 f_2) k_2$  in Table 1 shows the influence of mortar strength on masonry strength, i.e. masonry strength increases linearly with the increase of mortar strength, but as  $f_2 \leq a$  definite value, masonry strength will decrease further, it is due to that as units are stressed integrally with mortar in masonry, the lateral deformations of units and mortar should be consistent; the lateral deformation of low-strength mortar is so larger as to cause a lateral deformation in units larger than that they are stressed in compression separately, it results in transversely tensile stress in units, so leads to decrease masonry strength further [5], i.e. in this case,  $k_2 \leq 1$ ; but the mortar is constrained by units in masonry to cause transversely compressive stress resulting in a 3-dimensional compressive state so its compressive strength is increased very greatly, thus, the masonry strength with low-strength mortar is greater than mortar strength, so the masonry with  $mM = 0$  (in newly built masonry or the mortar strength in freeze-thaw stage as making check calculation for the masonry built with winter construction) has a definite values of strength and elasticity modulus (capacity against deformation).

For medium block masonry, as  $f_2 > 5$  MPa,  $k_2 > 1$ , i.e. masonry strength should be increased, it is suggested from analyzing test results. Writing the formula of  $f_m$  with  $k_2$  and finding  $\partial f_m / \partial f_2 = 0$  gives  $(f_2) = 12.024$  MPa to obtain the maximum calculated  $f_m =$



1.453  $k_1 f_1$  taking ( $f_2$ ) = 12 MPa; as  $f_2 > 12$  MPa, the calculated  $f_m$  will be contrarily decreased. Hence, author proposed to take  $f_m = 1.45 k_1 f_1$  [6] as  $f_2 > 12$  MPa.

For the masonries of all kind, as  $f_2$  are not over than the about boundary values, the influence of mortar is the same as  $(1 + 0.07 f_2)$ , if  $f_2 = 2.5-10$  MPa, the masonry strength will be increased equally to 1.45-folds. But following Russian theory, the influences of mortar strength are, respectively, in the order: rubble masonry, brick masonry, concrete block masonry and large block masonry (with thickness of a course being equal to and thicker than 600 mm), if  $f_1 = 10$  MPa (in Chinese Code, the influence of mortar strength is independent with relation to unit strength, but in Russian theory, the influence is dependent on unit strength), as mortar strength is in the above field, the increases of masonry strength are, respectively, 1.55, 1.42, 1.25 and 1.0, i.e. mortar strength has no influence on the strength of large block masonry.

For two batches of bricks with the same compressive strength, but one with higher flexural strength, of which the masonry has also high compressive strength than that of masonry of bricks with lower flexural strength. Hence for standard bricks, beside compressive strength, a corresponding requirement of flexural strength should be prescribed, i.e. the flexural strength is also an important mechanical index of standard bricks. For thicker hollow bricks and blocks, it is unnecessary to suggest the requirement of flexural strength.

## *1.2. Other factors influencing compressive strength of masonry*

The factors influencing masonry strength are very many, some mains are given in the following.

### *1.2.1. Regularity of units, density and uniformity of mortar joints*

These factors are mutually relative. Because the regularity of units influences the uniformity of built mortar joints. Both the density and uniformity of mortar joint will influence the uniform degree of compressive stress in units built in masonry, i.e. will relate to the flexural and shearing stresses caused in units being larger or smaller, so finally will decide masonry strength.

For an example, if taking the strength of crude-worked stone masonry as 1.0, the strengths of fine-finished stone masonry, semifinished stone masonry, rough-finished stone masonry and the masonry of stone with closed cracks around edges will be, respectively, 1.5, 1.3, 1.2 and 0.8 [7].

The density and uniformity of mortar joints will be improved under long-term compression (mainly in earlier stage), so the masonry strength will be increased [8]. Beside the quality of masonry built by high-level brick-layers being higher, one of the reasons of increasing strength is also because the newly built masonry is compressed in earlier stage [9] due to quickened masonry pace.

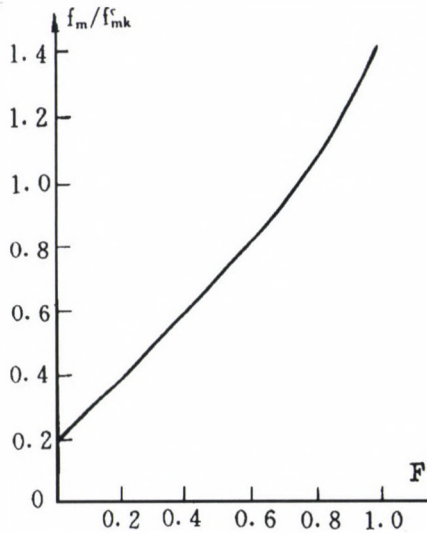


Fig. 3.  $f_m / f_{m,k}^c - F$  curve

### 1.2.2. Fullness of mortar joints

The fullness of mortar joints directly influences the uniformity of compression in units and the quantities of flexural and shearing stresses caused in them, so finally influences masonry strength.

Based on test results, the Buildings Research Institute of Sichuan Province gave [10]:

$$f_m = (0.2 + 0.8F + 0.4F^2)f_{m,k}^c, \quad (2)$$

where

$F$  – fullness of mortar in horizontal joints, in decimal;

$f_{m,k}^c$  – characteristic value of masonry strength in Design Code GBJ3-88, determined following unit and mortar strengths.

From Eq. (2), it can be seen that as  $F = 0.73$ ,  $f_m$  will be equal to the value given in Code. Chinese Technical Code of Work and Acceptance for Masonry Engineering [11] prescribes the fullness of mortar joints should be not smaller than 0.8, it allows for unforeseen circumstances.

The change curve of  $f_m / f_{m,k}^c - F$  is shown in Fig. 3.

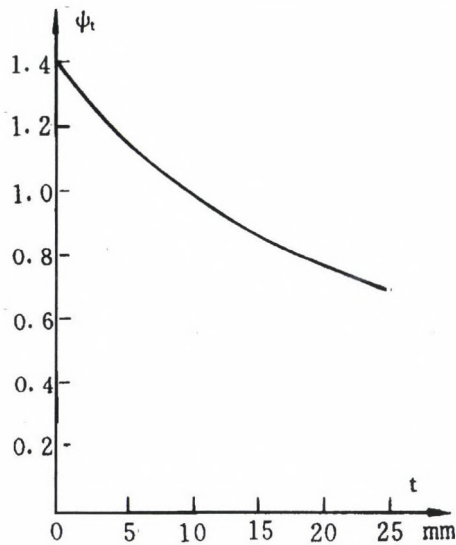


Fig. 4.  $\psi_t$ - $t$  curve

### 1.2.3. Thickness of mortar joints

The thicker the joint thickness is, the greater the flexural and shearing stresses caused by the effect of beams on elastic foundation will become, then the masonry strength will decrease more greatly. Basing on the statistics of test results, Hunan University proposed preliminary influence factor as follows [8]:

$$\psi_t = \frac{1.4}{1 + 0.04t} \quad (3)$$

where  $t$  – joint thickness, in mm.

The curve of  $\psi_t$ - $t$  is shown in Fig. 4.

It can be seen that the thinner the thickness  $t$  is, the higher the masonry strength will be. The strength of masonry of bricks with two built surfaces to be polished is much greater than the ordinary brick masonry with mortar joints thickness of 10 mm (attaining 60–90%) [8].

But it should not overestimate the favourable influence of very thin mortar joint, because this influence depends on the regularity of built surfaces of units. If the built surfaces of units are not regular, as the mortar joints are too thin, the nonuniformity of them will be increased, the very small stones in sand mixed in mortar will play a role of stiff points in joints to cause additional flexural and shearing stresses resulting in a compressive strength decrease of masonry. The standard joint thicknesses are 8–12 mm [8].



#### 1.2.4. Water content ratio of bricks

The water content ratio will also influence the compressive strength of masonry. As the water content ratio is low, the water retentivity of mortar is influenced, it will influence built quality of masonry and lead to a decrease of the compressive strength of masonry.

In ref. [8], based on tests an influence factor of water content ratio on the compressive strength of masonry was proposed as follows:

$$\psi_w = 0.8 + \frac{\sqrt[3]{\xi_w}}{10}, \quad (4)$$

where  $\xi_w$  – water content ratio of bricks as being built, in 1%.

Figure 5 shows  $\psi_w - \xi_w$  curve, as  $\xi_w$  is larger, the compressive strength of masonry will be higher. It is because the wetter the bricks are, even there is a flowing phenomenon of mortar on brick surfaces, but owing to so good mobility of mortar as to be able to pave thin uniform joints, then improve the complicated stress state in masonry, then the more compressive strength of masonry increases [8].

#### 1.2.5. Pattern of lapped joints

The pattern of lapped joints of units is also to influence the compressive strength of masonry. As the height of unlapped joints is not over 3 times of half-brick length, the influence is not great, otherwise the compressive strength will have a greater decrease [12].

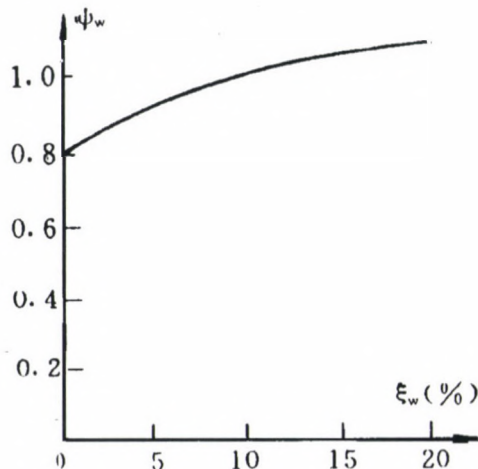


Fig. 5.  $\psi_w - \xi_w$  curve

For a square brick column with side of 2-brick length, it is not allowable to build it into a 1-brick core and a ½-brick out sleeve without any lapping along entire column height, which is called “ghost pushing millstones” (“Gui-tui-mo”).

#### 1.2.6. Pure cement mortar

Both the workability and water-retentivity of pure cement mortar are not good as those of compo mortar (cement-lime mortar), using this mortar to build masonry is difficult to pave, so it cannot build more uniform and dense joints, then the compressive strength of masonry will be decreased, generally by 15% [8].

#### 1.2.7. Skillfulness of brick-layers

The skillfulness of brick-layers will directly influence masonry quality, including all factors mentioned in the above, such as the fullness of mortar joints, the regularity and leveling, density and uniformity of mortar, as well as the thickness of joint, etc. The masonry members subjected to compression built by unskilled brick-layers may be with a larger initial bending, i.e. with a larger eccentricity, may also be with inconsistent quality in two sides resulting in a greater deviation of physical center axis of gravity, etc. will all decrease masonry strength.

#### 1.2.8. Masonry in engineering

The brick masonry specimens taken from existent building showed under equal built conditions, the compressive strength of masonry in engineering is higher than that built in laboratory conditions. It is because with the progression of building work, on the newly built masonry, the compression from dead load is continuously applied so as to improve the uniformity and density of mortar, then to increase the masonry strength. The tests in laboratory under compression at early stage indicated also this [9]. Generally the increase is about 15%.

### 1.3. *Compressive elasticity modulus of masonry*

A large number of tests conducted in China showed [13], the elasticity moduli of brick and block masonries are in relation to their compressive strength and reduce with the decrease of mortar strength. However, for stone masonry, because the elasticity modulus of stone is much higher than that of mortar, the deformations in stone masonry are mainly caused by mortar joints, so the elasticity modulus of stone masonry is only in relation to mortar strength and reduces with the decrease of mortar strength, the elasticity moduli of fine-finished and semifinished stone masonry are in 3-folds of those of rough-finished stone, crude-worked stone and rubble masonries.

## 2. Strength of local compression

The increase of local compression strength is both due to the confined effect and the spread of compressive stress [14–16]. The failure of masonry under local compression does not occur directly at the place of local compression, because the strength of this place has been greatly increased due to the strong constraint of cushion plate (in practice, the constraint of the above column or beam), but it begins at the crack(s) occurring generally in the second or third course of masonry. The uniform compressive stress  $\sigma_o$  on beams discharges due to the effect of “discharging arch” [16]. As  $\sigma_o/f_m \leq 0.4$ ,  $\sigma_o$  can be not considered, because this arch discharges the  $\sigma_o$  applying on beams to both sides of their supports and to increase the confined effect to the masonry on both sides.

## 3. Tensile and shearing strength of masonry

As the masonry built up with units and mortar is stressed under axial tension, the force may be normal to horizontal joints (Fig. 6a) or parallel to them (Fig. 6b, c). In the former case, failure occurs in horizontal joint, the tension will be taken by the normal adhesion strength of mortar and bricks, but the deviation of this strength is too large so as to be unreliable, so it is not allowable to design members bearing axial tension in engineering. As the strength of bricks is not too low, following Fig. 6b, failure will occur along toothed joint and the tension will be borne by the tangential adhesion strength of joints, so it is in relation to mortar strength.

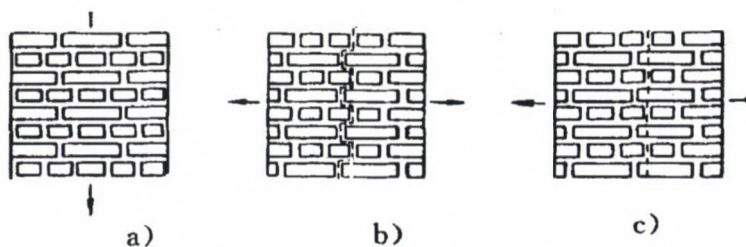


Fig. 6. Masonry under axial tension normal to horizontal joints

Based on a large number of the results of tests (amount of specimens attained 1378) conducted in China, a series of statistical formulas for determining average tensile and shearing strengths are given as follows [4]:

– average axial tensile strength of masonry:

$$f_{t,m} = k_3 \sqrt{f_2}, \quad (5)$$



– average flexural tensile strength of masonry:

$$f_{tm,m} = k_4 \sqrt{f_2}, \tag{6}$$

– average shearing strength of masonry:

$$f_{v,m} = k_5 \sqrt{f_2}, \tag{7}$$

where  $k_3$ – $k_5$  are listed in Table 2.

From Table 2, it can be seen that: (i) the average axial tensile strength of brick masonry is higher than its shearing strength; (ii) for the other masonries, both of these strengths are the same, but for rubble masonry, the shearing failure cannot occur along a horizontal straight joint, so its average shearing strength is much higher than its average axial tensile strength.

Table 2  
Average values of axial tensile strength  $f_{t,m}$ , flexural tensile strength  $f_{tm,m}$  and shearing strength

Se- quence	Kinds of masonry	$f_{m,m} = k_3 \sqrt{f_2}$	$f_{tm,m} = k_4 \sqrt{f_2}$		$f_{v,m} = k_5 \sqrt{f_2}$
		$k_3$	$k_4$		$k_5$
			along toothed joint	along straight joint	
1	clay brick, hollow brick	0.141	0.250	0.125	0.125
2	small hollow concrete block	0.069	0.081	0.056	0.069
3	medium hollow concrete block	0.053	0.063	0.044	0.053
4	medium solid ash-fly block	0.034	0.041	0.028	0.034
5	rubble	0.075	0.113	–	0.188

Besides, before a shearing failure, the shearing deformation has been large, so friction has taken a part of shear, so in the calculation of shearing along a horizontal joint, the influence of friction can be considered.

The influences of cleanliness of unit surfaces, wetness of units and consistency of mortar on tangential adhesion strength are greater. The optimum water content ratio for adhesion strength is equal to 8% or so. In dry and saturated masonry, the adhesion strengths will decrease, respectively, by 20% and 60% [8].



Fig. 7. Stressed in horizontal direction of reinforced masonry with fabrics

As the strength of bricks is lower, the failure under tension may occur along bricks and vertical joints as shown in Fig. 6c, in this case, the average axial tensile strength [18] depends on brick strength without consideration of vertical joints [2].<sup>1</sup>

The average tensile strength along units:

$$f_{tm} = 0.212\sqrt[3]{f_1}. \quad (8)$$

Flexural tensile strength is taken as  $1.5 f_{tm}$ :

$$f_{tm,m} = 0.318\sqrt[3]{f_1}. \quad (9)$$

#### 4. Reinforced masonry with steel fabrics

The strength increase of reinforced masonry with steel fabrics is not due to creating 3-dimensional compression state, but is because an instability of the small  $\frac{1}{2}$ -brick columns separated by vertical cracks but tied by the steel fabrics are prevented from occurring in earlier time. Before cracking, it is certain that there occurs a 3-dimensional compression state, i.e. in this time for equilibrium, steel stress  $\sigma_{s,cr}$  can cause horizontal compressive stress  $\sigma_x$  on masonry sections (Fig. 7a), hence it increases the cracking capacity. But after the vertical cracks occur, on cracked surfaces, it is impossible to produce  $\sigma_x$ , so there does not occur again a 3-dimensional compression state [14–16, 20].

The tensile stress  $\sigma_s$  of fabric in small corner column is anchored in joints by bond force (there is compressive stress applying on joints to increase this force much more);  $\sigma_s$  on both sides of small middle column will be balanced by itself [14–16, 20].

In reinforced masonry with fabric, the vertical cracks occur often in partial height of masonry (Fig. 8), it is favourable to strength increase.

Under eccentric compression and as slender ratio of member is greater, the effect of fabric reinforcements decreases, so it is necessary to limit both the eccentricity and slender ratio, i.e. to prescribe the eccentricity on safe side does not beyond the core of section (in rectangular section  $e \leq h/6$ ), the slender ratio not over 16.

<sup>1</sup>In vertical joints, the mortar is not fully filled, and owing to shrinkage influence, they are not considered in calculation, but as taking care of filling fully, tests showed they can bear tension or shear [19].



Fig. 8. Failure of reinforced masonry with fabric (test photo of NIT)

### 5. Masonry with longitudinal reinforcements

The stressed mechanism of longitudinally reinforced masonry (composite masonry) is similar to reinforced concrete member, tests indicated also this [21]. As the surface layers are made of mortar, considering when these layers attain failure moment, the yielding strain of compressive bars has not reached, so a utilization factor  $\eta_s = 0.9$  of steel strength should be introduced.

The stability factor  $\phi$  of axially compressive masonry with longitudinal reinforcements is determined with slender ratio  $\beta$  and reinforcement ratio  $\rho$ . It is evident that the greater  $\rho$  is, the greater the  $\phi$  will be. The tests of reinforced concrete columns showed also this [22], but for simplification, in Chinese Design Code of Concrete Structures (GBJ10-89) the influence of  $\rho$  on  $\phi$  is not considered.

For eccentrical compression members the longitudinal flexure is considered by means of additional eccentricity  $e_i$  [21].

$$e_i/h = \beta^2 (1 - 0.022\beta)/2200. \quad (10)$$

Equation (10) can be derived as follows [21]: according to the plane section hypothesis at ultimate state (corresponding to  $N \rightarrow \max.$ ), taking  $\varepsilon_y = f_y/E_s = 0.0016$  (approximately corresponding to bars of steel grade II),  $\varepsilon_{m,u} = 0.003$  and distribution factor of curvatures along member height  $H_0$  to be 11, then [21]

$$e_i/h = \frac{0.003 + 0.0016 - \beta_x 10^{-4}}{0.95 h} \frac{H_0}{11} \frac{1}{h} = \frac{\beta^2}{2272} (1 - 0.0217\beta) = \frac{\beta^2}{2200} (1 - 0.022\beta),$$



where  $-\beta \times 10^{-4}$  is equivalent to the modification factor  $\zeta_2$  to consider slender ratio influence in Chinese Design Code of Reinforced Concrete Structures (GBJ10-89).

$e_i/h - \beta$  curve is shown in Fig. 9.

From the above, it can be seen that there are two points worthy to be discussed. (i) taking  $\varepsilon_{m,u}$  and  $\varepsilon_y$  in calculation is assuming to be in the case of balanced design, generally,  $\rho$  is not great, the steel bars have undergone a definite flow deformation before attaining ultimate state, to take  $\varepsilon_y$  for calculating  $e_i$  is on small side, this problem exists also in the calculation of  $\eta$  (magnification factor of eccentricity of axial compression for considering the longitudinal flexure in eccentric reinforced concrete compression members); (ii) the calculation of longitudinal flexure of eccentric members without considering  $\rho$  may not conform to the actual [22, 23] and is also inconsistent with the determination of the stability factor  $\phi$  of axial compression members. But in concrete columns, a factor  $\zeta_1 = \frac{0.5f_c A}{N}$  is introduced to modify the magnification factor  $\eta$  of

eccentricity due to longitudinal flexure, as  $r$  is higher, the axial compression will be greater,  $\zeta_1$  then  $\eta$  will be smaller, i.e. using  $\zeta_1$  to consider the influence of  $\rho$  on additional eccentricity.

At the end of this article, author should say that owing to the limitation of his levels on theory and practice, this paper is still superficial even revised many times, it can be only to consider as the first draft. Author does heartily hope the colleagues in this field

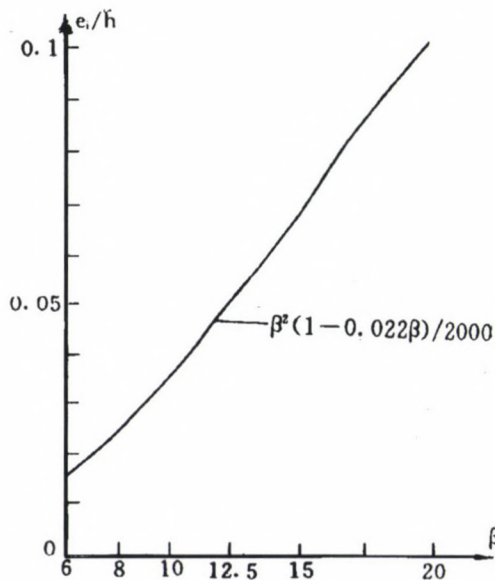


Fig. 9.  $e_i/h - \beta$  curve

at home and abroad would provide kind help to mail the author relevant materials and revising opinions, then author will make revision further, so as that it can become more complete and perfect.

## References

1. Semenzov, S. A.: Calculation of Masonries and Reinforced Masonries following Limited States (in Russian). Moscow, 1955
2. Ding Dajun (Editor-in-Chief): A Concise Book on Masonry Structures (in Chinese). Shanghai Science & Technique Press, Shanghai, 1981
3. Building Research Institute of Fujian Province. Strength, Elasticity Modulus and Longitudinal Flexure Factor. Building Technique (internal report, in Chinese), 1974, pp. 31–50
4. Qian Yiliang: Strength of Masonries and its Variation. Proceedings on Research on Masonries (in Chinese). Hunan University Press, Changsha, 1989, pp. 14–25.
5. Ding Dajun: Concise Book on Mansory Structures (in Chinese). Shanghai Science & Technique Press, Shanghai, 1957
6. Ding Dajun: Discussion on several problems in design of masonry structures. Journal of Nanjing Architectural and Civil Engineering Institute (in Chinese), (1994) 17–27
7. National Standard of PRC, Design Code of Masonry Structures (GBJ-88) (in Chinese). China Building Industry Press, Beijing, 1988
8. Shi Chuxian: Several Factors Influencing Compressive Strength of Masonries. Ibid. as No. 4, pp. 32–44
9. Chen Xingzhi–Li Yue: Influence of Early Compression of Brick Masonry on its compressive strength. Ibid. as No. 4, pp. 45–50
10. Sun Ming: Influence of Fullness of Mortar in Joints of Brick Masonry on its Compressive Strength. Building Research Institute of Sichuan Province. Building Information Materials (in Chinese), No. 7508, 1973, pp. 25–29
11. Technical Code of Work and Acceptance for Masonry Engineering (GBJ203-83) (in Chinese). China Building Industry Press, Beijing, 1983
12. Ding Dajun et al.: Masonry Structures (in Chinese). Dadong Press, Shanghai, 1954
13. Shi Chuxian: Compressive Elasticity Moduli of Masonries. Ibid. as No. 4, pp. 116–122
14. Ding Dajun: Study of failure mechanism of structures and creation of new theories of masonry strength industrial construction (in Chinese). Ibid. as No. 2, 1994, pp. 32–35
15. Ding Dajun: Two New Points of Views in Theory of Masonry Strength. Proceedings of the 8th International Conference on Brick/Block Masonry Structures, Dublin, Ireland, Sept. 1988, pp. 1531–1538
16. Ding Dajun: Same title as that of 15, Civil Engineering Review (Mélyépítéstudományi Szemle; in Hungarian), March, 1989, pp. 95–112
17. Tang Daixiu: Tests and Calculation Method of Local Compression of Masonry. Ibid. as No. 4, pp. 163–192
18. Zhi Qi: Three Strengths of Axial Tension, Flexural Tension and Shearing of Masonry. Ibid. as No. 4, pp. 51–61
19. Lu Xizhao: Compressive, Shearing Strengths and Deformation Characteristics of Clay Hollow Brick Masonry. Ibid. as No. 4, pp. 62–87
20. Ding Dajun et al.: Experimental research on strength of brick masonry reinforced with new-type transverse steel. Journal of Building Structures (in Chinese). Vol. 9, 53–58
21. Bai Aodong: Experimental Research on and Design of Longitudinally Reinforced Brick Columns. Ibid. as No. 4, pp. 282–296
22. Ding Dajun: Longitudinal bending of reinforced concrete compressive members. Journal of Nanjing Institute of Technology (in Chinese). Vol. 2, 1979, 49–61
23. Ding Dajun: Longitudinal bending of reinforced concrete compression members. Journal of Politecnico di Milano, Studi e Ricerche, 9 (1987), 443–460



# MODELING THE INFLUENCE OF ENVIRONMENTAL PARAMETERS AND RESIDENCE TIME ON MICROBIAL DRINKING WATER QUALITY IN DISTRIBUTION SYSTEMS

Dombay, G.-Piriou, Ph.-Dukan, S.-Kiene, L.

*Lyonnaise des Eaux CIRSEE. 38 rue du Président-Wilson, 78230 Le Pecq, France*

## 1. Introduction

During its distribution drinking water undergoes several quality changes which is accounted to the activity of the distribution system as a complex physicochemical and biological reactor. These phenomena can be classified as deterioration of organoleptic parameters, bacteriological water quality deterioration, nitrification, corrosion, post-precipitation, and deposit formation. Consequently the design and operation of drinking water distribution systems cannot be considered only from the hydraulic view, but reactor theory and process engineering measures are to be integrated.

Of the various causes of quality deterioration in distribution systems, microbiological parameters are the most closely studied and monitored, because of the short-term risks regarding to public health. Even if high heterotrophic plate counts (HPC) do not necessarily give rise to a health risk, they are the signature of a network in which undesirable microbial water quality changes may occur (Dukan *et al.*, 1996).

In the distribution system bulk water phase represents only a small fragment of the bacterial activity (Herson *et al.*, 1991). In the network bacteria colonizes surfaces, this attached biomass is often referred as *biofilm* (Characklis and Marshall, 1990). In biofilms bacteria is embedded in an extracellular polymer (EPS) matrix, forming an inhomogeneous layer on the substratum. Bacteriological water quality deterioration is mainly accounted to the biofilm activity in the distribution system. The presence of biofilm can protect indicator and/or pathogenic microorganisms (LeChevallier, 1990), it can generate biocorrosion (LeChevallier *et al.*, 1993), cause taste and odor problems (Burlingame and Anselme, 1995). Substantial biofilm activity in the network might promote the presence of higher organisms (fungi, yeast, protozoa, microzooplankton, invertebrates) in the network, resulting further water quality deterioration (Levy *et al.*, 1986; Fass *et al.*, 1996). Public health risk cannot only be characterized by accidental coliform occurrence, but there is a growing belief that some heterotrophic bacteria are opportunistic pathogens, and their growth should be controlled (Payment *et al.*, 1994). These research results



indicate the importance of understanding bacterial regrowth phenomena in the drinking water network.

To understand and describe bacterial regrowth phenomena in drinking water distribution systems, deterministic modeling is one of the most appropriate tools available to researchers to test hypotheses, and to engineers to implement water quality management in the network.

The objective of this article is to present the major concepts of bacterial regrowth modeling for distribution systems and to show a particular model application. In the first part the article describes the processes which were taken into account by the bacterial regrowth model developed by the CIRSEE research center of the Lyonnaise des Eaux group, and points out the aspects which are to be integrated in the forthcoming model development. In the second part the article shows how modeling can contribute to the understanding of the relationship between environmental parameters (temperature, substrate concentration, inlet bacteria concentration) and bacterial dynamics in the drinking water distribution system.

## 2. Modeling bacterial regrowth phenomena in the network

### 2.1. The influence of water quality parameters on bacterial dynamics

From the reactor engineering view, bacterial dynamics in the network is influenced by

- water quality characteristics, determined by environmental parameters and treatment technology, and
- hydraulic conditions, determined by properties and operation of the distribution system.

The behavior of a biofilm reactor is primarily influenced by the actual water quality parameters. These parameters determine the environment for the bacterial proliferation, the activity of the biofilm, hence the bacterial regrowth phenomena which is responsible for the microbial drinking water quality deterioration in the distribution system. Drinking water quality parameters are set by the resource of the water, modified by the treatment processes, and changing spatially and in time during the distribution.

To model bacterial regrowth phenomena in distribution systems, hydrodynamic modeling has to be coupled with biofilm kinetics, to describe the propagation and reaction of the water quality parameters.

The model was developed for an ideal plug-flow reactor. The applied transport equation for the bulk phase is the one-dimensional axial-dispersion model, where only convective transport is considered:

$$\frac{\partial C_i}{\partial t} + v \frac{\partial C_i}{\partial y} = \text{reaction}_i, \quad (1)$$

where  $C_i$  [ $\text{kg}/\text{m}^3$ ] is the concentration of a particular water quality parameter,  $v$  [ $\text{m}/\text{s}$ ] is the mean flow velocity,  $y$  [ $\text{m}$ ] is the distance in the axial direction,  $\text{reaction}_i$  [ $\text{kg}/\text{m}^3\text{s}$ ] is the rate of reaction for parameter  $i$ .

In the biofilm constituent transport occur due to molecular diffusion:

$$\frac{\partial C_i}{\partial t} - D_{\text{biof}} \frac{\partial^2 C_i}{\partial x^2} = \text{reaction}_i, \quad (2)$$

where  $D_{\text{biof}}$  [ $\text{m}^2/\text{s}$ ] is the effective diffusivity coefficient in the biofilm,  $x$  [ $\text{m}$ ] is the distance in the radial direction from the biofilm surface.

Major processes in a biofilm reactor include growth, death, adhesion and detachment, which represents a high grade of interactions between suspended and attached biomass. These processes are taken into account by the model according to Fig. 1 (Dukan *et al.*, 1996). Fixed biomass is modeled as a layer uniformly distributed over the pipe surface, expressed as an equivalent thickness of a carbon layer. By this way, it is possible to distinguish between phenomena depending on their locations: reactions in the bulk water phase, reaction at the water-biofilm surface interface and within the biofilm. Mass balance equations are set for substrate, suspended biomass (total [X] and active [XH] fraction), and fixed biomass (total [B] and active [Bib] fraction).

*Mass balance equations in the bulk phase* [applying Equation (1)]:

Substrate balance:

$$\text{reaction}_S = -S \text{ diffusion into the biofilm} + \text{release by XH} - \text{consumption by XH} \quad (3)$$

Free bacteria balance:

$$\text{reaction}_X = + \text{growth of XH} + \text{detachment of B} - \text{deposition of X} \quad (4)$$

Free active bacteria balance:

$$\begin{aligned} \text{reaction}_{\text{XH}} = & + \text{growth of XH} + \text{detachment of Bib} - \text{deposition of XH} - \\ & - \text{death of XH} \end{aligned} \quad (5)$$

*Mass balance equations in the biofilm* [applying Equation (2)]:

Substrate balance:

$$\text{reaction}_S = - \text{consumption by Bib} + \text{release by Bib} \quad (6)$$

Fixed bacteria balance:

$$\text{reaction}_B = + \text{growth of Bib} - \text{detachment of B} + \text{deposition of X} \quad (7)$$

Fixed active bacteria balance:

$$\begin{aligned} \text{reaction}_{\text{Bib}} = & + \text{growth of Bib} - \text{detachment of Bib} - \text{death of Bib} + \\ & + \text{deposition of XH} \end{aligned} \quad (8)$$



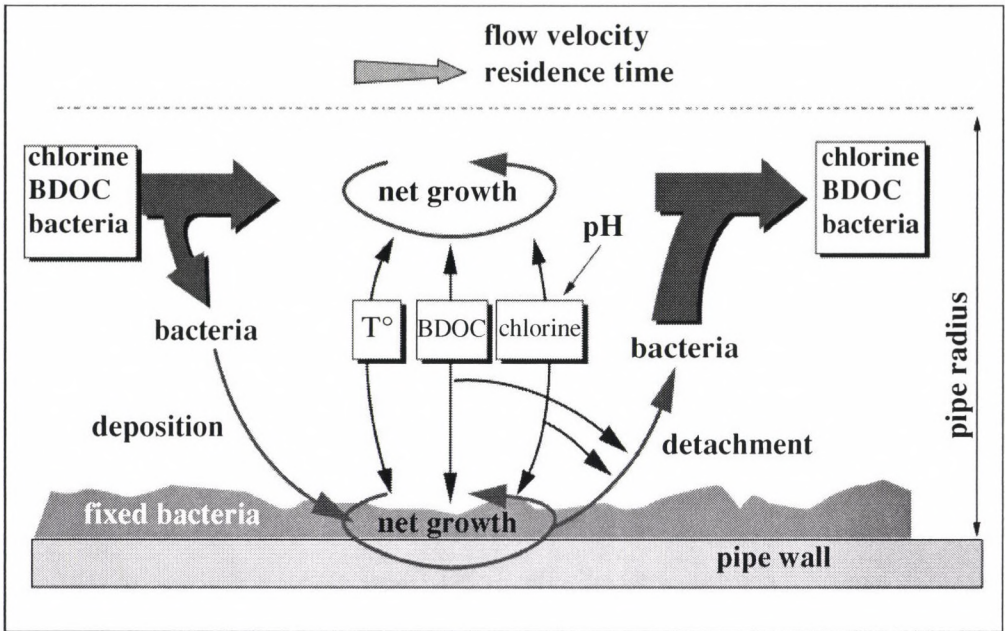


Fig. 1. Processes of the model

### Modeling bacterial growth

For heterotrophic (HPC) bacteria substrate is available in the drinking water as the biodegradable fraction of organic carbon compounds. The biodegradable dissolved organic carbon (BDOC) is generally considered the limiting factor of bacterial growth in drinking water (LeChevallier, 1990). Based on the assumption that the concentration of one single growth limiting nutrient determines the growth rate of bacterial proliferation, the model of Monod is most widely used:

$$\text{Consumption} = V_{\max}(T) \frac{S}{S + K_S} \cdot \text{active bacteria} . \quad (9)$$

Numerous studies showed correlation between BDOC and HPC values in distribution systems (Laurent *et al.*, 1993; van der Kooij *et al.*, 1995). Pilot scale experiments (Clark *et al.*, 1994; Piriou *et al.*, 1997) and modeling results (Dukan *et al.*, 1996; Bois *et al.*, 1997; Laurent *et al.*, 1997) indicate that biofilm evolution is mainly due to its growth, related to BDOC concentration. Peyton (1996) showed, that steady-state biofilm thickness is dependent on substrate loading rate. In most of his experiments the substrate utilization rate was nearly equal to the substrate loading rate. Based on literature data, the author emphasizes that over a certain level of substrate loading rate further increases do not effect steady-state biofilm thickness.



Bacterial growth is influenced by temperature. Temperature changes in the distribution system occur due to seasonal variations. Based on full scale experiments, bacterial regrowth problems are frequently associated with high temperature periods (Donlan *et al.*, 1994; Mathieu *et al.*, 1995; Amblard *et al.*, 1996). Contrary, HPC peaks in the network observed by Kerneis *et al.* (1995) did not occur in the warmest periods. According to the authors, high suspended HPC counts could not only be related to water temperature, but other factors, such as BDOC and inlet HPC, as well. Based on in situ biofilm examinations Holden *et al.* (1995) found that the increase in temperature amplified biofilm growth. It was noted that measured substrate concentrations decreased with increasing temperature, due to the increased bacterial substrate uptake.

The effect of temperature can be taken into account in Eq. (9) on the maximum growth rate ( $V_{\max}$ ) of bacteria with a sigmoidal relationship (see Appendix, A1). As Fig. 1 shows that the model takes into account BDOC consumption by attached and fixed bacteria, and also BDOC release during lysis. Equations (3)–(8) are based on (9) in terms of biomass growth and substrate consumption.

### Mortality

*Natural mortality* can be accounted to senescence and grazing by microzooplanktons. It can be modeled first order in relation to the quantity of active bacteria.

*Chlorine induced mortality.* To mitigate public health risk, during its distribution drinking water should contain biocide agents in an adequate concentration. Chlorine is the most widely used disinfectant in drinking water. Since the thorough chemical kinetics of chlorine decay is not known, chlorine decay models are generally based on pseudo-first order kinetics, assuming, that the concentration of the species reacting with the chlorine are much greater than that of chlorine (Chambers *et al.*, 1995). The model takes into account chlorine decay kinetics under the influence of pH, temperature, hydraulics and pipe materials (A18). The chlorine induced mortality of free bacteria is expressed by the inhibiting action of free chlorine as the "suicide inhibitor" (Dukan *et al.*, 1996). The mortality rate takes into account the different forms of chlorine in water ( $\text{HClO}/\text{ClO}^-$ ) depending on pH (A20).

In order to explain the greater resistance to chlorine of fixed bacteria compared with free bacteria (LeChevallier *et al.*, 1988), the hypothesis was made that chlorine diffusion within the biofilm takes place over a very low thickness of equivalent carbon. Hence the low diffusivity of free chlorine into the biofilm enables the identification of 2 biofilm layers: a chlorinated layer and a layer not attained by chlorine. This latter layer is accounted to the stronger resistance of fixed bacteria in relation to free bacteria against chlorine.

Consequently the mass balance equations (6) and (8) for the biofilm exist in two forms, for the chlorinated layer, (A5) and (A16), and the non-chlorinated layer, (A6) and (A14). The mortality is expressed by (A19).

## Detachment, deposition

*Detachment* is the key factor which

- balances attached microbial growth hence governs the steady-state biofilm condition, and
- responsible for the increase of suspended bacteria in the network.

For the chlorinated biofilm layer detachment rate was considered proportional to the number of fixed bacteria. For the non-chlorinated layer detachment rate is proportional to the number of fixed bacteria and their growth rate. (Detailed explication is given by Stewart, 1993; Dukan *et al.*, 1996.)

*Deposition* is the initial process of biofilm formation. Under dynamic conditions inlet total and active bacteria concentration has an effect on bacterial colonization of surfaces, hence biofilm formation. Transport could be a rate limiting step in bacterial colonization (Wolfaardt and Cloete, 1992; Mueller, 1996). Piriou *et al.* (1997) examined the influence of inlet bacteria concentration on biofilm behavior using a pipe loop pilot. The results indicate that the inlet total bacteria has a positive influence on biofilm evolution, but the inlet active bacteria has no impact on biofilm activity. They conclude that bacteria deposition has little effect on biofilm evolution under steady-state condition.

According to Characklis and Marshall (1990) bacterial deposition can be described by first-order kinetics in relation to bacteria concentration. Based on the aforementioned results it is to be mentioned that under dynamic conditions the first-order assumption might not describe the deposition process adequately. Deposition modeling under dynamic conditions has to be investigated.

Model equations are given in Appendix, based on (3)–(8) mass balance equations, and the described kinetic considerations. (More thorough details of each process are given by Dukan *et al.*, 1996.)

### 2.2. The influence of hydraulic parameters on bacterial dynamics

Presently the described bacterial regrowth model is solved for a tubular reactor, in which plug-flow conditions prevail. For the application of this model to drinking water distribution systems certain developments might be necessary which take into account the effect of network hydrodynamics

- on transport processes (axial and radial diffusion, mixing), and
- on biofilm kinetics.

Biofilms are influenced by hydraulic conditions, and in turn, biofilms may influence the hydrodynamic conditions close to the surface of and within biofilms (Wilderer *et al.*, 1995). During the operation of a drinking water network various hydraulic conditions occur in the system, determined by the actual water consumption and the status of operation (e.g. pumping). From the biofilm reactor's point of view, the magnitude and dynamic range of hydraulic parameters can be unique and significantly different from any biofilm reactors used for treatment processes or applied in laboratory examinations.



### Flow velocity

In tubular reactors increasing water velocity leads to

- an increase of the convective transport of bacteria, substrate, chlorine and other constituents in the reactor,
- an increase in turbulent diffusion,
- a decrease of the thickness of the boundary layer causing a reduction in the mass transfer resistance towards the biofilm,
- an increase of shear stress at the pipe wall.

Transport processes in the bulk are controlled by convection and turbulent diffusion, in the biofilm by molecular diffusion. Under turbulent conditions in the bulk virtually there is no concentration gradient in radial direction. The concentration gradient evolves in the concentration boundary layer (CBL) and reaches its maximum at the biofilm surface (Lewandowski *et al.*, 1994). In the CBL diffusion is the predominant mass transfer mechanism. The velocity profile in the hydrodynamic boundary layer (HBL), where the flow becomes unidirectional, is uncertain.

Bishop *et al.* (1997) showed, that there is a significant difference between the hydrodynamic boundary layer and the concentration boundary layer above a biofilm. CBL varied between 200–400  $\mu\text{m}$  depending on the water velocity in the bulk. HBL and flow velocity had only minimal relationship, the thickness is around 4000–5500  $\mu\text{m}$ . Increasing fluid velocity decreases the thickness of the CBL, causing a decrease in the mass transfer resistance (Characklis and Marshall, 1990).

In drinking water distribution systems the biofilm thickness is low (around 10–30  $\mu\text{m}$ , L'Hostis, 1996) and likely there is no transport limitation in the biofilm. Consequently it might be assumed, that variation in substrate transport due to different flow velocities does not have a significant effect on biofilm behavior. Under these conditions it would not be necessary to modelise the effect of hydrodynamics on substrate transport.

### Biofilm detachment and hydrodynamics

Despite its significance and the number of regarding studies, so far is very little has been understood about the mechanism of detachment (Peyton *et al.*, 1993; Ohashi *et al.*, 1996). Presently the effects of hydrodynamics on biofilm detachment are not clarified. Currently available analytical techniques provide a rather static picture, and the dynamic nature of biofilm systems is not adequately resolved (Wilderer *et al.*, 1995). For transient hydrodynamic conditions Peyton and Characklis (1993) proposed a detachment model, based on the results of Bakke (1986). Bakke showed that the variation of shear stress causes increased detachment, and after the transitory period detachment returns to its former value. Hence in the model not the magnitude of the shear stress ( $\tau$ ), but its derivative is included:

$$R_d = R_{de} + R_{d\Delta\tau}, \quad (10)$$



$$R_{d\Delta\tau} = \begin{cases} k_d \frac{d\tau}{dt}, & \frac{d\tau}{dt} > 0 \\ 0, & \frac{d\tau}{dt} \leq 0 \end{cases}, \quad (11)$$

where  $R_d$  is overall detachment rate [ $s^{-1}$ ],  $R_{de}$  is detachment due to erosion [ $s^{-1}$ ],  $R_{d\Delta\tau}$  is detachment due to transitory shear stress [ $s^{-1}$ ],  $k_d$  is coefficient [dimensionless].

Ohashi *et al.* (1996) examined the adhesion strength of biofilms. Adhesion strength is the resistance of the biofilm versus detachment when an external force is induced. Biofilms in conduits might be exposed to tensile stress ( $\sigma$ ) by lift force by fluid dynamics due to hydrodynamic vibrations or pressure fluctuations. The authors determined biofilm shear strength by collision force, and biofilm tensile strength by centrifugation force. The results indicate that shear strength was all the time much larger than tensile strength by a magnitude of two orders (2 log). Hence it is suggested that biofilm detachment is mainly caused by the decline of adhesion strength by tensile force rather than by shear force. Due to the applied methods, these result are rather applicable to transitory phenomena.

Based on this research result we propose the following modification of the (10) transitory biofilm detachment model:

$$R_d = R_{de} + R_{d\Delta\tau} + R_{d\Delta\sigma}, \quad (12)$$

$$R_{d\Delta\sigma} = \begin{cases} k_{d\sigma} \left| \frac{d\sigma}{dt} \right|, & \frac{d\sigma}{dt} < 0 \\ 0, & \frac{d\sigma}{dt} \geq 0 \end{cases}, \quad (13)$$

where  $R_{d\Delta\sigma}$  is detachment due to transitory tensile stress [ $s^{-1}$ ],  $k_{d\sigma}$  is coefficient [dimensionless].

The model expresses that increasing tensile stress toward the substratum cannot cause detachment, but a decreasing tensile stress (e.g. pressure drop or relaxation) can.

Presently the validation of the model is problematical due to the lack of adequate on-line biofilm monitors. In addition, the examination of the applicability of the model to distribution systems, and the description of the relationship between tensile and shear stresses on biofilm detachment are future research tasks.

### Residence time

Water quality in the distribution system changes spatially and changes in time. Hydraulic conditions determine residence time and residence time distribution in the network. In bacterial regrowth phenomena residence time certainly influences bacterial dynamics, since these processes are time dependent. In drinking water networks numerous studies examined the effect of residence time on water quality, and its correlation to

various parameters such as substrate concentration, HPC, and chlorine decay (e.g. Heraud *et al.*, 1997; Kerneis *et al.*, 1995; van der Kooij, 1992; Prévost *et al.*, 1997).

Considering the complexity of the phenomena influencing bacterial regrowth in the network (and taken into account by the model) it has to be noted that residence time cannot correlate directly with microbial water quality parameters. Consequently it is an excessive simplification to consider residence time as a single decisive parameter in microbial water quality changes, regardless to actual environmental parameters.

In Chapter 2 the effect of residence time on the relative influence of environmental parameters is examined, by the application of the bacterial regrowth model.

### **3. Model application: quantifying the influence of environmental parameters and residence time on bacterial regrowth phenomena**

The study focuses on four water quality parameters which thought to be the most influential, such as inlet total and active bacteria concentration, substrate concentration, and temperature. Due to their nature these parameters are referred as environmental parameters. As described in Chapter 1.1, the literature indicates that there might be significant interactions between these parameters. Despite their presumed importance, so far no effort has been made to quantify their relative influence in a comparable way.

The goal of the study was to quantify the influence of the aforementioned environmental parameters and residence time, as the most considerable hydraulic parameter, on bacterial dynamics in the distribution system, using the described bacterial regrowth model under steady-state conditions.

#### *3.1. The applied method*

In order to quantify influences (effects) of these environmental parameters on response variables, a complete experimental design at two levels was applied (Sado and Sado, 1991). For the application the following criteria had to be taken into account:

1. Four environmental parameters were to be investigated, namely X, XH, S, and T. Due to the expected high level of interactions between these parameters, fractional experimental designs were not applicable. (With fractional designs interactions are confounded with effects which would result in a non-quantifiable result.) Consequently a complete experimental design was carried out.

2. The experimental design was implemented at two levels. It made possible the quantification of the effects of parameters in a given range hypothesizing a linearity between the parameters and response variables. By this approach the  $2^4$  experimental design was realized as shown in Table 1.

The experimental range was defined by common values characterizing the drinking water. Temperature range was defined in a way to work with a closely linear section of the sigmoid curve (being  $T = 18$  °C the sigmoid temperature). The range for each parameters is given by Table 1.



*Table 1*  
The applied  $2^4$  experimental design

Experiment	X	XH	S	T
1	-	-	-	-
2	+	-	-	-
3	-	+	-	-
4	+	+	-	-
5	-	-	+	-
6	+	-	+	-
7	-	+	+	-
8	+	+	+	-
9	-	-	-	+
10	+	-	-	+
11	-	+	-	+
12	+	+	-	+
13	-	-	+	+
14	+	-	+	+
15	-	+	+	+
16	+	+	+	+

Level -	$10^4$	$10^1$	0.25	16
Level +	$10^6$	$10^3$	1.00	20
	bacteria/ml	bacteria/ml	mg C/l	°C

Residence time was taken into account up to 30 days with a 3-hour time step. Modeling was carried out with the absence of chlorine residual, for steady-state conditions.

3. Influences (effects) were examined for four response variables of the model, namely free and fixed bacteria, both total and active fraction (X, XH, B, Bib).

The matrix of the experimental design is the Hadamard matrix (H), based in Table 1. The additional 11 columns of the matrix are composed by the 6 first order, 4 second order, and 1 third order interactions of the four environmental parameters, and the 12th column is the unity vector for the calculation of the mean value (I).

For each response variable a result matrix ( $Y_X, Y_{XH}, Y_B, Y_{Bib}$ ) is composed. The rows of the response matrices include the values of the response variable of the actual experiment ( $n$ ), calculated by the model. Each column corresponds to a particular residence time ( $rst$ ). (E.g. the vector  $Y_{XH}(7, rst)$  represents the evolution of free active bacteria for the 7th experiment, where the environmental parameters were  $X=10^4$  bacteria/ml,  $XH=10^3$  bacteria/ml,  $S=1.00$  mg C/l, and  $T=16^\circ\text{C}$ .)

A matrix of effects is calculated for each response variable:



$$E_X = \frac{1}{16} H^t Y_X, \quad (14)$$

$$E_{XH} = \frac{1}{16} H^t Y_{XH}, \quad (15)$$

$$E_B = \frac{1}{16} H^t Y_B, \quad (16)$$

$$E_{Bib} = \frac{1}{16} H^t Y_{Bib}, \quad (17)$$

where  $H^t$  is the transposed Hadamard matrix. (Each row of a matrix of effects corresponds to the effects, interactions, and the mean value, determined by the transposed Hadamard matrix.)

The relative effects of the environmental parameters are expressed in percentage in relation to the mean value [ $I(rst) = E_i(16, rst)$ ] for each residence time step. For the  $i$  response variable:

$$RE_i^X(rst) = \frac{E_i(1, rst)}{E_i(16, rst)}, \% \quad (18)$$

$$RE_i^{XH}(rst) = \frac{E_i(2, rst)}{E_i(16, rst)}, \% \quad (19)$$

$$RE_i^S(rst) = \frac{E_i(3, rst)}{E_i(16, rst)}, \% \quad (20)$$

$$RE_i^T(rst) = \frac{E_i(4, rst)}{E_i(16, rst)}, \% \quad (21)$$

where  $i = \{X, XH, B, Bib\}$ .

With this method effects can be studied independently to the magnitude of the response variable.

### 3.2. Results and discussion

#### Effects on attached bacteria

The effects on *active attached biomass* can be seen in Fig. 2. The effect of substrate concentration is superior, its almost 100% influence is slightly decreasing in the course of residence time. Inlet active suspended bacteria has virtually no influence for the tested data range. These indicate that biofilm activity in the network is growth driven, the deposition of active bacteria plays no part under steady-state condition.

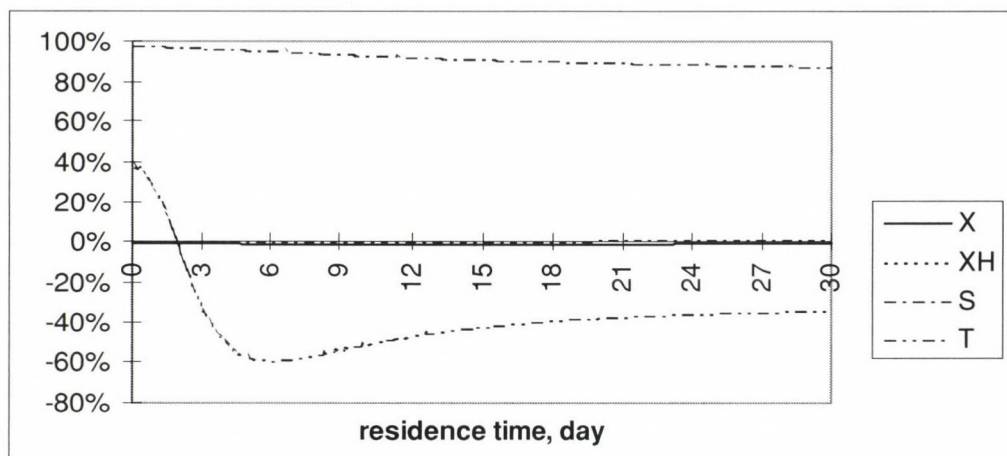


Fig. 2. Relative effects on active attached bacteria

While bacterial growth is influenced by temperature, the positive and substantial initial effect of temperature is evident. As can be seen in Fig. 2, this influence is characteristically decreasing with residence time, after 2 days alternates to negative, and stabilizing after about 3 weeks. It indicates that higher water temperature results in a less significant biofilm activity. This virtual paradox can be explained by Fig. 3. In this figure the results of experiment 8 and 16 are shown. All parameters are identical but temperature. In Fig. 3a the decrease in BDOC concentration is plotted. As expected, higher water temperature hence higher bacterial activity causes a more rapid substrate consumption. As biofilm growth is very strongly substrate dependent (as shown in Fig. 2), it is very responsive to changes in substrate concentration. Figure 3b shows that although higher temperature results in a more significant biofilm activity at low residence time, due to the rapid BDOC consumption, biofilm activity is constantly less after a certain period of time. This phenomena explains the negative effect of temperature on active attached bacteria.

Inlet total suspended bacteria obviously does not effect biofilm activity.

*Total attached bacteria* incorporates active and dead biofilm bacteria. Figure 4 shows that its amount is mainly determined by substrate concentration hence growth. For higher residence times deposition increases total attached biomass, this effect can be entitled to the less significant biofilm growth at higher residence times due to the diminution of BDOC. Consequently inlet total suspended bacteria has an effect on total biofilm mass. It is to be emphasized that total attached biomass does not impose a public health risk, and in distribution systems its amount does not cause substantial friction losses.

The effect of temperature has similar characteristics as for active attached biomass, although the amount of total biomass is not constantly inferior in case of higher temperature after a certain residence time, as in the experiment after 12 days its effect alternates to positive again. Inlet active suspended bacteria has no effect.

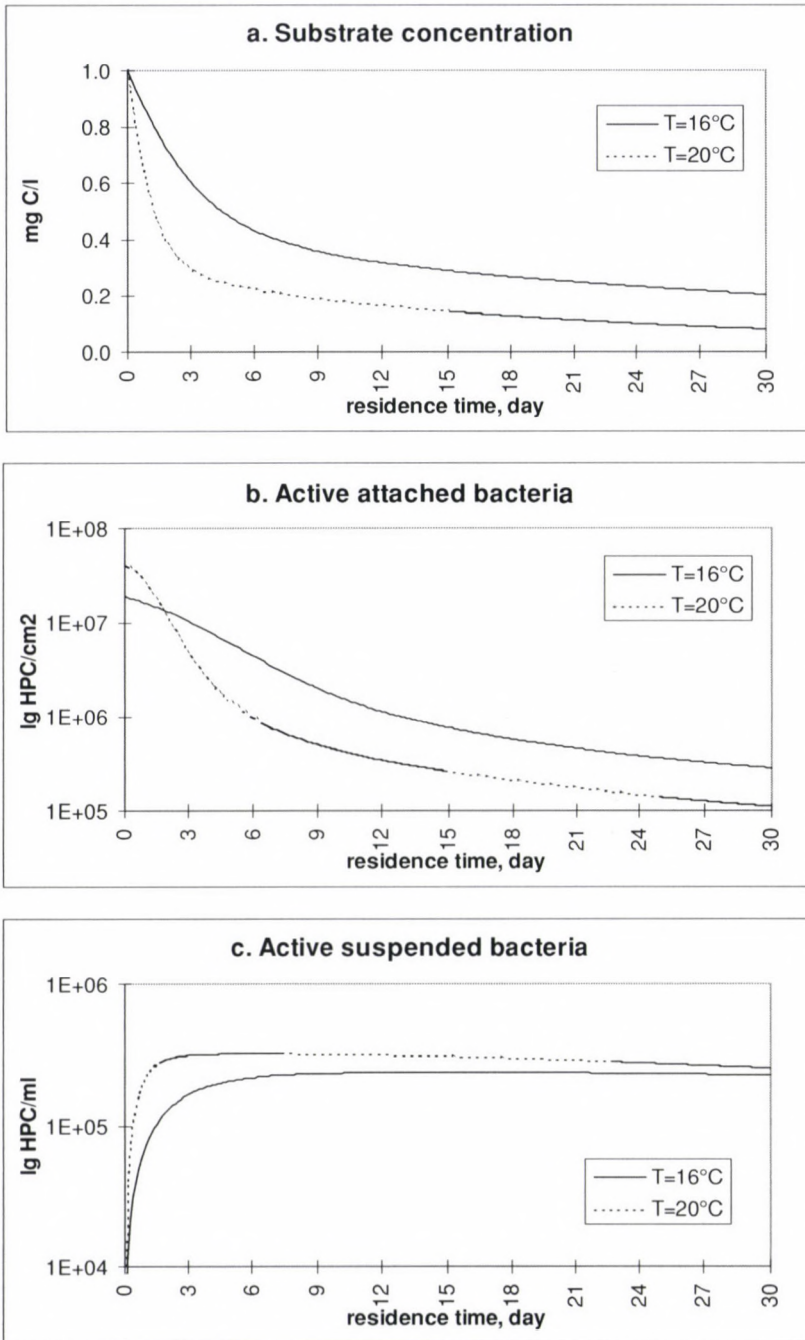


Fig. 3



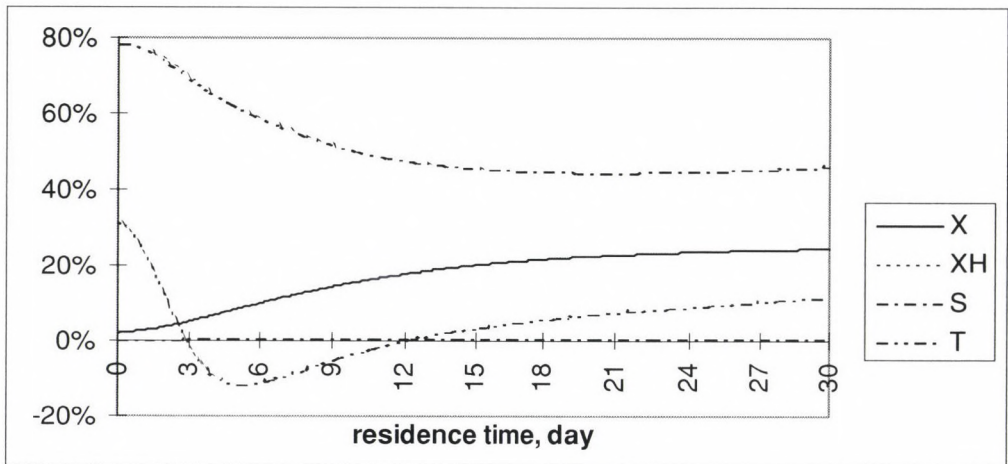


Fig. 4. Effects on total attached bacteria

#### Effects on suspended bacteria

*Active suspended bacteria* is the primary response variable of the model, since bacteriological problems are imposed by this fraction.

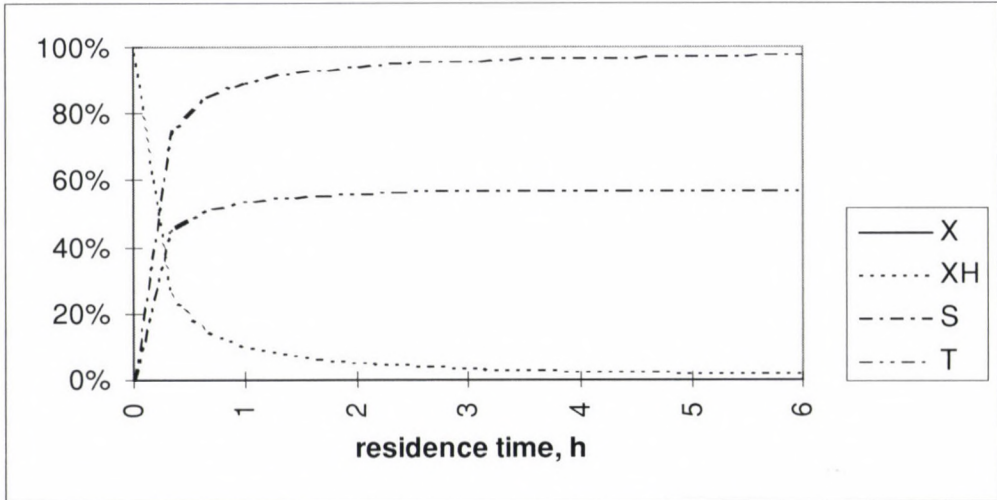
Figure 5a shows the period of the initial 6 hours residence time. At the inlet, the amount of the suspended active bacteria is determined by its inlet concentration. Its effect is rapidly decreasing in an exponential way in residence time, substituted by the effect of substrate concentration. It indicates that active bacteria concentration in the bulk is determined by biofilm activity, namely the detachment of active bacteria from the biofilm. As can be seen in Fig. 5b, in the long run the effect of initial active suspended bacteria is zero, hence inlet active bacteria has an influence only in the initial few hours, depending on its range. For the calculation low inlet HPC counts were used, between 1 log and 3 log, obviously for a concentration of 6 log HPC/ml its effect could not be neglected. (Although in practice, with disinfection, this concentration does not occur.)

These results show that biological stability of the drinking water is not only a relative concept in terms of inlet HPC concentration but depends on residence time as well.

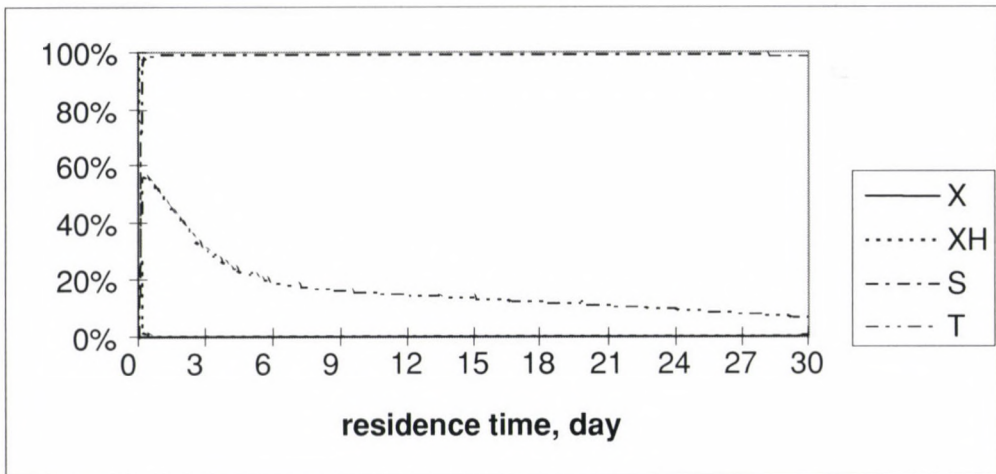
Temperature has a positive effect which is decreasing with residence time. Figure 3c shows the influence of temperature on active suspended bacteria based on experiment 8 and 16. Due to the initially higher biofilm activity HPC counts reach their maximum earlier in case of higher temperature. The impact of the more rapid BDOC decrease results in a slight decrease of HPC in the course of residence time. The graph shows that active suspended bacteria concentration is practically stabilizing after a few days of residence time. It indicates that longer residence times do not necessarily result in a worse bacteriological water quality. (Theoretically, for long enough residence times, after the consumption of BDOC hence the lack of bacterial growth, bacteriological water quality

improves. This decrease in active suspended bacteria can be seen in case of 20 °C in Fig. 3c.)

The effects on *total suspended bacteria* can be seen in Fig. 6. Due to the high total inlet bacteria concentrations (4 and 6 log/ml), the total suspended bacteria concentration



5a. For the initial period of 6 hours residence time



5b. For the period of 30 days residence time

Fig. 5. Effects on active suspended bacteria

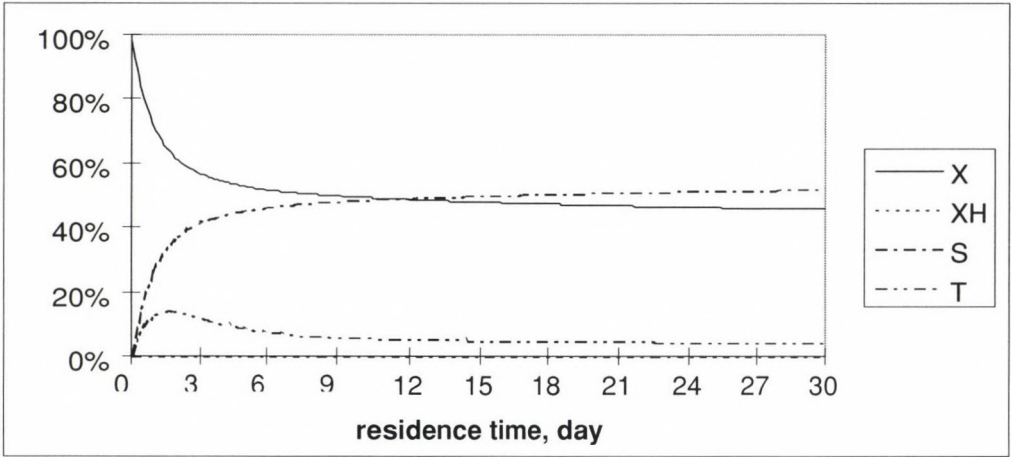


Fig. 6. Effects on total suspended bacteria

is determined by not only biofilm activity, but this inlet concentration as well. Due to mortality and biofilm detachment, the initially decisive influence of inlet total bacteria is decreasing in residence time. In this experiment after 12 days of residence time the influence of the biofilm becomes superior, but still very close to the influence of the inlet concentration.

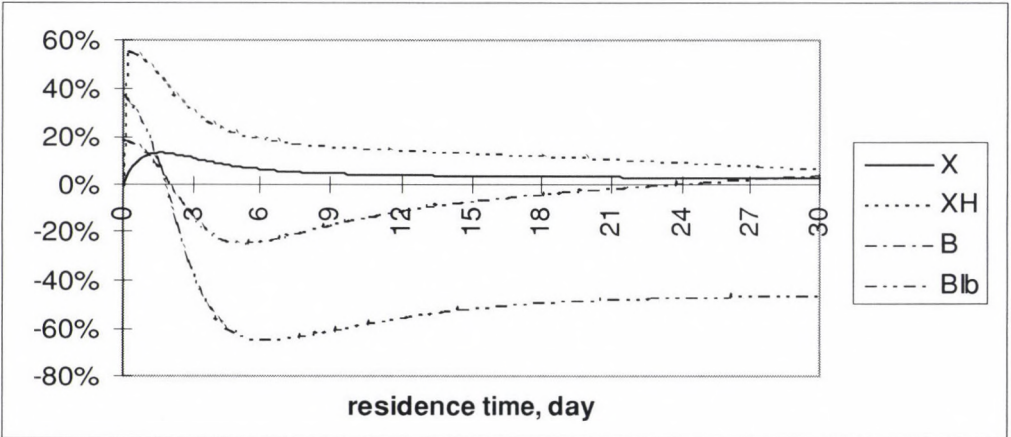


Fig. 7. Interactions between substrate concentration and temperature



### Interactions between parameters

Simulations showed a significant interaction between substrate concentration and temperature. These factors influence mainly the bacterial growth, hence biofilm activity, which is the major phenomena in bacterial dynamics in distribution systems.

Figure 6 shows the interactions for the response variables of total suspended bacteria, active suspended bacteria, total attached bacteria, and active attached bacteria. These interactions change according to residence time. Interactions are more significant for the active fractions. For suspended bacteria the interactions are positive. In case of the active attached biomass after a few days of residence time there is a substantial negative interaction between BDOC and temperature, which can be also explained by Fig. 3a, b.

Based on these results it can be seen that the application of complete experimental designs is necessary due to the high level of interactions. This might impose a problem in case of a sensitivity analysis, where due to the high number of constants a complete plan is not a feasible solution.

### 4. Summary and conclusions

Increasing levels of consumer expectations and new standards of regulation necessitate to consider drinking water supply system as a homogeneous system, including the water resource, treatment technologies and distribution conditions. To overcome unacceptable microbial water quality degradation, system operators should provide drinking water with the least possible amount of BDOC, and maintaining an adequate chlorine residual level in the network. Post-chlorination strategies can only be carried out with chlorine decay modeling in the network. The influence of inlet drinking water characteristics, post-chlorination, and network characteristics on the bacterial dynamics of the distribution systems can only be analysed by modeling approach.

Bacterial regrowth modeling is an indispensable tool to study and analyses scenarios of water quality changes. The model developed by the Cirsee research center incorporates the major phenomena that govern bacterial regrowth phenomena in the network. The development of the model to include the necessary influences of hydrodynamics on biofilm behaviour, furthermore the model application under dynamic conditions are future research tasks.

Associating the modeling approach with experimental design theory, the quantification of the relative influence of environmental parameters was carried out, in the course of residence time. Simulations showed that

- The relative influence of environmental parameters changes in residence time.
- Biofilm activity is primarily influenced by substrate concentration and temperature.
- Higher water temperature results in less significant biofilm activity for longer residence times, due to the rapid decrease of substrate consumption, induced by increased bacterial activity.

– Under steady-state condition, inlet active bacteria has an influence on bacterial water quality only in a relative way, which depends on its magnitude, residence time and biofilm activity in the network. Steady-state biofilm behavior is virtually not effected by active suspended bacteria.

– The simulations showed strong interaction between temperature and BDOC for each response variable. Based on these result it is to be indicated that sensitivity ana-lyses of the bacterial regrowth model has to be done in the course of residence time, with an adequate experimental design to prevent the confounding of effects with significant interactions.

– Drinking water can be considered biologically stable when bacterial regrowth phenomena is insignificant in the network. This concept does not only refers to a strictly nutrient limited milieu, but implies a relative measure. Bacterial regrowth can be considered insignificant in relation to inlet HPC counts. This indicates an indirect influence of inlet active bacteria.

## References

1. Amblard, C.–Bourdier, G.–Carrias, J.-F.–Maurin, N.–Quiblier, C. 1996. Evolution saisonniere de la structure des communautes microbiennes dans un reservoir d'eau potable. *Water Research* 30: 613–624
2. Bakke, R. 1986. *Biofilm Detachment*. Bozeman, MT: Montana State University
3. Bishop, P. L.–Gibbs, J. T.–Cunnigham, B. E. 1997. Relationship between concentration and hydrodynamic boundary layer. *Environmental Technology* 18: 375–386
4. Bois, F. Y.–Fahmy, T.–Block, J.-C.–Gatel, D. 1997. Dynamic modeling of bacteria in a pilot drinking-water distribution system. *Water Research* 31: 3146–3156
5. Burlingame, G. A.–Anselme, C. 1995. Distribution system tastes and odors. In: *Advances in Taste-and-Odor Treatment and Control*, eds I. H. Suffet, J. Mallevalle, E. Kawczynski. pp. 281–319. Denver, CO: AWWA
6. Chambers, V. K.–Creasey, J. D.–Joy, J. S. 1995. Modelling free and total chlorine decay in potable water distribution systems. *J Water SRT – Aqua* 44: 60–69
7. Characklis, W. G.–Marshall, K. C. 1990. *Biofilms*. John Wiley & Sons, Inc.
8. Clark, R. M.–Lykins, B. W.–Block, J. C.–Wymer, L. J.–Reasoner, D. J. 1994. Water quality changes in a simulated distribution system. *J Water SRT – Aqua* 43: 263–277
9. Donlan, R. M.–Pipes, W. O.–Yohe, T. L. 1994. Biofilm formation on cast iron substrata in water distribution systems. *Water Research* 28: 1497–1503
10. Dukan, S.–Piriou, P.–Guyon, F.–Villon, P. 1996. Dynamic modelling of bacterial growth in drinking water networks. *Water Research* 30: 1991–2002
11. Fass, S.–Dincher, M. L.–Reasoner, D. J.–Gatel, D.–Block, J.-C. 1996. Fate of *Escherichia coli* experimentally injected in a drinking water distribution pilot system. *Water Research* 30: 2215–2221
12. Heraud, J.–Kiene, L.–Detay, M.–Levi, Y. 1997. Optimised modelling of chlorine residual in a drinking water distribution system with a combination of on-line sensors. *J Water SRT – Aqua* 46: 59–70
13. Herson, D. S.–Marshall, D. R.–Baker, K. H.–Victoreen, H. T. 1991. Association of microorganisms with surfaces in distribution systems. *Journal AWWA*: 103–106
14. Holden, B.–Greetham, M.–Croll, B. T.–Scutt, J. 1995. The effect of changing inter process and final disinfection reagents on corrosion and biofilm growth in distribution pipes. *Wat. Sci. Tech.* 32: 213–220
15. Kerneis, A.–Nakache, F.–Deguin, A.–Feinberg, M. 1995. The effects of water residence time on the biological quality in a distribution network. *Water Research* 29: 1719–1727



16. Kooij, D. v. d. 1992. Assimilable organic carbon as an indicator of bacterial regrowth. *Journal AWWA*: 57-65
17. Kooij, D. v. d.-Vrouwenvelder, H. S.-Veenendaal, H. R. 1995. Kinetic aspects of biofilm formation on surfaces exposed to drinking water. *Wat. Sci. Tech.* 32: 61-65
18. Laurent, P.-Servais, P.-Prévost, M.-Gatel, D.-Clement, B. 1997. Testing the SANCHO model on distribution systems. *Journal AWWA* 89: 92-103
19. Laurent, P.-Servais, P.-Randon, G. 1993. Bacterial development in distribution networks - study and modeling. *Water Supply* 11: 387-398
20. LeChevallier, M. W. 1990. Coliform regrowth in drinking water: a review. *Journal AWWA*: 74-86
21. LeChevallier, M. W.-Cawthon, C. D.-Lee, R. G. 1988. Inactivation of biofilm bacteria. *Appl. Environ. Microbiol.* 54: 2492-2499
22. LeChevallier, M. W.-Lowry, C. D.-Lee, R. G.-Gibbon, D. L. 1993. Examining the relationship between iron corrosion and the disinfection of biofilm bacteria. *Journal AWWA*: 111-123
23. Levy, R. V.-Hart, F. L.-Cheetham, R. D. 1986. Occurrence and public health significance of invertebrates in drinking water systems. *Journal AWWA*: 105-110
24. Lewandowski, Z.-Stoodley, P.-Altobelli, S.-Fukushima, E. 1994. Hydrodynamics and kinetics in biofilm systems - recent advances and new problems. *Wat. Sci. Tech.* 29: 223-229
25. L'Hostis, E. 1996. Détection et caractérisation de biofilms par méthodes électrochimiques. pp. 226. Paris: Université Pierre et Marie Curie
26. Mathieu, L.-Block, J. C.-Prevost, M.-Maul, A.-Bischof, R. D. 1995. Biological stability of drinking-water in the city of Metz distribution system. *J Water SRT - Aqua* 44: 230-239
27. Mueller, R. F. 1996. Bacterial transport and colonization in low nutrient environments. *Water Research* 30: 2681-2690
28. Ohashi, A.-Harada, H. 1996. A novel concept for evaluation of biofilm adhesion strength by applying tensile force and shear force. *Wat. Sci. Tech.* 34: 201-211
29. Payment, P.-Coffin, E.-Paquette, G. 1994. Blood agar to detect virulence factors in tap water heterotrophic bacteria. *Appl. Environ. Microbiol.* 60: 1179-1183
30. Peyton, B. M. 1996. Effects of shear stress and substrate loading rate on *Pseudomonas aeruginosa* biofilm thickness and density. *Water Research* 30: 29-36
31. Peyton, B. M.-Characklis, W. G. 1993. A statistical analysis of the effect of substrate utilization and shear stress on the kinetics of biofilm detachment. *Biotechnology and Bioengineering* 41: 728-735
32. Piriou, P.-Kiene, L.-Levi, Y. 1997. Study & modeling of drinking water quality evolution using a pipe loop system. In: *AWWA Annual Conference*. Atlanta, Georgia
33. Prévost, M.-Rompre, A.-Baribeau, H.-Coallier, J.-Lafrance, P. 1997. Service lines: their effect on microbiological quality. *Journal AWWA* 89: 78-91
34. Sado, G.-Sado, M-C. 1991. Les Plans d'Expériences. AFNOR, Paris.
35. Stewart, P. 1993. A model for biofilm detachment. *Biotechnology and Bioengineering* 41: 111-117
36. Wilderer, P. A.-Cunningham, A.-Schindler, U. 1995. Hydrodynamics and shear stress: report from the discussion section. *Wat. Sci. Tech.* 32: 271-272
37. Wolfaardt, G. M.-Cloete, T. E. 1992. The effect of some environmental parameters on surface colonization by microorganisms. *Water Research* 26: 527-537



## Appendix

### Bacterial regrowth model equations

The effect of temperature

$$V_{\max.}(T) = V_{\max.}(T_{opt}) \cdot \exp\left(-\left(\frac{T_{opt} - T}{T_{opt} - T_i}\right)^2\right) \quad (A1)$$

Mass balance equation in the bulk

– Organic biodegradable substance (BDOC)

$$\frac{\partial S(y,t)}{\partial t} + V \cdot \frac{\partial S(y,t)}{\partial y} = -2 \cdot D_s \cdot \frac{(S(y,t) - S_{cl2}(y,t))}{L_{int} \cdot R} + a \cdot [K_{mort} + F(y,t)] \cdot X(y,t) \cdot H(y,t) - \frac{V_{\max.} \cdot S(y,t) \cdot H(y,t) \cdot X(y,t)}{S(y,t) + K_s} \quad (A2)$$

– Free bacteria concentration

$$\frac{\partial X(y,t)}{\partial t} + V \cdot \frac{\partial X(y,t)}{\partial y} = \frac{r \cdot V_{\max.} \cdot S(y,t)}{S(y,t) + K_s} \cdot H(y,t) \cdot X(y,t) + \frac{2}{R} \cdot [K_{det1} \cdot B(y,t) + K_{det2} \cdot h(y,t)] \cdot Z - K_{dep} \cdot X(y,t) \quad (A3)$$

– Active free bacteria concentration

$$\frac{\partial X(y,t) \cdot H(y,t)}{\partial t} + V \cdot \frac{\partial X(y,t) \cdot H(y,t)}{\partial y} = \frac{r \cdot V_{\max.} \cdot S(y,t)}{S(y,t) + K_s} \cdot H(y,t) \cdot X(y,t) + \frac{2}{R} \cdot [K_{det1} \cdot B(y,t) \cdot I_b(y,t) + K_{det2} \cdot h(y,t) \cdot I_{cl2}(y,t)] \cdot Z - K_{dep} \cdot X(y,t) \cdot H(y,t) - [K_{mort} + F(y,t)] \cdot H(y,t) \cdot X(y,t) \quad (A4)$$

BDOC balance equation in biofilm

– Mass balance equation in chlorinated biofilm

$$\frac{\partial S_{cl2}(x,y,t)}{\partial t} - D_{biof} \cdot \frac{\partial^2 S_{cl2}(x,y,t)}{\partial x^2} = -\frac{V_{\max.} \cdot S_{cl2}(x,y,t) \cdot I_{cl2}(y,t) \cdot Z}{S_{cl2}(x,y,t) + K_s} +$$

$$+ b \cdot (K_{\text{mort}} + G(y, t)) \cdot I_{\text{cl2}}(y, t) \cdot Z \tag{A5}$$

– Mass balance equation in non-chlorinated biofilm

$$\frac{\partial S_b(x, y, t)}{\partial t} - D_{\text{biof}} \cdot \frac{\partial^2 S_b(x, y, t)}{\partial x^2} = - \frac{V_{\text{max}} \cdot S_b(x, y, t) \cdot I_b(y, t) \cdot Z}{S_b(x, y, t) + K_s} + b \cdot K_{\text{mort}} \cdot I_b(y, t) \cdot Z \tag{A6}$$

With:

$$\left( \frac{\partial S_{\text{cl2}}(x, y, t)}{\partial x} \right)_{x=B(y, t)} = \left( \frac{\partial S_b(x, y, t)}{\partial x} \right)_{x=B(y, t)},$$

$$- D_{\text{biof}} \cdot \left( \frac{\partial S_{\text{cl2}}(x, y, t)}{\partial x} \right)_{x=e(y, t)} = D_s \cdot \frac{S_{\text{cl2}}(e(y, t), y, t) - S(e(y, t), y, t)}{L_{\text{int}}} \tag{A7}$$

and

$$\left( \frac{\partial S_b(x, y, t)}{\partial x} \right)_{x=0} = 0$$

(A8)

Fixed bacteria balance equation

$$v(x, y, t) = \int_0^x \left( r' \cdot V_{\text{max}} \cdot \frac{S_i(z, y, t)}{S_i(z, y, t) + K_s} \cdot I_i(z, y, t) - K_{\text{detr}} \right) dz + \frac{R}{2} \cdot \frac{X(y, t)}{Z} \tag{A9}$$

With:

$$h(y, t) = w \cdot \left( \frac{cl_2(y, t)}{1 + \frac{k_a}{h_3 o^+}} - cl_{2\text{lim}} \right), \quad e(y, t) = B(y, t) + h(y, t) \tag{A10}$$

$$K_{\text{detr}}(z, y, t) = K_{\text{detr1}}, \quad I_i(z, y, t) = I_b(y, t) \quad \text{if } z \in [0, B(y, t)] \tag{A11}$$

and

$$K_{\text{detr}}(z, y, t) = K_{\text{detr2}}, \quad I_i(z, y, t) = I_{\text{cl2}}(y, t) \quad \text{if } z \in [B(y, t), e(y, t)]$$

(A12)

$$\frac{\partial e(y, t)}{\partial t} = v(e(y, t), t) \tag{A13}$$

Active fixed bacteria balance equation

– in non-chlorinated biofilm

$$\frac{\partial I_b(y, t)}{\partial t} + \frac{\partial [v(x, y, t) \cdot I_b(y, t)]}{\partial x} = A(x, y, t) \cdot I_b(y, t) \quad (\text{A14})$$

With:

$$A(x, y, t) = r' \cdot V_{\max} \cdot \frac{S_b(x, y, t)}{S_b(x, y, t) + K_s} - K_{\text{det1}} - K_{\text{mort}} \quad (\text{A15})$$

– in chlorinated biofilm

$$\frac{\partial I_{cl2}(y, t)}{\partial t} + \frac{\partial [v(x, y, t) \cdot I_{cl2}(y, t)]}{\partial x} = D(x, y, t) \cdot I_{cl2}(y, t) + K_{\text{dep}} \cdot \frac{X(y, t) \cdot H(y, t)}{Z} \quad (\text{A16})$$

With:

$$D(x, y, t) = r' \cdot V_{\max} \cdot \frac{S_{cl2}(x, y, t)}{S_{cl2}(x, y, t) + K_s} - K_{\text{det2}} - K_{\text{mort}} - G(y, t) \quad (\text{A17})$$

Chlorine disappearance kinetic

$$-\frac{dcl_2(y, t)}{dy} = V \cdot \left[ K_{\text{helo}} \frac{h_3 o^+}{h_3 o^+ + k_a} + K_{\text{clo}} \cdot \left( 1 - \frac{h_3 o^+}{h_3 o^+ + k_a} \right) \right] \cdot \left( K_w + \frac{2}{R} \cdot K_p \right) \cdot cl_2(y, t) \quad (\text{A18})$$

Mortality of fixed bacteria

$$G(y, t) = K_2 \cdot \left[ 1 - \exp \left( - \frac{B \cdot \frac{cl_2(y, t)}{2 \cdot [1 + \frac{k_a}{h_3 o^+}] - cl_{2\text{lim}}}}{1 + A \cdot \frac{cl_2(y, t)}{2 \cdot [1 + \frac{k_a}{h_3 o^+}] - cl_{2\text{lim}}}} \right) \right] \quad (\text{A19})$$

Mortality of free bacteria

$$F(y, t) = K_1 \cdot \left[ 1 - \exp \left( - \frac{B \cdot \frac{cl_2(y, t)}{1 + \frac{k_a}{h_3 o^+}}}{1 + A \cdot \frac{cl_2(y, t)}{1 + \frac{k_a}{h_3 o^+}}} \right) \right] \quad (\text{A20})$$



### Nomenclature of the model

#### Constants:

$R$  = pipe radius (m)

$L_{int.}$  = boundary layer (m)

$D_s$  = effective diffusivity of substrate in the interface ( $m^2 \cdot s^{-1}$ )

$a$  = percentage of BDOC released / cell

$b$  = percentage of BDOC released / cell

$K_{mort.}$  = bacterial mortality rate ( $s^{-1}$ )

$V_{max.}$  = maximum growth rate at temperature ( $s^{-1}$ )

$K_s$  = Monod half saturation coefficient ( $mg\ C \cdot m^{-3}$ )

$D_{biof.}$  = effective diffusivity of substrate in the biofilm ( $m^2 \cdot s^{-1}$ )

$r$  = yield coefficient of free bacteria

$r'$  = yield coefficient of fixed bacteria

$K_{det.}$  = detachment rate coefficient ( $s^{-1}$ )

$A$  = constant

$B$  = constant

$K_1$  = chlorine bacterial mortality of free bacteria ( $s^{-1}$ )

$K_2$  = chlorine bacterial mortality of fixed bacteria ( $s^{-1}$ )

$k_a$  = dissociation constant HOCl/CIO<sup>-</sup>

$c/2_{lim.}$  = chlorine limiting concentration for bactericidal action (pH = 7 and T = 20 °C)

$Z$  = biofilm mass density ( $mg\ C \cdot m^{-3}$ )

$V$  = fluid velocity ( $ms^{-1}$ )

$c$  = coefficient

$\alpha$  = hydraulic constant

$T_{opt.}$  = optimum temperature (°C)

$T_i$  = coefficient (°C)

$w$  = chlorine diffusion transfert constant (m)

$k_{hclo}$  = reactivity constant of hypochlorous acid ( $s^{-1}$ )

$k_{clo}^-$  = reactivity constant of hypochlorite ( $s^{-1}$ )

$K_w$  = first order kinetic constant for chlorine consumption due to water

$K_r$  = first order kinetic constant for chlorine consumption due to pipes

$C$  = carbon

*Functions:*

$S(y, t)$  = concentration of BDOC ( $\text{mg C}\cdot\text{m}^{-3}$ )

$S_b(x, y, t)$  = concentration of BDOC in non-chlorinated biofilm ( $\text{mg C}\cdot\text{m}^{-3}$ )

$S_b'(y, t)$  = average concentration of BDOC in non-chlorinated biofilm ( $\text{mg C}\cdot\text{m}^{-3}$ )

$S_{cl2}(x, y, t)$  = concentration of BDOC in chlorinated biofilm ( $\text{mg C}\cdot\text{m}^{-3}$ )

$S_{cl2}'(y, t)$  = average concentration of BDOC in chlorinated biofilm ( $\text{mg C}\cdot\text{m}^{-3}$ )

$X(y, t)$  = free bacteria concentration ( $\text{mg C}\cdot\text{m}^{-3}$ )

$X(y, t) \cdot H(y, t)$  = active free bacteria concentration ( $\text{mg C}\cdot\text{m}^{-3}$ )

$cl_2(y, t)$  = chlorine concentration ( $\text{mg}\cdot\text{l}^{-1}$ )

$h(y, t)$  = chlorine diffusion depth (m)

$B(y, t)$  = non-chlorinated biofilm thickness (m)

$G(y, t)$  = Mortality rate for fixed bacteria caused by chlorine ( $\text{s}^{-1}$ )

$F(y, t)$  = Mortality rate for free bacteria caused by chlorine ( $\text{s}^{-1}$ )

# THE EFFECT OF DUCTILITY IN COMBINATION WITH THE P- $\Delta$ EFFECT, TO BE TAKEN INTO CONSIDERATION IN THE SEISMIC DESIGN OF COMPRESSED STRUCTURES

Dulácska, E.

*Budapesti Műszaki Egyetem, Szilárdságtani és Tartószerkezeti Tanszék,  
Műegyetem rkp. 3, H-1521 Budapest, Hungary*

(Received: 10 December 1995)

Investigating the earthquake stability of buildings, the effect of ductility, supposed to be favourable, is considered throughout the world: the force  $F$  substituting the horizontal earthquake impulse acting on the elastic system is divided by the  $\mu = \Delta_u/\Delta_e$  ductility factor. This ductility factor varies between 2 and 12, e.g. in the US for reinforced concrete structures  $\mu = 8$ . As a result, buildings are analysed on a low load level in elastic state, where displacements are small. Thus the unfavourable effect of the vertical load is either neglected or considered only approximately (as a linear effect). In fact, plasticity and buckling are both nonlinear and they must not be superposed. According to correct analysis, this is a serious mistake at the expense of safety. Due to this usual method, throughout the world many earthquake-engineered buildings fall or collapse, burying thousands of people under themselves. For more detailed analysis, we conducted different investigations considering the moment-increasing effect of the vertical load, applying a factual elasto-plastic model with a ductility factor  $\mu = 12$ . Calculations were made by the energy method and also by solving the differential equation of motion with step-by-step method. Results were checked by experiments. It can be seen from the results that the impulsive force has a critical value, thereafter the displacement is going to infinity, whereas the velocity is not decreasing to zero, thus vibration will not stop. The plastic spectrum curve (dynamic factor) is bending to infinity, whereas the elastic spectrum is decreasing with the vibration period. Our investigations indicate that the earthquake design applied throughout the world is false. Certain formulas were derived for correct design. Design is proper if the favourable effect of ductility that sometimes occurs is considered as a reserve and not taken into account. This paper prepared under OTKA project in the RC. Reseach Group of MTA.

## 1. Introduction

The fundamental problem in seismic design is to determine the loads due to the horizontal mass forces resulting from earthquake. As long as the structure remains elastic, the usual dynamic methods can be used and the necessary calculations can be made, although quite lengthy calculations (usually by computer) are required.

Some decades ago it was found that considerable energy could be absorbed in certain cases by elastic deformation and this effect was introduced to the seismic design [1, 2, 3].



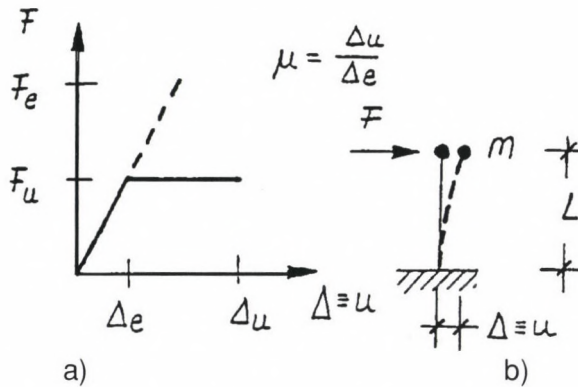


Fig. 1. Model of the usual seismic design method taking ductility into consideration

The plastic behaviour of the structural materials has been called ductility and the so-called ductility factor  $\mu$ , the quotient of limit plastic deformation  $\Delta_u$  and limit elastic deformation  $\Delta_e$ , that is

$$\mu = \Delta_u / \Delta_e \tag{1}$$

has been introduced.

The phenomenon discussed is illustrated in Fig. 1, where  $F_e$  is the replacement static force calculated on the bar of elastic material, which results in static deformation of a magnitude identical with that caused by the dynamic effect, in the case of an impulse excitation of constant value during the time  $t_0$ . We know that,  $F_e = \beta \cdot m \cdot a_g$  being the concentrated mass at the end of the bar,  $a_g$  the exciting acceleration which, multiplied by the mass, gives the force causing the impulse, while  $\beta$  is the dynamic factor. In the case of an elastic structure,  $\beta = 2$  if  $t_0/T > 1/2$ , and  $\beta = 2 \cdot \sin \pi \cdot t_0/T$  if  $t_0/T \leq 1/2$ ,  $T$  being the period of the structure. Furthermore,  $F_u$  is the limit plastic force,  $\Delta_e$  the limit elastic deformation of the ideally elastic-plastic bar and  $\Delta_u$  the limit plastic deformation. Note that also for a bar of continuous mass distribution, a replacement mass located at the end of the rod can be determined, which helps to simplify the dynamic calculations. The structure shown in Fig. 1 is the simplest model of one single mass and one degree of freedom, still suited, however, to describe the dynamic behaviour of the building.

The effect  $P - \Delta$  represents the fact that if a bar of elastic material is acted upon by a force  $P$ , then, due to the deformation  $\Delta$ , an incremental bending moment  $M = P \cdot \Delta$  also arises (see Fig. 2/a), causing further, additional deformation. If the force  $P$  approaches the Euler critical force  $P_{cr}$ , then  $\Delta$  tends towards infinity (Fig. 2/b). As the deformation increases and the stress in the bar reaches the ultimate stress, the bar fails.

Assume that the axial force in the bar is  $P = 0$  and that the horizontal static force  $F$  causes a static displacement  $\Delta_0$  of the bar end. As a rule, in seismic investigations, the force  $F$  is the product of the exciting acceleration  $a_g$ , replacing the dynamic effects, and the mass  $m$ . Note that, if instead of mass  $m$ , it is the clamping cross-section of the bar

which is excited by acceleration, then the force which excites the mass  $m$  as a result of acceleration  $a_g$  will be  $F = a_g \cdot m$  before displacement  $\Delta < \Delta_e$  is reached, while it will be  $F_U$  beyond this value of displacement, however, with a sign opposite to that valid in the case of a direct acceleration of mass  $m$ .

Should also a vertical axial force  $P$  act upon the bar, then deformation  $\Delta_0$  will increase to the value

$$\Delta = \frac{\Delta_0}{1 - P/P_{cr}} = \psi \cdot \Delta_0 \tag{2}$$

where for the structure investigated, that is, for a cantilever structure clamped at the bottom, the critical force  $P_{cr}$  is

$$P_{cr} = \pi EI / (2L)^2 \tag{3}$$

Here  $E$  is the modulus of elasticity and  $I$  is the moment of inertia of the cross-section of the bar. This increase of  $\Delta_0$  is called  $P - \Delta$  effect. (Displacement  $\Delta$  will be denoted by  $u$  as is customary in analytical investigations.)

If, for the sake of simplification again, a rigid bar clamped by a spring at the bottom is used to replace the bar in question, then the spring constant of the bar has to be chosen in such a way as to describe the behaviour of the elasto-plastic bar. In this simplified case, the displacement caused by the force  $F$  is  $\Delta_0 = F / k$ , and the critical force is  $P_{cr} = k \cdot L$ .

As shown by theoretical studies so far, the behaviour of the structure differs in some cases from what has been described in the literature [3, 12, 13]. The point in this behaviour is that not only the vertical force but also the horizontal force has a critical

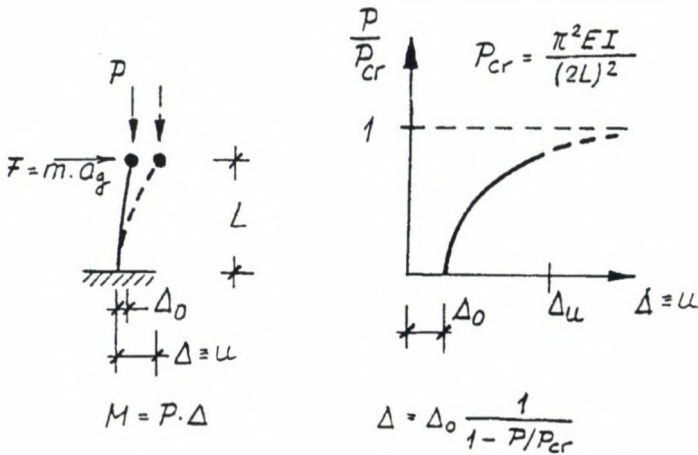


Fig. 2. Displacement of the model under compression

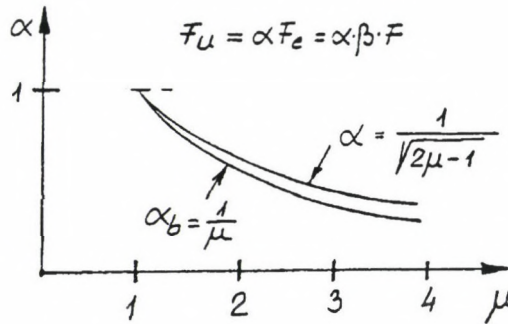


Fig. 3. Use of load reduction factor  $\alpha$  to take the effect of ductility factor  $\mu$  into consideration in the usual way

value. Either force above this critical value will certainly cause the collapse of the structure. This two-sided effect is due to the vertical force  $P$ , and this influence has not been properly taken into consideration in research so far. Since this phenomenon had not been mentioned in the literature available for us, we deemed it advisable to investigate this phenomenon by calculation and also experimentally, taking into consideration both the effect of force  $P$  and the plastic behaviour. This investigation is dealt with in the chapters to follow.

## 2. The usual ductile design procedure

The seismic analysis of buildings takes the expectedly favourable effect of ductility into consideration. In so doing, force  $F_e$  determined for the elastic structure is multiplied with a reducing factor  $\alpha$ . This reducing multiplier is specified in the regulations [4, 5] for the different materials. The reducing multiplier  $\alpha$  has been obtained by equalizing the work  $E_{el}$  (elastic work) performed by force  $F_e$  and the work  $E_{pl}$  (plastic work) performed by force  $F_u$ . In this way the reducing factor

$$\alpha = \frac{F_u}{F_e} = \frac{1}{\sqrt{2\mu - 1}} \quad (4)$$

has been obtained [1, 2, 3]. If the deformations are equalized, then the relationship

$$\alpha_b = \frac{F_u}{F_e} = \frac{1}{\mu} \quad (5)$$

will be arrived at. A comparison of the two factors is shown in Fig. 3. The US seismic design regulations use the factor  $\alpha_b$  with  $\mu = 8$  [4].



The effect  $P - \Delta$  has been either neglected while limiting the deformation, or it has been taken into consideration as a multiplier  $\psi$  of elastic buckling [4]. Accordingly, the seismic force is calculated by means of the formula

$$F = F_e \cdot \alpha \cdot \psi. \quad (6)$$

Note that different vibration patterns are possible in the case of multimass structures and this necessitates a combined calculation. However, from the point of view of our investigation, it is enough to study the single-mass system.

The usual calculation method described above is erroneous in that it super-imposes the results obtained for nonlinear systems, although a superposition can be used in the case of linear systems only. The erroneous calculation results in an approximation *to the detriment of safety*, in some cases quite considerably. Another problem of the usual method is that, instead of using the effect of ductility to increase the resistance of the structure, it reduces the load, and thus the calculation results in small deformations where the nonlinear effects are insignificant and thus the problem remains concealed in general.

### 3. The correct ductile design

As is well known, the critical state of an elasto-plastic structure can also be assessed by looking for the critical transverse force  $F_{cr}$  which causes the structure to go from the indifferent state to the instable state with continually increasing deformation (see e.g. in [6]).

This phenomenon had been demonstrated by computer simulation studies of elasto-plastic systems in compliance with the methods of the probability theory [7] and it has been proved experimentally by shaking table tests [8].

Here the point is that, if a system undergoes plastic deformation (it yields) in one direction as a result of the seismic shock, then the next yield would eventually take place as a result of a smaller seismic shock but in the direction of the original plastic deformation.

In this work, a single-mass system excited by a finitely small portion of the shock, by the impulse  $F \cdot t_0$ , will be investigated, taking into consideration the work  $P \cdot s$  of the force  $P$  performed will in the course of vertical displacement  $s$  caused by horizontal displacement  $\Delta$ . Here  $t_0$  is the finitely short time and  $s$  is the displacement of the bar end in the direction of force  $P$ .

#### 3.1. Investigation by means of the energy method

Let us consider the weightless, rigid cantilever bar shown in Fig. 4, stabilized by an elasto-plastic spring of the rigidity (spring constant)  $k$  at the lower end and having a mass  $m$  at the upper end.

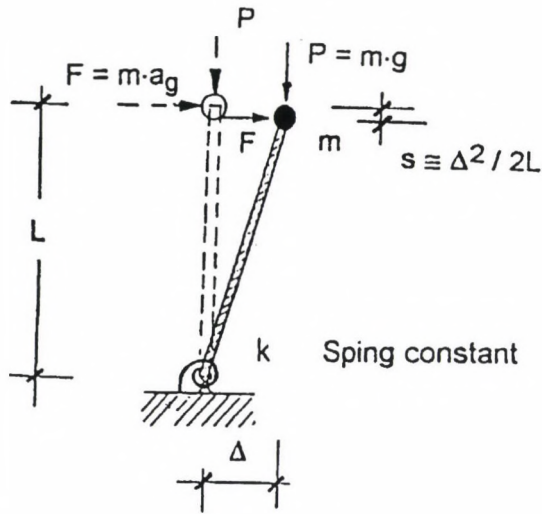


Fig. 4. Simplified model of the verification test

On the basis of the energy conservation law, the following equation system can be written [9]:

$$\Pi = \Pi_{ex} - \Pi_{in} = E_{kin} \quad (7)$$

Here  $\Pi$  is the total potential energy,  $\Pi_{ex}$  is the potential energy of external forces,  $\Pi_{in}$  is the potential energy of internal forces, while  $E_{kin}$  is the kinetic energy. Note that  $\Pi_{ex} = L_{ex1}$  (work of external forces) and  $\Pi_{in} = L_{in}$  (work of internal forces). The values of energies are

$$\Pi_{ex} = L_{ex1} + L_{ex2}, \quad (8)$$

where  $L_{ex1}$  is the work of horizontal force  $F = m \cdot a_g$ , that is,  $L_{ex1} = \int_0^{\Delta} a_g \cdot m \cdot d\Delta$ ,  $a_g$  being the exciting acceleration,  $m$  the mass, and  $\Delta$  the path length. If  $a_g$  and  $m$  are constant, then

$$L_{ex1} = a_g \cdot m \cdot \Delta. \quad (9/a)$$

If  $t_0$  is the time for which the constant force  $F = m \cdot a_g$  is acting, then the impulse will be

$$I_t = a_g \cdot m \cdot t_0. \quad (10)$$

In the case of a short impulse time  $t_0$  ( $t_0 \leq T/5$ ), the velocity will be  $v = I_t/m$  (with the small work performed by the spring force during this time neglected) and thus, the work of the force causing the impulse can be expressed as [10, 11], where  $T$  is the period.

$$L_{ex1} \approx a_g \cdot m \cdot t_o \cdot \left( \frac{1}{2} a_g t_o^2 \right) = \frac{I_t^2}{2 \cdot m} \quad (9/b)$$

The second part  $L_{ex2}$  of the external work can be described as the work performed by force  $P$  on the subsidence (vertical) path  $s$ , that is,

$$L_{ex2} = P \cdot s, \quad (11)$$

where  $P = m \cdot g$  is the gravitational force with  $g$  as the gravity acceleration.

The mass  $m$  moves along a circular arc and the value  $s$  can be expressed from the equation  $\Delta^2 + (L - s)^2 = L^2$  of the circle. If we confine us to values  $\Delta \leq L/2$ , then

$$s \approx \Delta^2 / 2L. \quad (12)$$

Taking this into consideration, the external work will be

$$L_{ex} = L_{ex1} + L_{ex2} = \frac{I_t^2}{2m} + P \frac{\Delta^2}{2L}, \quad (13)$$

In the case of a short-time impulse of constant value  $I_t = a_g \cdot m \cdot t_0$ , while

$$L_{ex} = a_g \cdot m \cdot \Delta + P \frac{\Delta^2}{2L}. \quad (14)$$

in the case of a shock of constant intensity and longer time.

Expression

$$\Pi_{in} = L_{in,el} + L_{in,pl} + E_{damp} \quad (15)$$

can be used to describe the internal potential energy which the structure exerts against the movement. Here

$$L_{in,el} = \frac{1}{2} F_{in} \cdot \Delta, \quad (16)$$

is the internal elastic energy stored over the elastic section  $\Delta \leq \Delta_e$ . Force  $F_{in}$  can be expressed by the spring constant  $k$ , thus  $F_{in} = k \cdot \Delta$ .

The internal plastic energy stored over the plastic section  $\Delta \geq \Delta_e$  is

$$L_{in,pl} = F_u \cdot \left( \Delta - \frac{\Delta_e}{2} \right), \quad (17)$$



where  $F_u = k \cdot \Delta_e$  is the plastic limit force. In the course of the movement of the structure, the internal plastic energy  $L_{in,pl}$  is converted into heat and thus it does not result either in elastic energy accumulation or in a restoring force.

$E_{damp}$  is the energy absorbed by the internal damping of the system which is neglected in this investigation because of its negligibly small value.

Thus, over the elasto-plastic section, the internal potential energy is

$$\Pi_m = L_m = L_{in1} + L_{in2} = \frac{1}{2} F \cdot \Delta_e + F_u \left( \Delta - \frac{\Delta_e}{2} \right), \quad (18)$$

while the kinetic energy is

$$E_{kin} = \frac{1}{2} m \cdot v^2, \quad (19)$$

where  $v$  is the first derivative of  $\Delta$  with respect of time.

The investigation of the state of movement of the structure leads to the following conclusions: Taken the condition of equilibrium into consideration, the structure will move as long as

$$\Pi > 0, \quad (20)$$

because there is kinetic energy present in the system. If

$$\Pi = 0, \quad (21)$$

the structure will stop moving and the state of equilibrium will set in. We have now to determine whether the state of equilibrium is stable or indifferent. If  $d\Pi_{ex} = d\Pi_{in}$ , that is, if in the course of a small displacement  $d\Delta$ , the external work is equal to the internal energy stored, the system will be in an indifferent state. This state is called critical state or critical position, and the displacement  $\Delta_{cr}$  associated with this state is called critical displacement. If  $\Pi = 0$  exists and  $d\Pi = d\Pi_{ex} - d\Pi_{in} \leq 0$ , then the system investigated is in equilibrium and it will be stable because after a small displacement it assumes its original position again. However, if  $\Pi > 0$  or if in case of  $\Pi = 0$ ,  $d\Pi = d\Pi_{ex} - d\Pi_{in} > 0$ , the system is instable and it will keep on moving. Hence, to find the critical position, conditions  $\Pi = 0$  and  $d\Pi = 0$  have to be met. To investigate values of  $d\Pi$ , the derivative can be used in general.

What has been said above is illustrated as a function of displacement in Fig. 5. The external work  $L_{ex}$  can be seen in Fig. 5/a. If  $P = 0$ , then  $L_{ex}$  remains constant after displacement  $\Delta(t_0)$ . However, if there exists a vertical force  $P$ , then  $L_{ex}$  will also increase as  $\Delta$  increases. The internal energy  $L_{in}$  can be seen in Fig. 5/b. If the value of the spring constant  $k$  is high, then the internal energy  $L_{in}$  will also be larger, while if a value of  $k$  is low, then the internal energy  $L_{in}$  will be smaller for the same value of  $\Delta$ . Figure 5/c shows the difference between  $L_{ex}$  and  $L_{in}$  which is the kinetic energy  $E_{kin}$ . The movement will not stop before this value becomes zero.

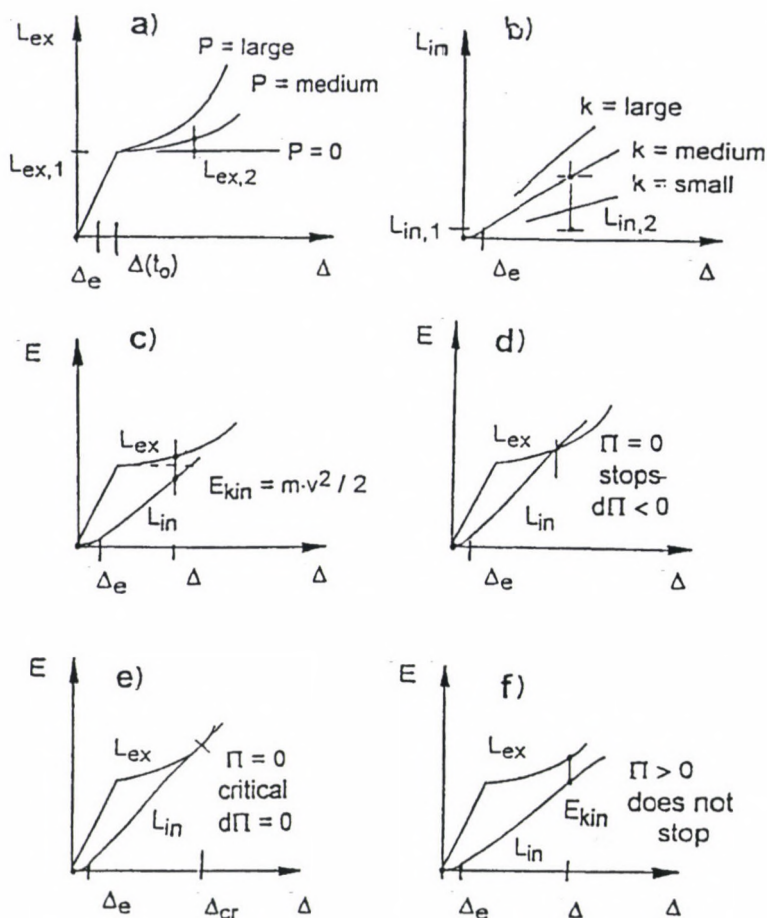


Fig. 5. Investigation by means of the energy method

In Fig. 5/d, a larger spring constant is associated with a smaller force  $P$ . Thus the two curves intersect and the structure stops moving. This case will always take place in the case of  $P=0$  whenever the value of  $t_0$  is low. In the case of a larger force  $P$  or a reduced rigidity  $k$  of the spring, the case shown in Fig. 5/e occurs where, as seen, the two lines meet at the critical displacement  $\Delta_{cr}$ . If  $P$  increases further or  $k$  decreases further, then the two values will not intersect, so that some part of the kinetic energy  $E_{kin}$  still exists, that is,  $\Pi > 0$ . Accordingly, the structure keeps on moving and thus it gets destroyed. Note that if  $P \geq P_{cr}$ , the structure will buckle even without any horizontal force. Here  $P_{cr} = k \cdot L$ , the Euler force.

Investigated below are the effects of a short-time impulse and those of a longer shock lasting at least until  $\Delta_{cr}$  is reached, separately for the elastic and the elasto-plastic case.

Since we are interested in the state when the structure stops moving, the velocity is  $v = 0$  and the kinetic energy need not be dealt with. Thus the energy equation is  $L_{ex} = L_{in}$ .

Let first the cases  $P = 0$  investigated ( $T$  being the period).

**E, 0, s** (elastic bar,  $P = 0$ , short-time impulse).

Assume that  $t_0 \leq T/5$ .

The works are:

$$L_{ex} = \frac{I_t^2}{2m} = \frac{1}{2} m \cdot a_g^2 \cdot t_o^2,$$

$$L_{in} = \frac{1}{2} F \cdot \Delta = \frac{1}{2} k \cdot \Delta^2.$$

With the two values made equal and after simplification we obtain

$$\Delta_{\max} = \frac{I_t}{\sqrt{m \cdot k}} = a_g \cdot t_o \sqrt{\frac{m}{k}}. \quad (22)$$

**E, 0, ℓ** (elastic bar,  $P = 0$ , longer shock).

The works are:

$$L_{ex} = a_g \cdot m \cdot \Delta,$$

$$L_{in} = \frac{1}{2} k \cdot \Delta^2.$$

From the two values,

$$\Delta_{\max} = \frac{2m \cdot a_g}{k}. \quad (23)$$

**PI, 0, s** (plastic bar,  $P = 0$ , short-time impulse). Assume that  $t_0 \leq T/5$ , where  $T$  is the period of the free vibration and  $F_u = k \cdot \Delta_e$  (the plastic force).

The works are:

$$L_{ex} = \frac{I_t^2}{2m} = \frac{1}{2} m \cdot a_g^2 \cdot t_o^2,$$

$$L_{in} = k \cdot \Delta_e \left[ \Delta - \frac{\Delta_e}{2} \right].$$

With the values of  $L_{ex}$  and  $L_{in}$  made equal and after simplification we obtain

$$\Delta_{\max} = \frac{1}{2} \left[ \Delta_e + \frac{m \cdot a_g^2 \cdot t_o^2}{F_u} \right]. \quad (24)$$



**PI, 0,  $\ell$**  (plastic bar,  $P = 0$ , longer shock)

$$L_{ex} = a_g \cdot m \cdot \Delta,$$

The works are:

$$L_m = k \cdot \Delta_e \left[ \Delta - \frac{\Delta_e}{2} \right].$$

From the two values:

$$\Delta_{\max} = \frac{\Delta_e}{2} \left[ \frac{1}{1 - \frac{a_g \cdot m}{F_u}} \right]. \quad (25)$$

If  $a_g \cdot m < F_u/2$ , we will have an elastic case, while if  $a_g \cdot m > F_u/2$ , we will have a plastic case.

If  $a_g \cdot m > F_u$ , the structure will not stop moving because  $\Pi > 0$  and thus  $\Delta \rightarrow \infty$ .

*Let now the cases  $P \neq 0$  be investigated.*

**E, P, s** (elastic bar, force  $P$ , short-time impulse)

The works are:

$$L_{ex} = \frac{1}{2} m \cdot a_g^2 \cdot t_o^2 + P \frac{\Delta^2}{2L},$$

$$L_m = \frac{1}{2} k \cdot \Delta^2.$$

With the two values made equal

$$\Delta_{\max} = \sqrt{\frac{L \cdot m \cdot a_g^2 \cdot t_o^2}{k \cdot L - P}}. \quad (26)$$

is obtained.

If the denominator becomes zero, then  $\Delta \rightarrow \infty$ . This case will occur if  $P = k \cdot L$ . Therefore,  $d\Pi = 0$  in this case and thus  $P = P_{cr}$ , the critical force.

**E, P,  $\ell$**  (elastic bar, force  $P$ , long-time shock)

The works are:

$$L_{ex} = a_g \cdot m \cdot \Delta + \frac{P \cdot \Delta^2}{2L},$$

$$L_m = \frac{1}{2} k \cdot \Delta^2.$$

From the two values,

$$\Delta_{\max} = \frac{2a_g \cdot m}{k - \frac{P}{L}} = \frac{2a_g \cdot m}{k \left(1 - \frac{P}{P_{cr}}\right)}. \quad (27)$$

As can be seen, relationship (23) will be obtained if  $P = 0$ , while in the critical case  $P = P_{cr}$ ,  $\Delta \rightarrow \infty$  if  $a_g \cdot m \neq 0$ . The dynamic factor is  $\beta = 2$ .

**PI, P, s** (plastic bar, force  $P$ , short-time impulse)

$$L_{ex} = \frac{1}{2} m \cdot a_g^2 \cdot t_o^2 + \frac{P \cdot \Delta^2}{2L},$$

The works are:

$$L_{in} = F_u \left( \Delta - \frac{\Delta_e}{2} \right).$$

Setting equal the two values we obtain a second degree equation which yields

$$\Delta_{\max} \frac{F_u \cdot L}{P} \left[ 1 - \sqrt{1 - \frac{P}{F_u \cdot L} \left( \frac{m \cdot a_g^2 \cdot t_o^2}{F_u} + \Delta_e \right)} \right]. \quad (28)$$

This remains valid as long as  $m \cdot a_g^2 \cdot t_o^2 < F_u \cdot \Delta_e \left( \frac{P}{P_{cr}} - 1 \right)$ , otherwise  $\Delta \rightarrow \infty$ . This may occur if  $F_u \leq P \cdot \Delta_e / L$  or  $P/P_{cr} \geq 1,0$ .

**PI, P,  $\ell$**  (plastic bar force  $P$ , long-time shock)

$$L_{ex} = a_g \cdot m \cdot \Delta + P \cdot \frac{\Delta^2}{2L},$$

The works are:

$$L_{in} = F_u \left( \Delta - \frac{\Delta_e}{2} \right).$$

Setting equal the two equations, and solving the second degree equation,

$$\Delta_{\max} = B \cdot \left[ 1 - \sqrt{1 - \frac{F_u \cdot L \cdot \Delta_e}{P \cdot B^2}} \right], \quad (29)$$

is obtained, where

$$B = (F_u - a_g \cdot m) \cdot \frac{L}{P}. \quad (30)$$

As long as  $a_g \cdot m \leq \frac{\Delta_e(k - P/L)}{2}$ , the structure deforms elastically only. However, if the value of  $a_g \cdot m$  is higher, plastic deformation will take place as well, and if

$$a_g \cdot m \geq F_u \left[ 1 - \sqrt{\frac{P \cdot \Delta_e}{L \cdot F_u}} \right], \quad (31)$$

the structure will not stop moving,  $\Delta > B$  and thus  $\Delta \rightarrow \infty$ . In the case of  $P \geq P_{cr} = k \cdot L$ , the structure buckles as well, because  $\Delta \rightarrow \infty$  also in this case.

On the basis of the above analyses, the behaviour of the structure depends obviously not on whether the material of the structure can endure a relatively large plastic deformation or not, but it depends on the rigidities of the structure, on the magnitude and duration of the shock, and furthermore, on the value of quotient  $P/P_{cr}$ , that is, on the magnitude of the safety against buckling. The plastic deformation capacity of the material sets a limit only in the sense that if  $\Delta_u < \Delta_{cr}$ , then the structure will collapse earlier, but the condition  $\Delta_u > \Delta_{cr}$  will not increase the destructive effect at all.

With the energy equation written in the case of  $PI, P, \ell$  we obtain

$$a_g \cdot m \cdot \Delta + \frac{P \cdot \Delta^2}{2L} = F_u \left( \Delta - \frac{\Delta_e}{2} \right). \quad (32)$$

Assuming an arbitrary value of  $\Delta < \Delta_u$  we can determine from the equation the pertaining value of  $F_u$  which stops the structure moving:

$$F_u = \frac{a_g \cdot m \cdot \Delta + \frac{P \cdot \Delta^2}{2L}}{\Delta - \frac{\Delta_e}{2}}. \quad (33)$$

The structure is properly designed if the value of  $F_u$  becomes a minimum. This minimum value of  $F_u$  can be calculated by means of extremum calculation or it can be chosen from among values of  $F_u$  calculated for different values of  $\Delta < \Delta_u$ . This solution complies with the solutions  $\Pi = 0$  and  $d\Pi = 0$ .

### 3.2. Investigation by means of equilibrium differential equation

When writing the equilibrium-deformation differential equation, displacement will be denoted by  $u$  as is usual in analytical calculations. Accordingly,  $\Delta \equiv u$ . In the equation, the term expressing the effect of force  $P$  is  $P \cdot u / (L - s)$ . Assuming a parabolic movement  $s = u^2/2L$  for  $s$  instead of the accurate movement along a circular arc, then the error will be less than 1% provided that  $u \leq 0.5 L$ . Moreover, because of its small value, also the internal damping arising in the elastic domain can be neglected. Thus in the elastic domain, the equilibrium differential equation becomes



$$m \cdot a_g - m \cdot \ddot{u} - k \cdot u + \frac{P}{L-s} u = 0. \quad (34/a)$$

In the plastic domain, the difference lies in that after reaching the elastic limit  $u_e$ , the force in the bar stops increasing, and the value  $k \cdot u_e = F_u$  enters the equation instead of the term  $k \cdot u$ . Thus the differential equation of the plastic domain will be

$$m \cdot a_g - m \cdot \ddot{u} - F_u + \frac{P}{L-s} u = 0. \quad (34/b)$$

Of course, the two equations must join each other at  $u = u_e$ . Should the  $s$  be neglected as compared with  $L$ , then the error resulting from this will increase quadratically with  $u$  and it will reach 10% in the term expressing the effect of the force if  $u \leq 0.5 L$ . Taking this into consideration, the  $s$  as compared with  $L$  will be neglected.

The analytical solution of equations (34) is unknown. Different numerical "step-by-step" methods have therefore been developed to solve equations of this type, among them e.g. the average acceleration method of NEWMARK, which uses an acceleration of constant value over the differentially small section investigated [12], or the method of linear acceleration where the change in acceleration is considered to be linear over the differentially small section in question [13]. These methods have been developed for manual calculation using 0.1 to 0.05 sec as a step of time. However, a computer is capable of handling much smaller steps at a fast rate and thus the method can be re-garded sufficiently accurate. Therefore, in our example, the solution has been produced by rewriting the differential equation so as to obtain a difference equation which has then been solved using the step-by-step method. Thus, essentially, the solution is similar to the NEWMARK method. Care should be taken to use a step of time not verging on the sensitivity limit of the computer because in this case the accuracy might be affected unfavourably. Use of a step of 1/500 to 1/1000 sec results in a sufficient accuracy, as a rule. Considering that the calculation of  $P/(L-s)$  with value  $u$  in  $s$  in the denominator would have been measurements, we used an  $L-s \approx L$  approximation, taking into consideration what has been said earlier. This approximation would result in an error of max. 5-6 % within the limits of the experimental measurements. Thus the following difference equations have been obtained:

Difference equation of the elastic domain:

$$m a_{g(t)} - m \frac{u_{i-1} - 2u_i - u_{i+1}}{\Delta t^2} - k_i \cdot u_i + \frac{P}{L} u_i = 0. \quad (35/a)$$

Difference equation of the plastic domain:

$$m a_{g(t)} - m \frac{u_{i-1} - 2u_i - u_{i+1}}{\Delta t^2} - F_u + \frac{P}{L} u_i = 0. \quad (35/b)$$

Problems would arise in the solution if every value were zero at time  $t = 0$  because, in this case, the computation would not start and some trick is necessary for starting. Such a problem has not been encountered in our case because the acceleration  $a_g$  has been constant.

With difference equations (35), the computation can be started because  $u_{i-1} = 0$ ,  $u_i = 0$ ; and  $u_{i-1}$  can be calculated. In the next step, that is at time  $\Delta t$ , only  $u_{i-1} = 0$  can be calculated, while from the third step on, neither of the values is computable. This introduces an error in the computation at the beginning, but this error is very small because the values of  $\Delta t$  are very small as well.

The computation will change if instead of the mass it is the support which is excited. Over the elastic domain, the effect of the support acceleration is the same as if the mass were accelerated at the same rate in the opposite direction [14]. In the plastic domain, limits are set to excitation by the value of  $F_u$  in the case of  $a_g \cdot m \cdot F_u$ .

## 4. Experiments

The theoretical investigations discussed so far have shown that, in certain cases, the behaviour of the structure differs from what has been described in the literature [3, 12, 13]. Here the point is that not only the vertical force but also the horizontal force has a critical value, causing the structural to collapse. This "two-sided" effect results from vertical force  $P$  which has not been correctly taken into consideration so far. Considering that this phenomenon had not been mentioned in the literature known to us, we found it necessary to investigate the phenomenon experimentally to check whether our calculations were correct. We thought sufficient in the first step to include in the experiment the model investigated theoretically.

### 4.1. Planning of the experiments

Originally, we wanted to use the shaking table of the Department of Strength of Materials and Load-bearing Structures of the Technical University Budapest. However, in the course of detailed planning, the actuator of the shaking table proved to be incapable of producing the required impulses. Therefore, a shock transfer device was developed, where the impulse resulted from the energy of a mass falling from a definite height  $h$  and converted into a horizontal shock by means of a rocker system.

This device is shown in Fig. 6. The cantilever structure investigated was built onto the fixed shaking table. The frame structure had also a leg hinged above and below parallel to the bar clamped at the bottom, in order to provide room for mass  $m_1$  to be placed on the beam of the frame and to eliminate the rotational inertia of the mass. Steel disks fastened by bolts on the beam were used to build up the mass. The measuring device was mounted on an independent stand (not shown in Fig. 6).

The mass  $m_2$  giving the impulse was fastened on the left arm of the cross-shaped rocker fastened to the base on the right side (see Fig. 7). This mass consisted of disc-



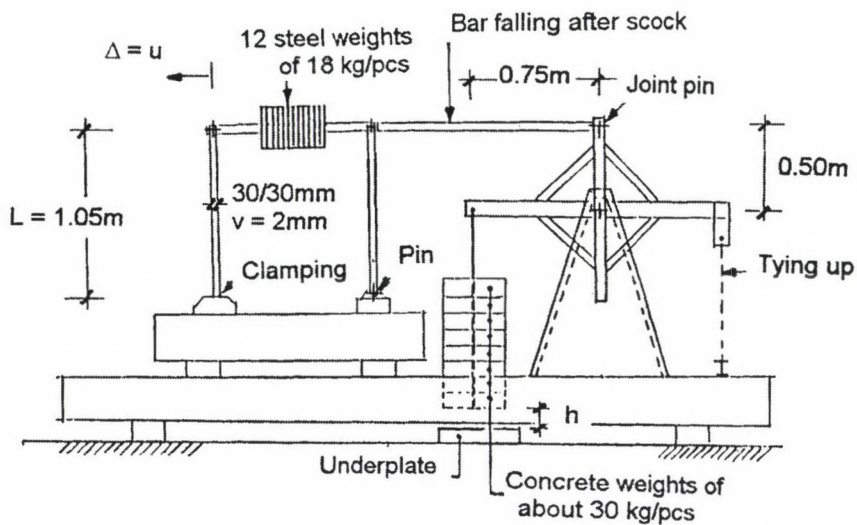


Fig. 6. Schematic illustration of the experimental device

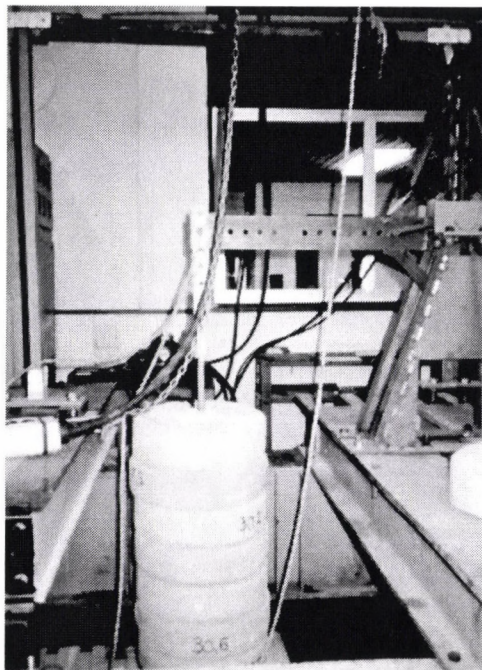


Fig. 7. Concrete weights resulting in mass  $m_2$

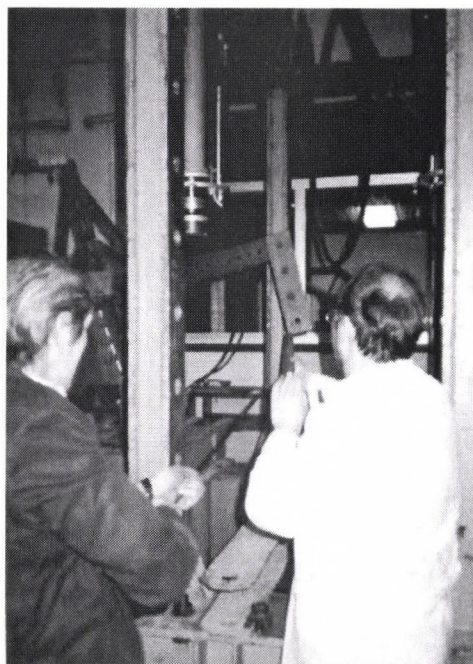


Fig. 8. Tying up of the aim of the rocker



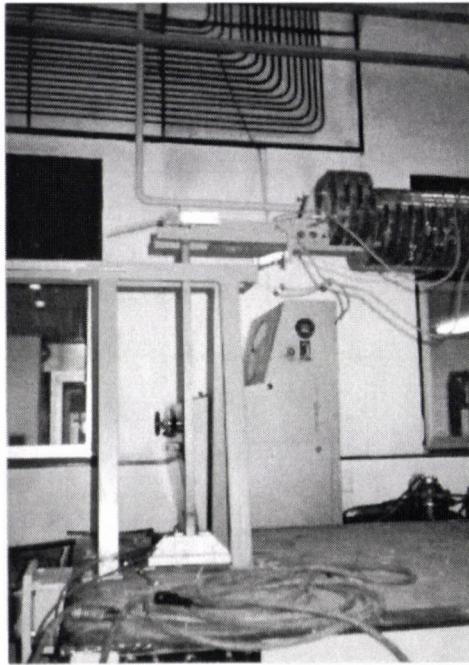


Fig. 9. The experiment as adjusted together with the steel weight resulting in the bulk of mass  $m_1$

shaped concrete weights weighing 30 kg each. The impulse strength was increased by increasing the number of the concrete discs, while the duration of the impulse ( $t_0$ ) was regulated by changing the height of fall ( $h$ ). Five and 9 pieces of concrete weight were used in the experiments and the height of fall was 0.10 m and 0.15 m, respectively. The mass of the bar and rocker arms, properly transformed (into reduced mass), was included in the mass ( $m_1$ ) on the horizontal bar and thus a mass of  $m_1 = 235$  kg resulted. The right arm of the rocker was tied up (Fig. 8), then untied for impulse production. The interconnecting bar was designed so as to disengage after the shock to permit the system investigated to move freely. The experiment set-up is shown in Fig. 9.

In the experiments, the height of fall and the acceleration were measured while the velocity was determined by differentiation of the height and integration of the acceleration (properly coordinating the two calculations).

#### 4.2. Calculation of acceleration transferred to the test bar

The rocker system used is similar to the ATWOOD free-fall apparatus but it is more sophisticated as far as kinetics is concerned. The model shown in Fig. 10 was used to calculate the acceleration transferred to mass  $m_1$  on the test bar.

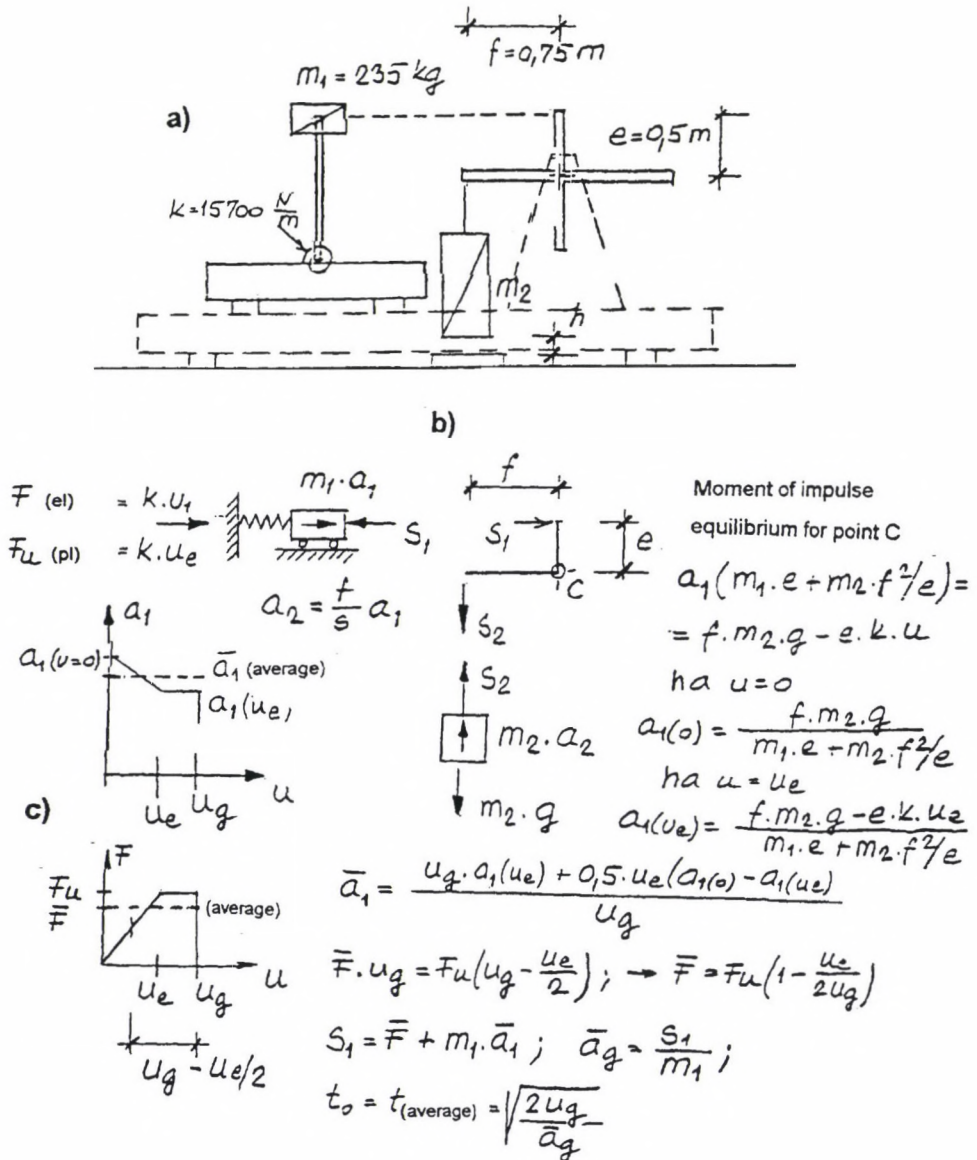


Fig. 10.

- a) Schematic model of the experimental setup
- b) Calculation model of the test impulse
- c) Calculation of the values and averages of the test impulse and internal force, and determination of the (average) accelerating excitation  $a_g$  for the time  $t_0$

The geometry of the model being shown in Fig. 10/a, while the forces acting upon the masses as well as the dynamic model are given in Fig. 10/b. The effect of the rocker was determined by means of the impulse moment equilibrium equation written for the centre of rotation C. As seen in Fig. 10/c, the acceleration of mass  $m_1$  decreases first over the elastic domain  $u < u_e$  and becomes constant over the plastic domain  $u > u_e$ . The average value of acceleration can be calculated by means of the expressions presented. The force  $F$  acting upon the bar increases first over the elastic domain until  $F_u$  is reached. Thereafter its value remains constant, the average being  $\bar{F}$ . By setting equal the work of the average force  $\bar{F}$  to that of the trapezoidal diagram, the value of  $\bar{F}$  can be calculated. The value of the accelerating force  $S_1$  can be obtained as the sum of  $\bar{F}$  and of d'Alembert's force produced by multiplying the acceleration  $a$  with the mass  $m_1$ , and thus the average exciting acceleration  $a_g$  and excitation time  $t_0$  can be calculated. For details of the calculation see Fig. 10/c.

A cold rolled 30 × 30 mm square tube of a wall thickness of 2 mm was used for the experiment. Over the section where plastic deformation was to be expected, the steel tube was heated by flame to become soft and thus comparable with the ideally elasto-plastic material. The force-displacement diagram of the bar obtained experimentally shown in Fig. 11. This is, of course, not ideally elasto-plastic because of the gradual plasticization of the cross-section. At a certain value of displacement, the wall of the tube buckled and, as a result, the rigidity decreased and thus the curved section of the

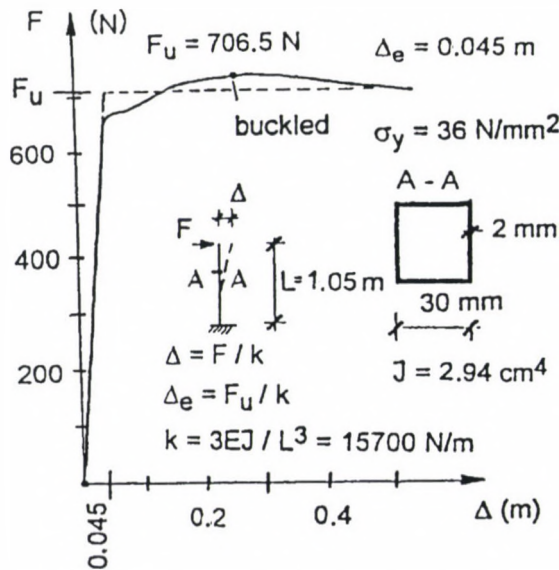
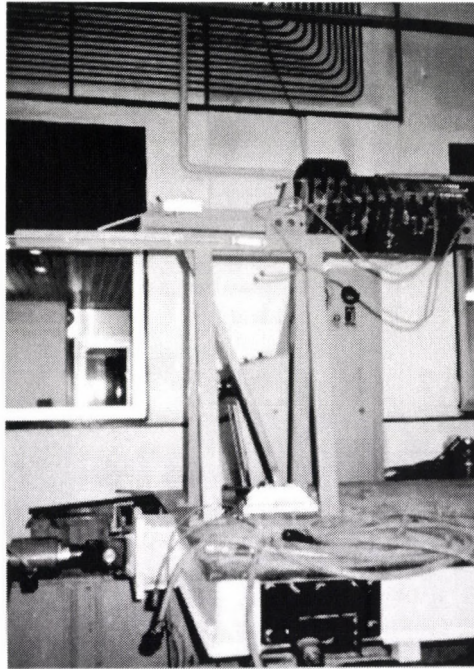


Fig. 11. Force-displacement diagram of the specimen, obtained in static experiment





*Fig. 12. Specimen permanently curved after the experiment*

diagram started to bend downwards. In the calculations, the idealized elasto-plastic bi-linear line (see the broken line in Fig. 11) has been taken into consideration with a plastic force of  $F_u = 706.5$  N. Thus  $u_c = 0.045$  m has been obtained for the limit of plastic deformation.

A limit deformation of  $u_T = 0.54$  m has been assumed, resulting in a ductility factor of  $\mu = 0.54/0.045 = 12$ . The yield point of the softened material of the bar has been determined by a tensile test which resulted in  $\sigma_y = 36$  N/mm<sup>2</sup>.

#### *4.3. The experiments carried out*

Originally, 15 experiments have been planned so that the critical phenomenon could be included in a range of parameters adjusted. Unfortunately, 4 experiments proved to be unsuccessful due to different problems (collision of parts, loss of data by the controlling computer). Thus 11 tests results could be evaluated. The increment of the test impulse magnitude were determined by the size of the concrete weights available. To increase the fineness of the stepping, mass  $m_2$  was dropped from two different heights  $h$ , 0,10 and 0.15 m.

The data of the experiments which could be evaluated are compiled in Table 1, showing masses  $m_2$  used, height  $h$ , excitation path  $u_g$  of mass  $m_1$  when dropped from

Table 1  
Initial accelerations  $a_g$  and excitation times  $t_0$

Sign	h [m]	$u_g$ [m]	$m_2$ [kg]	$a_g$ [m/sec <sup>2</sup> ]	$t_0$ [sec]
D <sub>1</sub>	0.10	0.0667	184	5.443	0.1565
D <sub>2</sub>	0.10	0.0667	215	5.743	0.1524
D <sub>3</sub>	0.10	0.0667	249	6.011	0.1489
D <sub>4</sub>	0.15	0.10	249	6.249	0.1789
D <sub>5</sub>	0.15	0.10	249	6.249	0.1789
D <sub>6</sub>	0.15	0.10	215	5.970	0.1830
D <sub>7</sub>	0.15	0.10	215	5.970	0.1830
D <sub>8</sub>	0.15	0.10	215	5.970	0.1830
D <sub>9</sub>	0.15	0.10	215	5.970	0.1830
D <sub>11</sub>	0.10	0.0667	280	6.214	0.1465
D <sub>13</sub>	0.15	0.10	152	5.257	0.1950

$$m_1 = 235 \text{ kg}, k = 15,700 \text{ N/m}, u_e = 0.045 \text{ m}, f/e = 1.5$$

height  $h$ , exciting acceleration  $a_g$  computed according to Sect. 4.2 as well as excitation time  $t_0$  associated with it.

The experimental results were recorded by the measuring device at 0.02 sec intervals. For the check calculations according to Sect. 3.2, a time step of  $dt = 0.005$  has been assumed. The values of displacement calculated ( $u_{\text{cal}}$ ) and obtained experimentally ( $u_{\text{exp}}$ ) are shown at intervals of 0.2 sec in Table 2.

Data calculated for, and obtained experimentally in, one experiment before, and another experiment after, the critical state (*D I* and *D II*, respectively) are illustrated in Fig. 14 and 15. Indicated in the Figures are the impulsive exciting acceleration as well as the chronological history of acceleration  $a$ , velocity  $v$  and displacement  $u$ , of the mass  $m_1$ . Thick lines were used for the experimental results while thin broken lines for the calculated values.



Fig. 13. The basis of the specimen, permanently yielded and buckled after the experiment

Table 2  
Displacements  $u$  in experiments D<sub>1</sub>–D<sub>13</sub>

Sign D	Data: sec m/s <sup>2</sup> kg m	t → sec	0	0.2	0.4	0.6	0.8	1.0	1.2	1.4	max.
1	$t_0 = 0.1565$ $a_g = 5.443$ $Q = 184$ $h = 0.10$	$u_c$	0	0.090	0.177	0.209	0.201	0.178	0.184	0.207	0.21
		$u_{exp}$	0	0.093	0.186	0.219	0.214	0.197	0.198	0.214	0.22
2	$t_0 = 0.1524$ $a_g = 5.743$ $Q = 215$ $h = 0.10$	$u_c$	0	0.094	0.186	0.226	0.231	0.212	0.206	0.224	0.23
		$u_{exp}$	0	0.072	0.175	0.207	0.188	0.157	0.186	0.204	0.21
3	$t_0 = 0.1489$ $a_g = 6.011$ $Q = 249$ $h = 0.10$	$u_c$	0	0.098	0.192	0.237	0.252	0.243	0.231	0.238	0.25
		$u_{exp}$	0	0.076	0.179	0.211	0.199	0.171	0.161	0.193	0.21
4	$t_0 = 0.1789$ $a_g = 6.249$ $Q = 249$ $h = 0.15$	$u_c$	0	0.108	0.251	0.370	0.509	–	–	–	$\infty$
		$u_{exp}$	0	0.082	0.231	0.218	0.193	0.253	0.217	0.213	0.26
5	$t_0 = 0.1830$ $a_g = 6.249$ $Q = 249$ $h = 0.15$	$u_c$	0	0.108	0.251	0.370	0.509	–	–	–	$\infty$
		$u_{exp}$	0	0.101	0.248	0.340	–	–	–	–	–
6	$t_0 = 0.1830$ $a_g = 5.970$ $Q = 215$ $h = 0.15$	$u_c$	0	0.103	0.242	0.351	0.474	–	–	–	$\infty$
		$u_{exp}$	0	0.106	0.246	0.343	–	–	–	–	$\infty$
7	$t_0 = 0.1830$ $a_g = 5.970$ $Q = 215$ $h = 0.15$	$u_c$	0	0.103	0.242	0.351	0.474	–	–	–	$\infty$
		$u_{exp}$	0	0.106	0.246	0.343	–	–	–	–	$\infty$
8	$t_0 = 0.1830$ $a_g = 5.970$ $Q = 215$ $h = 0.15$	$u_c$	0	0.103	0.242	0.351	0.474	–	–	–	$\infty$
		$u_{exp}$	0	0.101	0.256	0.375	–	–	–	–	$\infty$
9	$t_0 = 0.1830$ $a_g = 5.970$ $Q = 215$ $h = 0.15$	$u_c$	0	0.103	0.242	0.351	0.474	–	–	–	$\infty$
		$u_{exp}$	0	0.085	0.243	0.364	–	–	–	–	$\infty$
11	$t_0 = 0.1465$ $a_g = 6.214$ $Q = 280$ $h = 0.10$	$u_c$	0	0.101	0.202	0.259	0.293	0.318	0.344	0.380	$\infty$
		$u_{exp}$	0	0.099	0.203	0.262	0.297	0.324	0.353	–	$\infty$
13	$t_0 = 0.1950$ $a_g = 5.257$ $Q = 152$ $h = 0.15$	$u_c$	0	0.090	0.216	0.303	0.385	0.495	–	–	$\infty$
		$u_{exp}$	0	0.072	0.211	0.312	–	–	–	–	$\infty$

$P = 2350$  N.  $P_{cr} = 16,480$  N.  $m = 235$  kg.  $L = 1.05$  m.  $u_e = 0.045$  m.  $u_T = 0.54$  m  
 $k = 15.700$  N/m.  $P/P_{cr} = 0.1426$ ; c = computed. exp = experimental



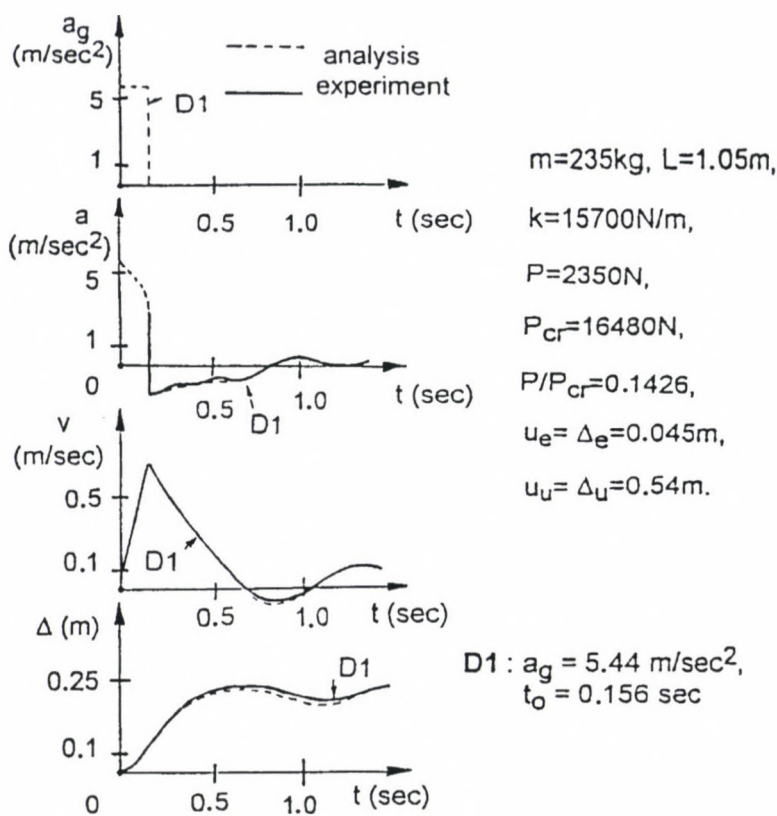


Fig. 14. Acceleration, velocity and path (displacement) diagrams of mass  $m_1$  of specimen D1 as functions of time excited by an impulse smaller than the critical impulse (experimental results: thick line, calculated results: thin broken line)

It can be seen in the Figures that, on the one hand, there is a satisfactory agreement between the calculation and the experiment and, on the other hand, that the critical phenomenon does exist. In experiment D1, the velocity decreased to zero at about 0.7 sec as a result of an impulse  $I_t = 200$  mkg/sec and the displacement stopped. Hence, the phenomenon is stable. However, in experiment D11, the velocity due to the impulse  $I_t = 214$  did not decrease to zero and the displacement increased beyond any limit boundlessly. That means that the phenomenon is unstable.

The critical state lies between the two values. According to relationship (28) and to condition  $m \cdot a_g^2 \cdot t_o^2 = F_u \cdot \Delta_e (P_{cr}/P - 1)$ , the critical impulse is  $I_t = 212$  mkg/sec. These values show that the theoretical calculation is reliable.

For the sake of interest, the values of displacement have been calculated for an excitation time of  $t_0 = 0.15$  and for different values of  $a_g$ , and illustrated in a diagram in

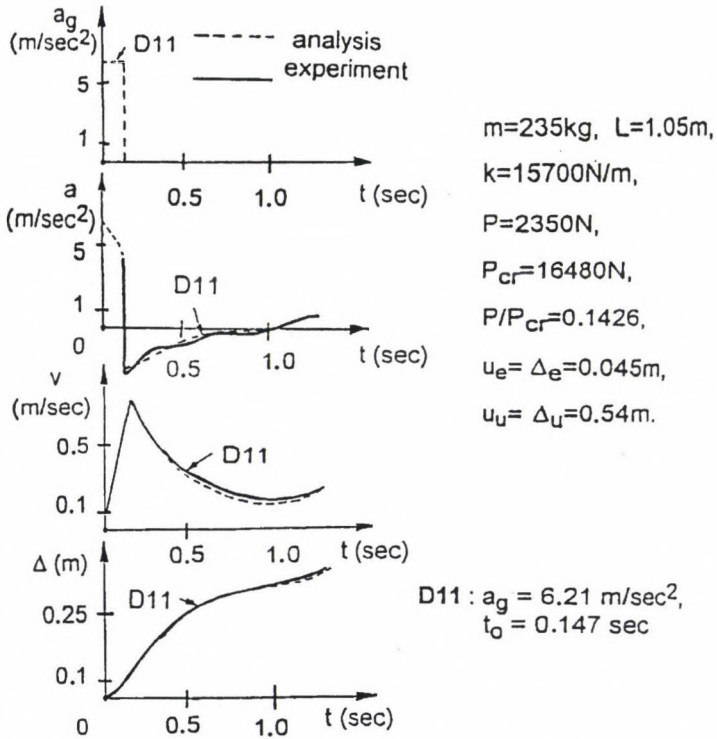


Fig. 15. Acceleration, velocity and path (displacement) diagrams of mass  $m_1$  of specimen D11 as functions of time excited by an impulse larger than the critical impulse (experimental results: thick line. calculated result: thin broken line)

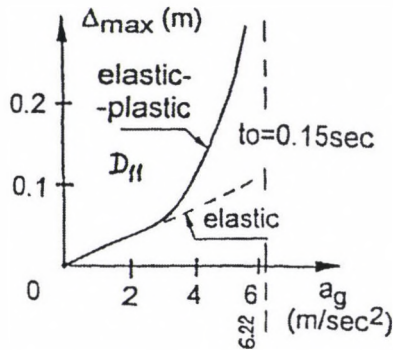


Fig. 16. Maximum displacement of the elements investigated as a function of the exciting acceleration

Fig. 16. Below the elastic limit, displacement  $u_{\max}$  is proportional to  $a_g$ , while beyond the elastic limit, it increases rapidly until the critical exciting acceleration  $a_g = 6.22$  m/sec is reached where, theoretically, the moving increases infinitely. Of course, the infinite value is a result of an approximation since, in fact, the end of the bar moves along a circular arc (and not parabolically), and thus  $u_{\max}$  could be not higher than  $L$ .

## 5. Conclusions

As clearly follows from what has been said so far, the usual calculation method of load reduction with the effect of ductility taken into consideration is erroneous and it results in some cases in a considerable underdimensioning. The error is a result of several factors described below:

– Instead of using the effect of ductility to increase the resistance of the structure, it is used to reduce the load in the usual seismic design method. As a result, the displacements of the structure under reduced load, calculated by elastic dynamics, are small enough to make the designer to believe that the effect  $P - \Delta$  can be taken into consideration on the basis of the elastic principles. In fact, much larger displacements take place than those calculated in this way, and in the case of these larger displacements, the effect of  $P - \Delta$  is significant and cannot be described on the basis of the linear theory. Thus the use of a more accurate plastic dynamic method, taking the effect of force  $P$  as well as the effect of plasticity into consideration, seems to be justified.

– In the usual design method, the calculation is based on the elastic response spectrum, and the internal forces (and from these the deformations) are calculated in general from the acceleration spectrum. The elasto-plastic response spectrum differs considerably from the elastic one and it increases instead of decreasing as the period increases. The displacement answer spectrum calculated for the experimental setup is shown in Fig. 17. The decreasing elastic answer spectrum has led the designer to believe that with the damage experienced by the structure (degradation in strength and rigidity), also the internal forces decrease. On the contrary, a certain reduction in the plastic system is followed by increase again as a result of which the structure under increased forces collapses. The acceleration spectrum cannot be taken as a basis for design in the plastic state for the very reason that there is no definite relation between acceleration and displacement in a plastic structure and therefore seismic design on the basis of the acceleration spectrum may be considerably misleading.

– In the case of a plastic system, the behaviour of the structure depends on several factors, among them e.g. on the rate of ductility, the ratio between load  $P$  and critical load  $P_{cr}$ , the relationship between the critical load and the plastic horizontal limit force, the duration of the shock impulses etc. These data become known only at the end of the design procedure and thus the calculation must be repeated again and again to make sure that the structure in question is suited for the actual internal forces and complies with the deformation limits.



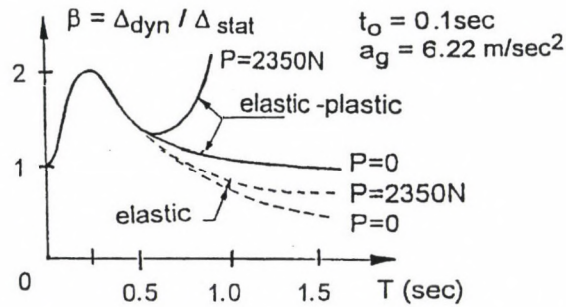


Fig. 17. Displacement spectra of elements  $T$  of different vibration periods (different spring constants) each in the elastic and elasto-plastic case, with and without compressive force  $P$  (with an excitation of an acceleration of  $a_g = 6.22 \text{ m/sec}^2$  for a time of  $t_0 = 0.1 \text{ sec}$ )

– In the case of reinforced concrete structures, repeatedly arising internal forces cause a degradation in strength and rigidity and thus a considerable reduction in the resistance of the structure. Hence the behaviour of the reinforced concrete structure is more disadvantageous than that of the theoretically elasto-plastic structure.

– In the case of steel structures, the elements of the cross-section always undergo buckling under plastic (ductile) load. Although this buckling does not necessarily result in a reduction in strength, it always results in a reduction in rigidity. This affects the behaviour of the structure unfavourably and thus the question arises whether it is allowed to take into consideration the effect of ductility in the case of steel structures at all?

– While the elastic system responds to short-time and long-time shocks identically, the response of the elasto-plastic system to shocks of different duration is different and, as compared with the short-time shock: a long-time shock of a much lower value may be enough to destroy the structure.

– In ductility based design, investigation of the velocity of movement of the structure is indispensable because the stability of the structure depends on whether the velocity of movement will be come zero or not. Therefore, should ductility be taken into consideration, determination of the value of critical displacement  $\Delta_{cr}$  can hardly be avoided. The limit ductile deformation plays a role in the problem only in that its value must be higher than the value of the critical deformation, that is.  $\Delta_u > \Delta_{cr}$  is required.

On the basis of what has been said above, the widely used ductility based method of seismic design seems to need a revision. Described below briefly is a method of seismic design which is, in our opinion, much more correct than the method used today.

## 6. Recommended method of seismic design

First of all, it is reasonable to consider whether it is economically justified to make use of the load bearing capacity reserve resulting from ductility. Also, we must remember that a quake of an intensity of  $I = 5-6$  exerts a load to the structure of buildings which is approximately equal to the load caused by wind. Since the buildings must be designed at any rate for windload, we need not deal with earthquakes below an intensity of  $I = 5$ . With a rather strong earthquake of a magnitude of  $M = 7.5$ , an intensity of  $I = 11$  is associated in the epicentre according to the 12-division scale. For an earthquake more intensive than that, an appropriate load bearing structure cannot be designed at all. As compared with an intensity  $I = 5-6$ , seismic design for an earthquake of an intensity  $I = 11$  means design for a load 40 to 50 times as much, however, the increase in costs is only 20 to 25% in the total costs of the building. Considering that the load increases quadratically with the intensity, the increase in costs can be estimated at 5-6% in case of an intensity  $I = 8-9$ .

A correct design making use of the ductility would result in saving half of the additional costs of max. 20-25%. Thus in highly seismic regions, investments costs about 10-11% could be saved if a correct seismic design of ductility, were applied.

The structure of the building forced to undergo ductile deformation suffers a permanent strain after which a restoration of the building is usually impossible. The time of recurrence of large earthquakes complies with the lifetime of the buildings and thus the earthquake is expected to result in serious damages to at least 20 to 50% of the buildings. And if the building must be demolished, this will affect 100% of the value of the building. Thus 20 to 50% of the value of the buildings can be opposed as a loss to savings of 10-12%.

At the same time, the deformations will be small if the building keeps its elastic state in the case of an earthquake. The damages can be repaired simply since the load bearing structure is not damaged. The situation is different in a region where the recurrence time of earthquakes is much shorter than the lifetime of the buildings, In such regions, the ductility might possibly be taken into consideration since the cost ratios might be different.

Taking all what has been said above into consideration, it can be recommended that the buildings be not permitted to get into plastic (ductile) state under standard earthquake conditions. The effect of ductility, which might be favourable, shall remain a reserve for saving human life. Although the construction costs of the buildings may increase somewhat, but the damages due to earthquake will decrease considerably for the same period, and certainly much less human lives will fall victim to earthquake, Note that on the basis of an erroneous analysis, the effects of ductility have been overestimated and utilized for the existing buildings all over the world at present, and thus the benefits of what has been said would appear not before decades.

Remember that ductility has been introduced to seismic design because of its energy absorbing and vibration damping effect. An examination of the differential equation shows that this effect is similar to the Coulomb (frictional) damping effect. Best use can



be made of the favourable consequences without unfavourable side-effects if the damping system is separated from the load bearing structures. By so doing we do not endanger the stability of the building. Optimum damping can be achieved in this way, and in the case of a possible damage, the friction brakes can be replaced in a relatively simple way and at reasonable costs. This system is similar to a motor vehicle where the brake system is built in independently. Whoever would think of reducing the production costs of a motor vehicle by building no brake in, relying on the vehicle to stop anyway when it runs into a tree.

Nor should we forget that so far based isolation systems of quite a number have been developed which are capable of eliminating 70 to 80% of the load exerted to the building by earthquake.

By using the different methods in combination, damages due to earthquake can be almost entirely avoided.

## 7. Notation

$a$	= acceleration	$a_g$	= acceleration of impulse excitation
$g$	= gravitational acceleration	$h$	= height of fall
$k$	= spring constant	$m$	= mass
$s$	= vertical displacement	$t$	= time
$t_0$	= time of impulse excitation	$u = \Delta$	= horizontal displacement
$v$	= velocity	$E$	= elastic (Young) modulus
$E_{(i)}$	= energy	$E_{kin}$	= kinetical energy
$F$	= horizontal force	$F_e$	= replacement static force
$F_u$	= plactic force	$I$	= intensity of earthquake
$I_t$	= impulse	$L$	= length of bar
$L_{ex}$	= work of external force	$L_{in}$	= work of internal forces
$M$	= bending moments	$P$	= gravitation force
$P_{cr}$	= critical force	$T$	= vibration period
$\alpha$	= reducing factor	$\beta$	= dynamic factor
$\mu$	= ductility factor	$\sigma_y$	= yield strenght
$\Delta = u$	= horizontal displacement	$\Delta_e$	= limit of elastic deformation
$\Delta_0$	= deformation without P force	$\Delta_u$	= limit of plastic deformation
$\Pi$	= potential energy	$\Pi_{ex}$	= potential energy of external forces
$\Psi$	= factor of buckling deformation	$\Pi_{in}$	= potential energy of internal forces
$du / dt = u \dot{\phantom{u}}$	= differential quotient of $u$ displacement by $t$ time		



### References

1. Newmark, M. N.–Rosenbluth, E.: *Fundamentals of Earthquake Eng.* Prentice-Hall, Inc, London, 1971
2. Müller, F. P.–Keintzel, E.: *Erdbebensicherung von Hochbauten.* Ernst und Sohn, Berlin 1984
3. Paulay, T.–Bachmann, H.–Moser, K.: *Erdbebenbemessung von Stahlbetonhochbauten.* Birkhauser V., Basel–Boston–Berlin, 1990
4. NEHRP: *Recommended Provision for the Development of Seismic Regulations for New Buildings.* Building Seismic S.C. Washington, D.C., 1991
5. Eurocode 8. *Structures in Seismic Regions.* Commission of the European Communities, 1988
6. Bognár, L.: Stability of elasto-plastic plane bar system. *Acta Tech. Ac. Hung.* 101 (1988) 5–47
7. Bognár, L.: Calculation for frame structures exposed to earthquake on the basis of the random vibration theory (In Hungarian). *Építés-Építészettudomány XXIII* (1992–93) No. 1–2
8. Bognár, L.: Investigation of the state of stability of discrete structural systems of nonelastic material (In Hungarian). Vol. 4. Experiments. SÁMSON Co-op. 1995
9. Ziegler, H.: Die Stabilitätskriterien der Elastomechanik. *Ing. Archiv XX.* (1952) 49–56.
10. Kaliszky, S.: *Plasticity Theory and Engineering Applications,* Akadémiai Kiadó–Elsevier Sci. Publ. Amsterdam, 1989
11. Csák, B.: Ultimate analysis of structures exposed to seismic effects. *Periodica Politechnica* 22 (1987) 81–116
12. Smith, J. W.: *Vibration of Structures.* Chapman and Hall, London–New York, 1988
13. Roy, R.–Craig, Jr.: *Structural Dynamics.* J. Wiley and SONS, New York, 1981
14. Vértes, Gy.: *Structural Dynamics.* Elsevier, Amsterdam–Oxford–New York–Tokyo, 1985



## ON THE DIVISION OF THE KINETIC ENERGY (POWER) CONTENT IN TURBULENT FLOW (An Approximate Analytical Review)

Hankó, Z.

*Water Resources Research Plc. (VITUKI Plc.), H-1095 Budapest, Kvassay Jenő út 1, Hungary*

(Received: 11 March, 1997)

The kinetic energy (power) content in turbulent flow can be divided between the turbulent fluctuation and the main (mean) flow. In spite of that the former one is a few per cent only of the total one its role is significant in e.g. sediment carrying or pollutant dispersal (generally speaking: transport) capacity of the flow. The approximate analytical review enlightens on the mutual interaction of the turbulent fluctuation and the main (mean) flow. The most important result of this review is the clear picture of the composition of the kinetic power. This shows that the kinetic power of the turbulent fluctuation consists not only of the generally known local and convective components but also of a second convective term which indicates a kinetic energy convection of the turbulent fluctuation by the main (mean) flow velocity. And depending on the signs of the space derivatives it can indicate energy transport in both directions [either from the main (mean) flow to the turbulent fluctuation or vice versa].

### Notation

$\frac{D}{Dt}$	– material (total) derivative
$e$	– subscript, refers to ergodic state
$f_i$	– $i$ -th component of body force of elementary water body of unit mass
$i, j = 1, 2, 3$	– subscripts refer to components parallel to Cartesian coordinate axes in Eulerian description
$N$	– number of experiments (tends to infinity)
$P$	– instantaneous hydrostatic pressure
$\bar{p}$	– main (time-mean) part of hydrostatic pressure
$p'$	– instantaneous part of the hydrostatic pressure by turbulent fluctuation
$\bar{p}_t$	– normal part of Reynolds stresses
$t$	– running time coordinate
$t_0$	– definite time point



$T$	– time limit of time span (tends to infinity)
$T_f$	– time limit (finite, in order of magnitude of minutes) for calculating time-mean characteristics in turbulent water flow
$T_{mf}$	– kinetic power of elementary body of unit mass of main (mean) turbulent water flow
$T_n$	– period time of natural or man-made periodicity in instationary water flow
$T_r$	– relative transport capacity of turbulent water flow
$T_t$	– period time(s) of turbulent fluctuation
$T_{tf}$	– kinetic power of elementary body of unit mass of (mean) turbulent fluctuation of water flow
$U; U_1=U_x, U_2=U_y, U_3=U_z; U_i, U_j$	– instantaneous velocity; components according to co-ordinate axes
$\bar{U}$	– main velocity (time-mean over finite time span: $T_f$ )
$\bar{u}; \bar{u}_i; \bar{u}_j$	– main (mean) part of instantaneous velocity; and $i$ -th and $j$ -th components of it
$u'; u'_i; u'_j$	– instantaneous part of velocity by turbulent fluctuation; and $i$ -th and $j$ -th components of it
$X, Y, Z$	– limits of space spans (tend to infinite) of space coordinates
$x, y, z$	– running space coordinates
$x_o, y_o, z_o$	– space coordinates of a definite point
$x_i, x_j$	– $i$ -th, $j$ -th space coordinates
$\epsilon_n$	– apparent, turbulent, eddy viscosity
$\xi, \eta, \zeta$	– space-spans according to $x, y, z$ space coordinates
$\partial$	– symbol of partial derivative
$\delta_{ij}$	– Kronecker delta (if $j \neq i$ then $\delta_{ij} = 0$ and if $j = i$ then $\delta_{ij} = 1$ )
$\mu$	– constant molecular dynamic viscosity of water
$\nu$	– constant molecular kinematic viscosity of water
$\rho$	– constant density of water
$\overline{\sigma_{ij}} = -\rho \overline{u'_i u'_j}$	– Reynolds stresses
$\Sigma$	– sum up symbol
$\tau$	– time span according to $t$ time coordinate

## 1. Introduction

It is well known that in hydraulic engineering practice both the supercritical/subcritical state of flow (in open channel flow) and the laminar/turbulent behaviour of flow have their significance in the relevant cases.

When constructing a chute during training a creek the state of the flow in the head-water race may be subcritical while running down the chute its state may be supercritical. Again following the energy dissipator the state of flow may be also subcritical in

the tailwater race of the trained creek. Along the path described above the behaviour of flow is turbulent. Generally "open surface" flow in hydraulic engineering is turbulent. Probably the only exception is the "vail-like" waterflow if the order of magnitude of the vail depth is not more than millimeter and the order of magnitude of its velocity is not more than millimetres per second.

The state of water flow in closed conduits is generally turbulent. An important exception is seepage flow of water in a system of closed conduits which has a range (seepage flow of Darcy) the state of flow of which is laminar.

Thus it can be stated that in point of view of hydraulic engineering the state of turbulence in the water flow is essential.

## 2. Definitions and statements

In formulating a general definition regarding turbulent flow the best starting point can be Hinze's (1959) definition. Following his row of thoughts it can be stated: "Turbulent fluid motion is an irregular condition of flow in which the various quantities of the fluid and that of the flow show a random variation with time and space coordinates but so that statistically distinct mean values can be discerned."

Comparing the hydraulic engineering aspects with the definition of the turbulent flow above and neglecting the exceptions it can be stated that water flow may be treated as an incompressible fluid flow with viscosity constant. Thus only the statistical means of the flow characteristics have any relevance.

Three different statistical means can be defined e.g. regarding *flow velocity* (Korn–Korn, 1968).

(a) The *time-mean* of velocity in a definite point of the space in a *stationary* turbulent flow is depending only of the space coordinates:

$$\overline{U(x_o, y_o, z_o)}^* = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{+T/2} u(x_o, y_o, z_o; t + \tau) d\tau, \quad (1/a)$$

in which  $U$  refers to the instantaneous velocity,  $x_o, y_o, z_o$  – space coordinates of a definite point,  $t$  – the running time coordinate,  $\tau$  – time span over which the statistical mean will be calculated,  $T$  – limit of  $\tau$ .

(b) The *space-mean* of velocity in a definite time point in a *homogeneous* turbulent flow is depending only of the time coordinate:

\*Overbar indicates statistical mean, and the symbol above it refers to the quantity over which the statistical mean is calculated.

$$\frac{\bar{\xi}, \bar{\eta}, \bar{\zeta}}{U(t_0)} = \lim_{z \rightarrow \infty} \frac{1}{z} \int_{-z/2}^{+z/2} \lim_{y \rightarrow \infty} \frac{1}{y} \int_{-y/2}^{+y/2} \lim_{x \rightarrow \infty} \frac{1}{x} \int_{-x/2}^{+x/2} U(x + \xi, y + \eta, z + \zeta; t_0) d\xi d\eta d\zeta, \quad (1/b)$$

in which  $\xi$ ,  $\eta$  and  $\zeta$  refer to the space-span of the running space coordinates:  $x$ ,  $y$ ,  $z$  over which the statistical means are calculated,  $X$ ,  $Y$ ,  $Z$  – limit of  $\xi$ ,  $\eta$ ,  $\zeta$ , respectively, and  $t_0$  – time coordinate of a definite time point.

(c) If the turbulent flow is neither stationary nor homogeneous then the *average-mean* can be defined as the mean of  $N$  experiments (performed amongst identical circumstances), e.g.

$$\frac{a}{U(x_0, y_0, z_0; t_0)} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N U(x_0, y_0, z_0; t_0). \quad (1/c)$$

(d) According to the definition of the ergodic hypothesis of random phenomena if the flow is both stationary and homogeneous then the time-mean and the space-mean to be equal with the average-mean. That means

$$\frac{\tau}{U_e(x_0, y_0, z_0)} = \frac{\bar{\xi}, \bar{\eta}, \bar{\zeta}}{U_e(t_0)} = \frac{a}{U_e(x_0, y_0, z_0; t_0)}. \quad (2)$$

where subscript “e” refers to the ergodic state.

### 3. Basic assumptions in interpreting the disorder of turbulence analytically in hydraulic engineering

The real turbulent flow is never stationary in the strict sense of the word but it can often be supposed as quasi-stationary (during short-time spans) because of the slow variance of the flow in time. Simultaneously the homogeneity of turbulence cannot also be supposed because at least along the boundaries (in the laminar and turbulent shear layers) the statistical means of flow characteristics perpendicular to the boundary have significant gradients showing anisotropy based on (gradual) inhomogeneity. Furthermore, the statistical means over infinite spans cannot be interpreted in practical applications.

For analysing turbulence it seems to be the most relevant fact that the time-mean values serve the most valuable information. First of all, it is important to mention that according to the results gained from experimental evidence the period of turbulent fluctuation ( $T_f$ ) varies between two distinct limits, namely

$$10^{-4} < T_f [\text{sec}] < 1. \quad (3)$$



This range of time exceeds e.g. with many orders of magnitude the period of collision due to the Brown motion in gaseous fluids which can be characterized as  $10^{-9} \div 10^{-10}$  [sec]. At the same time, the period ( $T_n$ ) of natural or man-made instationarity of water flow is with many orders of magnitude longer because its characteristic value may be hour.

Thus, it is possible to find a time span which is much longer than the longest period of turbulent fluctuation and much shorter than the natural or man-made period of instationarity of water flow:

$$T_{max} \ll T_f \ll T_{min} \quad (4)$$

This finite  $T_f$  time-span (in order of magnitude of minutes) can be the basis in calculating time-mean and this can be accepted as characteristic mean in analysing turbulent flow in hydraulic engineering, thus (exemplifying by velocity):

$$\bar{U} = \frac{1}{T_f} \int_{-T_f/2}^{+T_f/2} U(t + \tau) d\tau. \quad (5)$$

Here and in the following *overbar* (without any superscript) indicates a time-mean value calculated over a finite  $T_f$  time span defined previously.

In the following the instantaneous value (exemplified by velocity) will be indicated by  $U$  which is the sum of the characteristic mean defined by Eq. (5) and indicated by  $\bar{u}$  plus the instantaneous value of turbulent fluctuation designed by  $u'$  thus

$$U = \bar{u} + u' \quad (6/a)$$

and by definition

$$\overline{\bar{u}} = \bar{u} \quad \text{and so} \quad \overline{u'} = 0. \quad (6/b, c)$$

#### 4. Equilibrium of forces acting on an elementary water body of unit mass

There are two relationships describing the interrelation between the motion characteristics of water flow (incompressible fluid with viscosity constant): Equation of continuity and Navier–Stokes equation (interrelating the acting forces on a water body of unit mass). Applying the Cartesian coordinate system and the Eulerian description form the equation of continuity says (if there is no sink or source in the flow field) (Hughes–Gaylord, 1964):

$$\sum_{i=1}^3 \frac{\partial U_i}{\partial x_i} = 0 \quad (7/a)$$

and the  $i$ -th component of the Navier–Stokes equation:

$$\frac{D}{Dt} U_i = \frac{\partial U_i}{\partial t} + \sum_{j=1}^3 U_j \frac{\partial U_i}{\partial x_j} = f_i - \frac{1}{\rho} \frac{\partial P}{\partial x_i} + \nu \sum_{j=1}^3 \frac{\partial^2 U_i}{\partial x_j^2} \quad (7/b)$$

where  $U_i, U_j$  [m/sec] –  $i$ -th and  $j$ -th components of the instantaneous  $\mathbf{U}$  velocity vector;  $t$  [sec] – independent time coordinate;  $x_i, x_j$  [m] –  $i$ -th and  $j$ -th components of independent space coordinates;  $P$  [N/m<sup>2</sup> = kg/m sec<sup>2</sup>] – instantaneous hydrostatic pressure;  $\rho$  [kg/m<sup>3</sup>] – constant water density;  $\nu = \mu/\rho$  [m<sup>2</sup>/sec] – constant kinematic viscosity of water;  $f_i$  [m/sec<sup>2</sup>] –  $i$ -th component of body force vector:  $\mathbf{f}$  (acting on a water body of unit mass). In this relationship  $i = 1, 2, 3$  and  $j = 1, 2, 3$  where  $1 \equiv x$  (flow direction, horizontal, forward);  $2 \equiv y$  perpendicular to  $1 \equiv x$ , horizontal and left directed;  $3 \equiv z$  perpendicular to  $x, y$  plan, vertical, upward direction.

In this equation system consisting of  $1 + 3 = 4$  equations there are four unknown quantities:  $U_1 = U_x, U_2 = U_y, U_3 = U_z$  and  $P$ , and if the boundary and initial conditions are known then the equation system can be solved. Because any general solution does not exist, therefore numerical evaluation prevail first of all in one- or two-dimensional approximations.

*This equation system is valid both for subcritical/supercritical flow and for laminar/turbulent state of flow.* Taking into account Eqs (5) and (6/a–c) it can be seen that this system of equations at turbulent state of flow has to be valid also for statistical mean values. This means that

$$U_i = \bar{u}_i + u_i'; \quad U_j = \bar{u}_j + u_j'; \quad P = \bar{p} + p' \quad (8/a-c)$$

and in calculating means Eqs (6/a–c) are to be applied.

The equation of continuity for mean values reads as follows

$$0 = \sum_{i=1}^3 \frac{\partial \bar{u}_i}{\partial x_i} = \sum_{i=1}^3 \frac{\partial (\bar{u}_i + \overline{u_i'})}{\partial x_i} = \sum_{i=1}^3 \frac{\partial \bar{u}_i}{\partial x_i} \quad (9/a)$$

As the consequence of Eq. (9/a) it can be stated that

$$0 = \sum_{i=1}^3 \frac{\partial U_i}{\partial x_i} = \sum_{i=1}^3 \frac{\partial \bar{u}_i}{\partial x_i} = \sum_{i=1}^3 \frac{\partial u_i'}{\partial x_i} \quad (9/b-d)$$

which means that equation of continuity is valid simultaneously both for instantaneous velocity components and for time-mean values of velocity components and for the instantaneous values of turbulent velocity fluctuation components.

Similar to the equation of continuity the time-mean form of the Navier–Stokes equation can also be evaluated

$$\begin{aligned} \frac{D\bar{U}_i}{Dt} &= \frac{D(\bar{u}_i + u'_i)}{Dt} = \frac{\partial(\bar{u}_i + u'_i)}{\partial t} + \sum_{j=1}^3 \overline{(\bar{u}_j + u'_j)} \frac{\partial(\bar{u}_i + u'_i)}{\partial x_j} = \\ &= f_i - \frac{1}{p} \frac{\partial(\bar{p} + p')}{\partial x_i} + \nu \sum_{j=1}^3 \frac{\partial^2(\bar{u}_i + u'_i)}{\partial x_j^2}. \end{aligned} \tag{10/a}$$

Taking into account Eqs (6/a–c) and rearranging Eq. (10/a) it reads

$$\begin{aligned} \frac{D\bar{u}_i}{Dt} &= \frac{\partial \bar{u}_i}{\partial t} + \sum_{j=1}^3 \bar{u}_j \frac{\partial \bar{u}_i}{\partial x_j} = f_i - \frac{1}{p} \frac{\partial \bar{p}}{\partial x_j} + \\ &+ \sum_{j=1}^3 \left( \nu \frac{\partial^2 \bar{u}_i}{\partial x_j^2} - \overline{u'_j \frac{\partial u'_i}{\partial x_j}} \right). \end{aligned} \tag{10/b}$$

Adding the following terms

– to the second term, left side:

$$0 = \sum_{j=1}^3 \bar{u}_i \frac{\partial \bar{u}_j}{\partial x_j}, \tag{11/a}$$

– to the last but one term on the right side:

$$0 = \sum_{j=1}^3 \nu \frac{\partial^2 \bar{u}_j}{\partial x_i \partial x_j}, \tag{11/b}$$

– and to the last term on the right side:

$$0 = \sum_{j=1}^3 -\overline{u'_i \frac{\partial u'_j}{\partial x_j}} \tag{11/c}$$

which (all the three) are special forms of continuity the following formula will be got

$$\begin{aligned} \frac{D\bar{u}_i}{Dt} &= \frac{\partial \bar{u}_i}{\partial t} + \sum_{j=1}^3 \frac{\partial(\bar{u}_i \bar{u}_j)}{\partial x_j} = f_i - \frac{1}{p} \frac{\partial \bar{p}}{\partial x_i} + \\ &\text{(i) (ii) (iii) (iv) (v)} \\ &+ \sum_{j=1}^3 \left\{ \frac{\partial}{\partial x_j} \left[ \nu \left( \frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \right) - \overline{u'_i u'_j} \right] \right\} \\ &\text{(vi) (vii)} \end{aligned} \tag{12/a}$$



where (i) denotes the  $i$ -th, total time-mean component of acceleration due to the acting forces on a water body of unit mass which equals the sum of (ii) the local and (iii) plus the convective acceleration. The right-hand side of the equation consists of (iv)  $i$ -th component of the body forces acting on a water body of unit mass and (v) acceleration due to hydrostatic pressure. The last two term on the right-hand side of the equation represent the effect of resistance. The term (vi) indicates acceleration due to body deformation by viscous forces and (vii) acceleration by turbulent velocity fluctuations.

Equation (12/a) clearly indicates the effect of turbulent fluctuation on the main (mean) flow. This form has been introduced first by Osborn Reynolds therefore its characteristic term is often called as Reynolds stresses:  $\overline{\sigma'_{ij}} = -\rho \overline{u'_i u'_j}$ . J. Boussinesq has been the first one who supposed that resistive forces caused by the Reynolds stresses can be explained by the acceleration due to the deformation of the main (mean) flow (similar to the viscous stresses), only instead of the molecular viscosity coefficient ( $\nu = \mu/\rho$ , material characteristic) an “apparent”, “turbulent” or “eddy” viscosity coefficient has to be introduced (designed by  $\epsilon_m$  [m<sup>2</sup>/sec] and it is “motion characteristic” because it is effective only then if motion exists). Taking into account that Reynolds stresses include normal components, too, and these can be characterized by spherical symmetry, thus, they can be expressed as

$$\frac{1}{\rho} \sum_{j=1}^3 \frac{\partial \overline{\sigma'_{ij}}}{\partial x_j} = \sum_{j=1}^3 \frac{\partial (-\overline{u'_i u'_j})}{\partial x_j} = \sum_{j=1}^3 \left\{ \frac{\partial}{\partial x_j} \left[ \epsilon_m \left( \frac{\partial \overline{u}_i}{\partial x_j} + \frac{\partial \overline{u}_j}{\partial x_i} \right) - \frac{\delta_{ij}}{\rho} \overline{p}_t \right] \right\}, \quad (12/b)$$

in which – besides the foregoing explained ones –  $\delta_{ij}$  is the Kronecker delta ( $\delta_{ij} = 0$  if  $j \neq i$  and  $\delta_{ij} = 1$  if  $j = i$ ) and  $\overline{p}_t$  [N/m<sup>2</sup> = kg/m sec<sup>2</sup>] – normal part of Reynolds stresses.

In a significant group of turbulent water flows (wide, shallow, open channel flow  $\overline{p}_t = \text{constant}$  and  $\epsilon_m = \text{constant}$  supposition yields a reasonable agreement between theoretical and experimental results. But there are other situations at which higher- (e.g. fourth-) order tensor is needed instead of  $\epsilon_m = \text{constant}$ , scalar for acceptable agreement in spite of that its correct physical meaning is not fully understood as yet.

Substituting Eq. (12/b) into Eq. (12/a) and rearranging it reads

$$\frac{D\overline{u}_i}{Dt} = \frac{\partial \overline{u}_i}{\partial t} + \sum_{j=1}^3 \frac{\partial (\overline{u}_i \overline{u}_j)}{\partial x_j} = f_i - \frac{1}{\rho} \frac{\partial (\overline{p} + \overline{p}_t)}{\partial x_i} + (\nu + \epsilon_m) \sum_{j=1}^3 \frac{\partial^2 \overline{u}_i}{\partial x_j^2}. \quad (12/c)$$

As it is to be seen  $\overline{p}_t$  and  $\epsilon_m$  are completing the relevant original terms with physical meaning as  $\overline{p}$  and  $\nu$ . In wide, shallow, open channel flow  $\overline{p}_t = \text{constant}$ , scalar has the same order of magnitude as  $\overline{p}$ , at the same time in closed conduit under high overpressure  $\overline{p}_t$  will probably be negligible. The eddy viscosity ( $\epsilon_m$ ) compared to the mo-

lecular viscosity ( $\nu$ ) may be of order(s) of magnitude higher in water flow. Using any experimental measurement(s) for recalculating the numerical value of eddy viscosity and if it results in a negative value (supposing reasonable numerical range of  $\overline{p}_t$ ) then it proves that  $\varepsilon_m = \text{constant}$ , scalar supposition is not a solid one.

### 5. Kinetic energy (power) in turbulent water flow

Starting with Eq. (12/a) expressing time-mean equilibrium of forces acting on an elementary water body of unit mass it is easy to reach an equation which is the time-mean equilibrium of kinetic energy, exactly speaking, equilibrium of kinetic power. Namely, it is well known that force multiplied by velocity with identical direction: results in kinetic power. Thus, multiplying Eq. (12/a) with  $\overline{u}_i$  velocity the result will be the *i*-th component of kinetic power of time-mean equation of main flow. Therefore after rearranging

$$\begin{aligned}
 \frac{1}{2} \frac{D\left(\frac{2}{u_i}\right)}{Dt} &= \frac{1}{2} \frac{\partial\left(\frac{2}{u_i}\right)}{\partial t} + \sum_{j=1}^3 \frac{1}{2} \frac{\partial\left(\frac{2}{u_i u_j}\right)}{\partial x_j} = \overline{u}_i f_i - \frac{\overline{u}_i}{p} \frac{\partial \overline{p}}{\partial x_i} + \\
 &\quad \text{(i)} \qquad \qquad \text{(ii)} \qquad \qquad \text{(iii)} \qquad \qquad \text{(iv)} \qquad \qquad \text{(v)} \\
 + \sum_{j=1}^3 \left\{ \nu \left[ \frac{\partial}{\partial x_j} \left[ \overline{u}_i \left( \frac{\partial \overline{u}_i}{\partial x_j} + \frac{\partial \overline{u}_j}{\partial x_i} \right) \right] - \frac{\partial \overline{u}_i}{\partial x_j} \left( \frac{\partial \overline{u}_i}{\partial x_j} + \frac{\partial \overline{u}_j}{\partial x_i} \right) \right] \right\} + &\quad \text{(13/a)} \\
 &\quad \text{(vi)} \qquad \qquad \qquad \qquad \qquad \qquad \text{(vii)} \\
 + \frac{\partial}{\partial x_j} \left[ \overline{u}_i \left( -\overline{u'_i u'_j} \right) \right] - \frac{\partial \overline{u}_i}{\partial x_j} &\quad \text{(viii)} \qquad \text{(ix)}
 \end{aligned}$$

In this relationship related to the elementary water body of unit mass and time-mean value of the *i*-th component (i) refers to the total kinetic power which is the sum of (ii) local and (iii) convective kinetic power. This sum is in equilibrium with the sum of kinetic power terms of (iv) body forces, (v) hydrostatic pressure (convective diffusion) and the sum of resistive forces, respectively. The sum of kinetic power of the resistive forces consists of (vi) the kinetic power due to viscous deformation and (vii) viscous dissipation furthermore with (viii) the convective transport of the Reynolds stresses by the main (mean) flow velocity and (ix) the kinetic power due to the Reynolds stresses exerted on deformation of the main (mean) flow.

Similarly, Eq. (7/b) force-equilibrium regarding the instantaneous values can be transformed to instantaneous kinetic power equilibrium; and forming its time-mean value for the *i*-th component it reads as follows



$$\begin{aligned} \frac{1}{2} \frac{D}{Dt} \left( \frac{2}{u_i} + \overline{u_i'^2} \right) &= \frac{1}{2} \frac{\partial}{\partial t} \left( \frac{2}{u_i} + \overline{u_i'^2} \right) + \frac{1}{2} \sum_{j=1}^3 \frac{\partial}{\partial x_j} \overline{(\overline{u_i} + u_i')^2 (\overline{u_j} + u_j')} = \\ &= \overline{u_i} f_i - \frac{\overline{u_i} + u_i'}{p} \frac{\partial (\overline{p} + p')}{\partial x_i} + v \sum_{j=1}^3 \overline{(\overline{u_i} + u_i')} \frac{\partial^2 (\overline{u_i} + u_i')}{\partial x_j^2}. \end{aligned} \tag{13/b}$$

After rearranging Eq. (13/a) may be subtracted, and then the residuum shows the *time-mean equation of kinetic power of the i-th component of turbulent fluctuation* as follows

$$\begin{aligned} \frac{1}{2} \frac{D \overline{u_i'^2}}{Dt} &= \frac{1}{2} \frac{\partial \overline{u_i'^2}}{\partial t} + \frac{1}{2} \sum_{j=1}^3 \frac{\partial}{\partial x_j} \overline{u_i'^2 u_j'} = - \frac{\overline{u_i'} \partial p'}{p \partial x_i} + \\ &\quad \text{(i) \quad (ii) \quad (iii) \quad (iv)} \\ &+ \sum_{j=1}^3 \left\{ v \left[ \frac{\partial}{\partial x_j} \overline{u_i' \left( \frac{\partial u_i^l}{\partial x_j} + \frac{\partial u_j'}{\partial x_i} \right)} - \frac{\partial u_i^l}{\partial x_j} \overline{\left( \frac{\partial u_i^l}{\partial x_j} + \frac{\partial u_j'}{\partial x_i} \right)} \right] - \right. \\ &\quad \text{(v) \quad (vi)} \\ &\quad \left. - \frac{1}{2} \frac{\partial}{\partial x_j} \overline{(u_i'^2 u_j')} + \overline{u_i' u_j'} \frac{\partial \overline{u_i}}{\partial x_j} \right\}. \\ &\quad \text{(vii) \quad (viii)} \end{aligned} \tag{13/c}$$

This relationship summarizes the *i*-th time-mean component of kinetic power exerted by the turbulent fluctuation. Term (i) expresses the total kinetic power of elementary water body of unit mass which is the sum of (ii) the local and (iii) the convective terms, respectively. And this is in equilibrium with the sum of (iv) the kinetic power of turbulent fluctuation of hydrostatic pressure (convective diffusion) and the sum of kinetic power resulted in by resistance; furthermore the extraordinary convective term of kinetic power. Among the resistance terms (v) expresses the viscous deformation and (vi) viscous dissipation caused by turbulent fluctuation and (viii) kinetic power exerted on the deformation of the main flow by the turbulent fluctuation. And, last but not least, the (vii) extraordinary convective term: convection of turbulent energy by the main (mean) flow velocity.

Equation (13/c) clearly shows how much the kinetic power of turbulent fluctuation is affected by the main (mean) flow. First, the extraordinary convective term [term (vii) in Eq. (13/c)] may have a double role. Depending on the signs of the space derivatives of the kinetic energy of the turbulent fluctuation  $\left( \frac{\overline{u_i'^2}}{2} \right)$  and the sign of the convective main (mean) flow velocity  $\left( \overline{u_j} \right)$  it can indicate power transport both from main flow into the fluctuation or vice versa. Second, the kinetic power exerted on the viscous de-



formation of the turbulent fluctuation [term (v) in Eq. (13/c)] is completed by the kinetic power exerted on the deformation of the main flow by the Reynolds stresses [term (viii) in Eq. (13/c)]. This later one is the clear explanation of the transition from laminar to turbulent flow. If main flow gradient in the flow direction is positive the Reynolds number increases and at a certain variation the laminar state of flow will go over turbulent state (simultaneously the extraordinary convective term of kinetic power [term (vii) of Eq. (13/c)] must also increase for maintaining equilibrium).

Equation (13/c) expresses – as stated above – the *i*-th component of the time-mean kinetic power owing to turbulent fluctuation. But this relationship can be rearranged in a more conservative form (Rodi, 1984 and Rátky, 1995) as

$$\begin{aligned}
 \frac{1}{2} \frac{D \overline{u_i'^2}}{Dt} &= \frac{1}{2} \frac{\partial \overline{u_i'^2}}{\partial t} + \frac{1}{2} \sum_{j=1}^3 \frac{\partial}{\partial x_j} \left( \overline{u_i'^2 u_j'} \right) = \frac{\overline{u_i'} \partial p^1}{p \partial x_j} - \frac{1}{2} \sum_{j=1}^3 \frac{\partial}{\partial x_j} \left( \overline{u_i'^2 u_j'} \right) + \\
 \text{(i)} \quad \text{(ii)} \quad \text{(iii)} \quad \text{(iv)} & \\
 &+ \sum_{j=1}^3 \left[ \left( \overline{-u_i' u_j'} \right) \frac{\partial \overline{u_i'}}{\partial x_j} + \nu u_i' \frac{\partial^2 \overline{u_i'}}{\partial x_j^2} \right] \quad \text{(v)} \quad \text{(vi)}
 \end{aligned} \tag{13/d}$$

Equation (13/d) corresponds to Prandtl's (1945) time-mean kinetic energy transport equation owing to turbulent fluctuation. According to his explanation (cited by Rodi, 1984) term (i) corresponds to the total kinetic power of elementary body of unit mass owing to turbulent fluctuation which is the sum of term (ii) the local and term (iii) the convective constituents. And this is in equilibrium with the sum of term (iv): the so-called diffusive transport, plus term (v): production by shear and term (vi): viscous dissipation if the present *i*-th component is summed up according to the Cartesian coordinate axes as  $\sum_{i=1}^3$

This rearrangement does not cause any modification in the physical content, it alters the denomination of the terms only. In Eq. (13/d) – according to the *k-epsilon* model –

$$k = \sum_{j=1}^3 \frac{\overline{u_j'^2}}{2} = \overline{e_k} \tag{13/e}$$

and

$$\epsilon = \sum_{j=1}^3 \nu \sum_{i=1}^3 \overline{u_i' \frac{\partial^2 u_i'}{\partial x_j^2}} \tag{13/f}$$

and both terms (iv) plus (v) and term (vi) can only be approximated by semiempirical formulas and, thus, the *k-epsilon* model of the time-mean kinetic energy transport ow-

ing to turbulent fluctuation can numerically be solved as exemplified by a recent EDF study (1996/97), too, which compared the results of *various empirical approximations*.

A comparison of terms (i) = (ii) + (iii) in equation (13/c and d) deserve a little more attention. The difference can be found at tem (iii). In Eq. (13/c) the factors of the cross-correlation between the various components belong to the same domain of the velocity constituents: turbulent fluctuation, while in Eq. (13/d) the cross-correlation is a mixed up one because the convection of the turbulent fluctuation is generated by the main (mean) flow and, therefore, the equation (i) = (ii) + (iii) is not true (it is invalid). By changing term (iii) with the second part of term (iv) in Eq. (13/d) the equilibrium is re-established and – in this case – the diffusive transport [term (iv) in Eq. (13/d)] is mixed up term. And we arrived to the same consequences which are explained regarding Eq. (13/c) by deriving, the transition from laminar to turbulent flow and vice versa.

## 6. Transport capacity of turbulent water flow

Dividing the kinetic power of the turbulent flow between the constituents of the flow, namely between the main (mean) flow and the turbulent fluctuation a relative transport capacity term may be defined as

$$T_r = \frac{T_{tf}}{T_{mf} + T_{tf}} \quad (14/a)$$

where  $T_r$  – (dimensionless) relative transport capacity of turbulent flow,  $T_{mf}$  [kg m<sup>2</sup>/sec<sup>3</sup>] – kinetic power of main (mean) flow in turbulent water flow and  $T_{tf}$  [kg m<sup>2</sup>/sec<sup>3</sup>] – kinetic power of turbulent fluctuation (mean) in turbulent water flow.

Summing up the  $i = 1, 2, 3$  equations the total kinetic energy of an elementary water body of unit mass can be got. Denoting

$$\bar{e}_k = \sum_{j=1}^3 \frac{\left(\frac{2}{u_j}\right)}{2} \quad \text{and} \quad \bar{e}'_k = \sum_{j=1}^3 \frac{\overline{u_j^2}}{2} \quad (14/b, c)$$

where  $\bar{e}_k$  [m<sup>2</sup>/sec<sup>2</sup>] represents the kinetic energy of elementary water body of unit mass in the main (mean) water flow while  $\bar{e}'_k$  [m<sup>2</sup>/sec<sup>2</sup>] is the same for turbulent fluctuation (time-mean value, too). The kinetic power of the same constituents can be defined as follows:

$$T_{mf} = \frac{D\bar{e}_k}{Dt} = \frac{\partial \bar{e}_k}{\partial t} + \sum_{j=1}^3 \bar{u}_j \frac{\partial \bar{e}_k}{\partial x_j} \quad (14/d)$$

and

$$T_{tf} = \frac{D\bar{e}'_k}{Dt} + \sum_{j=1}^3 \bar{u}_j \frac{\partial \bar{e}'_k}{\partial x_j} = \frac{\partial \bar{e}'_k}{\partial t} + \sum_{j=1}^3 \left( \overline{u_j \frac{\partial e'_k}{\partial x_j}} + \bar{u}_j \frac{\partial \bar{e}'_k}{\partial x_j} \right) \tag{14/e}$$

Thus, the relative transport capacity of the flow reads

$$T_r = \left( 1 + \frac{T_{mf}}{T_{tf}} \right)^{-1} = \left[ 1 + \frac{\frac{\partial \bar{e}'_k}{\partial t} + \sum_{j=1}^3 \bar{u}_j \frac{\partial \bar{e}'_k}{\partial x_j}}{\frac{\partial \bar{e}'_k}{\partial t} + \sum_{j=1}^3 \left( \overline{u_j \frac{\partial e'_k}{\partial x_j}} + \bar{u}_j \frac{\partial \bar{e}'_k}{\partial x_j} \right)} \right]^{-1} \tag{14/f}$$

And the kinetic power of the constituents can be expressed also with the right-hand terms of the relevant relationships as follows:

$$T_{mf} = \sum_{j=1}^3 \left[ \bar{u}_i f_i - \frac{\bar{u}_i}{\rho} \frac{\partial (\bar{p} + \bar{p}_i)}{\partial x_i} + (\nu + \epsilon_m) \sum_{j=1}^3 \bar{u}_i \frac{\partial^2 \bar{u}_i}{\partial x_j^2} \right] \tag{14/g}$$

and

$$T_{tf} = \sum_{j=1}^3 \left\{ -\frac{\overline{u'_i \frac{\partial p'}{\partial x_i}}}{\rho} + \sum_{j=1}^3 \left[ \overline{\nu u'_i \frac{\partial^2 u'_i}{\partial x_j^2}} + \epsilon_m \frac{\partial \bar{u}_i}{\partial x_j} \left( \frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \right) \right] \right\} \tag{14/h}$$

Both Eq. (14/e) and Eq. (14/h) prove that the kinetic power of elementary water body of unit mass of the time-mean of turbulent fluctuation is not only depending on the quantities of fluctuation but also that of the main (mean) flow. In contrary, the relevant relationships of the main (mean) flow [Eq. (14/d) and Eq. (14/g)] depend (formally) only on the quantities of the main (mean) flow. But it must not be forgotten that  $\bar{p}_i$  and  $\epsilon_m$  refer to the Reynolds stresses which express the effect of turbulent fluctuation exerted on the main (mean) flow.

### 7. Conclusions

Applying a relevant, approximate analytical review on the equation system (equation of continuity and Navier–Stokes equation) of turbulent water flow the following conclusions may be summarized:

(i) The turbulent water flow may be divided into two imaginary constituents: time-mean equation of main flow and time-mean equation of turbulent fluctuation.



(ii) The time-mean equations of the constituents clearly indicate that a mutual interaction exists between the two imaginary parts of the turbulent water flow.

(iii) The  $i$ -th component of the Navier–Stokes equation (expressed in Cartesian coordinates with Eulerian description form) multiplied by the relevant  $i$ -th component of the velocity vector results in the  $i$ -th component of the kinetic power equation. Summing up these equations according to the coordinate axes it results in the kinetic power of elementary water body of unit mass. The kinetic power of the main (mean) flow consists of the well-known local and convective terms [Eq. (14/d)] while that of the mean flow of turbulent fluctuation has an extraordinary convective term, too, which expresses a kinetic energy (of turbulent fluctuation) convection by the main (mean) flow velocity [Eq. (14/e)].

(iv) The kinetic power of the constituents can be expressed by the relevant fluid and flow characteristics [ $T_{mf}$ , Eq. (14/g) for main (mean) flow;  $T_{ff}$ , Eq. (14/h) for turbulent (mean) fluctuation]. The relationship for main flow consists of terms by characteristics of the fluid (constants) and that of the main flow, except the quantities which refer to the Reynolds stresses ( $\bar{p}_i$  and  $\varepsilon_m$ ). The relationship referring to the turbulent fluctuation is similarly constructed (characteristics by the fluid and that of the turbulent fluctuation) except that kinetic power by the Reynolds stresses is expressed as deformation of the main flow. Thus, the mutual interaction between the constituents is formally demonstrated, too.

(v) Interesting to note that while the kinetic power of turbulent (mean) fluctuation shows a direct dependence on the main (mean) flow such a direct interaction from turbulent (mean) fluctuation to the main (mean) flow does not exist. This is to demonstrate the primary importance of the main flow (e.g. in laminar flow a main flow does exist but any turbulent fluctuations do not; by increasing main flow – after transition – also turbulent fluctuation comes into existence and this is created – later on maintained – by the main flow itself; the counter-effect is the increase in resistive forces in the main flow).

## 8. Closing remarks and future outlooks

Bernoulli, D. published the results of his research on mathematics and physics in 1738, among others on the dynamic equation system of an inviscid fluid jet bounded by streamlines. This relationship – called later as Bernoulli equation – expressed the constancy of energy of elementary fluid volume of unit weight viz. along its path its energy content did not change only the ratio(s) of the energy constituents (potential-, pressure- and kinetic energy) was (were) varying.

Euler, L. published the three-dimensional equation system of inviscid fluid flow in 1755. The dependent variables (the three components of the velocity vector and the pressure) are functions of the independent space and time coordinates. The equation system was closed because the fourth equation was the equation of continuity.

These theoretical results were for the practice of no use because of the unacceptable difference between the theoretical forecasts and the practical measurements; which was the consequence of the supposition of inviscidity. The researchers were misled by the coincidence between theoretical forecasts and practical measurements regarding hydrostatics; but molecular structure of the fluid resulted in internal and wall friction at motion and the turbulent behaviour of the flow produced further resistance; and without considering these resistive forces theory could not meet the real processes.

The first – semiempirical – result was already born in 1775. Chézy published namely his relationship between flow velocity and the hydraulic and geometrical characteristics of open channel flow. The effect of all resistive forces was taken into account by the empirical – so-called – *Chézy coefficient*. Nobody could overpass this achievement till today.

Navier, C. (1826) and Stokes, G. (1847) – similarly but independently to each other – completed Euler's equation system with terms expressing internal friction caused by molecular viscosity. This extended equation system was applicable in laminar flow and very far from any rigid (solid) boundary, therefore – in practical point of view – it has a very narrow band of usefulness (in hydraulic engineering).

Reynolds, O. was who recognized that the turbulent behaviour of flow evoked additional resistance in form of stresses; and Boussinesq, J. was who proposed to take into account the effect of this Reynolds stresses in a similar way as the effect of viscous stresses were taken into account (1877). The result of this way of thought was the introducing of additional terms of apparent viscosity and pressure ( $\varepsilon$  and  $\bar{p}_t$ ), respectively. These terms were the consequences of a hypothesis with not any physical meaning or content.

For more than the next one hundred years the basic research in physics of fluid flow or the fundamental research in applied fluid dynamics tried to solve the remaining problems, e.g.

- regarding wall friction (as one of the body forces) along the rigid (solid) boundary and
- to find any relevant and reliable calculation method for the Boussinesq's additional terms.

Neglecting some special attempts not too many publications could be found on wall friction which did not based on Chézy's theorem. It seemed that research could not exceed so-far Chézy's creative genius.

Many attempts could be found in the literature which proposed various solutions for calculating the Boussinesq's additional terms; but neither of them brought closer this hypothesis to practical applications (Rodi, 1984 and Rátky, 1995).

Thus – according to the author's opinion – this way of thoughts led to a deadlock and a fundamentally new research attempt must be found. Thus – in the author's modest view – the foregoing study (from Introduction to Conclusions) is a summary of the present knowledge which shows the hole in which science has dug itself and from which a fundamentally new way of thinking could lead to further development. The new way of thinking must probably taken into account *the molecular structure of the fluids* instead



of the *Eulerian continuum* or *Lagrangian substantial particle* (but this later one can probably lead in many cases farther than the *Eulerian continuum* concept). And this needs a new and young master mind in theoretical physics of fluid flow and experimentalists who are able to prove (or disprove) the theoretical results.

### References

1. EDF (Electricité de France, Laboratoire National d'Hydraulique): Rapport d'activité 1966, Chatou, 1997
2. Hinze, J. O.: Turbulence. McGraw-Hill Series in Mechanical Engineering, New York, etc. 1959
3. Hughes, W. F.–Gaylord, E. W.: Basic Equations of Engineering Science. Schaum's Outline Series. McGraw-Hill Book Company, New York, etc. 1964
4. Korn, G. A.–Korn, T. M.: Mathematical Handbook for Scientists and Engineers. Definitions, Theorem and Formulas for Reference and Review. McGraw-Hill Book Company, New York, etc. 1968
5. Rátky, I.: Mathematical basis of turbulent flow (in Hungarian). Vízügyi Közlemények (Hydraulic Engineering), Vol. LXXVII, 1995
6. Rodi, W.: Turbulence Models and their Application in Hydraulics. A State of the Arts Review. Karlsruhe, 1984



# APPLICATION OF A NEW OPTIMIZATION PROCEDURE FOR THE DESIGN OF A PROPORTIONAL WEIR

Keshava, M. K. and Shesha, P. M. N.

*Dept. of Civ. Engrg., Indian Inst. of Sci., Bangalore - 560 012, India*

(Received: 11 March, 1997)

A new general optimization technique to obtain the maximum linearity range in the head-discharge characteristics subject to a maximum permissible error, for flows through weirs which exhibit an extended *S* curve type theoretical head-discharge relationship is presented. It is shown that with this technique it is possible to fix the terminal points of linearity range exactly and thus obtain the maximum linearity range. This method incidentally is superior to the earlier methods used such as range of points or tangent point or graphical methods. The technique presented can be successfully used for the analysis of inverted V-notch or bell-mouth weir, etc. This procedure is adopted here for the analysis of the semi-circular weir (with the diameter as its crest). It is shown that for all flows through the weir in the range of  $0.38R \leq h \leq 1.19R$ , the discharges are proportional to the heads reckoned from a reference plane situated at  $0.16R$  above the crest. The usefulness of this weir with a simple profile and linear characteristics in minor irrigation and open drainage is highlighted. Experiments with two weirs are in good compliance with the theory by giving a constant average coefficient of discharge of 0.620-0.623.

## 1. Introduction

Of recent, there have been several studies on geometrically simple weirs exhibiting a near linear characteristics: a practical proportional weir [8], quadrant plate weirs [6], inverted V-notch, chimney weir, bell-mouth weir [1-3], constant accuracy practical linear weir, constant accuracy chimney weir [4, 5] are among the important ones which have been analysed. It has been shown that these weirs exhibit a near linear head-discharge relationships within certain ranges of head and within a prefixed permissible percentage of error. These weirs have their discharge-head relationship in the form of an extended *S* curve with a point of inflection. The near linear characteristics in the middle portion has been analysed by three procedures previously, viz. range of points, tangent points [1-3] and graphical method [4-5]. Although these procedures are adequate, they are basically search methods or trial and error methods, which not only give slightly varying results, but also, invariably consume a lot of CPU (computer processing unit) time on computers, which is highly uneconomical. Hence there was a necessity to evolve an exact procedure which would economise on CPU time and give better results. In what fol-

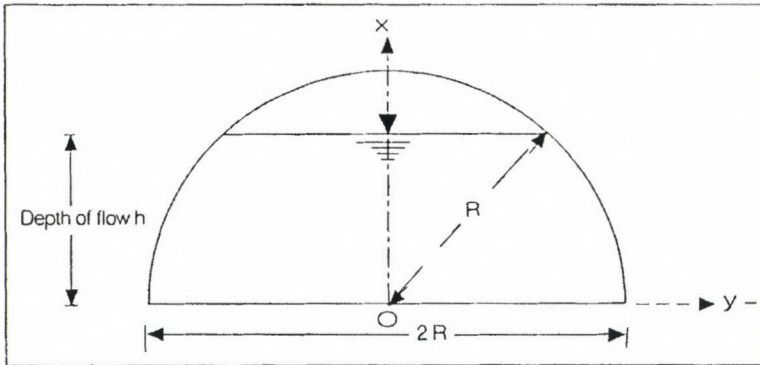


Fig. 1. Typical inverted semicircular weir

flows is given an algebraic procedure which has been applied to one of the examples – viz. flow through semicircular weir with diameter as its crest.

### 2. Discharge through the semicircular weir

Let us consider the discharge through the semicircular weir with a diameter ‘2R’ as its crest, shown in Fig. 1 given by,

$$q = 2C_d \sqrt{2g} \int_0^h \sqrt{R^2 - x^2} \sqrt{h - x} dx \quad 0 \leq h \leq R \quad (1)$$

where  $q$  = total discharge,  $h$  = head above the crest,  $g$  = acceleration due to gravity,  $C_d$  = coefficient of discharge.

For sharp crested weirs and streamed lined flows, the coefficient of discharge,  $C_d$  can be assumed to be a constant (which will be confirmed by experiments).

For convenience Eq. 1 can be expressed in nondimensional forms as

$$Q = \int_0^H \sqrt{H - X} \sqrt{1 - X^2} dH \quad 0 \leq H \leq 1 \quad (2)$$

where

$$Q = \frac{q}{KR^{\frac{5}{2}}}; \quad K = 2C_d \sqrt{2g}; \quad X = \frac{x}{R} \quad \text{and} \quad H = \frac{h}{R}.$$

The above integral has been evaluated by using standard IMSL routines, in CD4360 unix system. The theoretical head-discharge curve is shown in Fig. 2. It is seen from

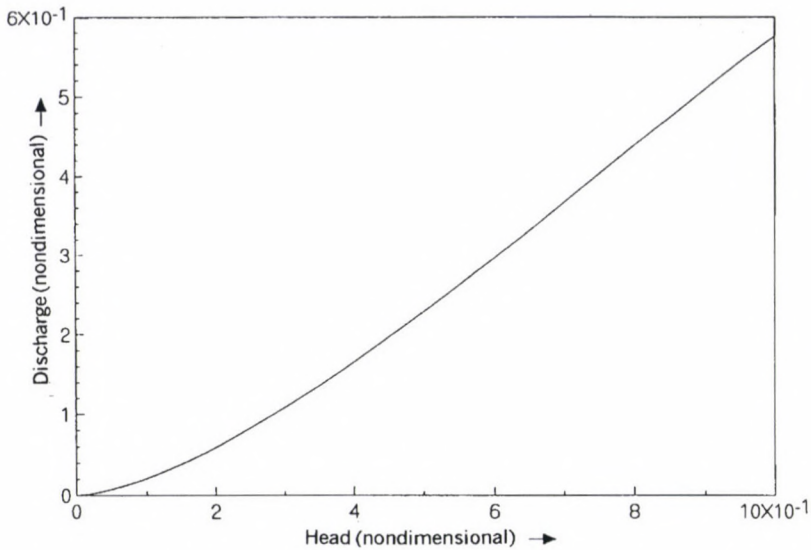


Fig. 2. Theoretical H-Q relationship showing near linear-relation in a certain range

Fig. 2 that the discharge-head relationship is near linear in a certain ranges of head beyond a certain minimum value. This property of the weir is exploited to evolve a linear relationship between  $Q$  and  $H$  in a certain range of head, such that the relative percentage deviation between theoretical discharge and the one given by linear relation does not exceed a prefixed maximum permissible error.

### 3. Procedure

Let

$$Q_L = mH + C \quad (3)$$

be the proposed optimal linear head-discharge relationship (where  $m$  is the constant of proportionality and  $C$  is the discharge intercept) to substitute the theoretical head-discharge relationship

$$Q = f(H) \quad (4)$$

in a certain range. Letting  $K_u = \left(1 + \frac{E}{100}\right)$  and  $K_d = \left(1 - \frac{E}{100}\right)$  where  $E$  is the prefixed maximum permissible relative deviation of the proposed linear function and the theoretical head-discharge function. These define two explicit curves  $f_1(H)$  and  $f_2(H)$



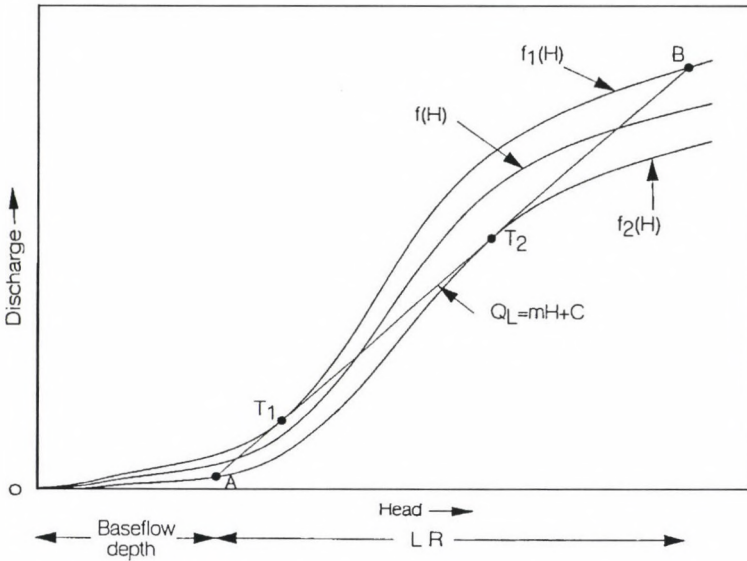


Fig. 3. Optimisation procedure

forming the lower and upper bounds for the linear function as shows in Fig. 3. Mathematically,

$$f_1(H) = K_u f(H) \tag{5a}$$

$$f_2(H) = K_d f(H) \tag{5b}$$

Step 1: Find the point of inflection of the curve  $Q = f(H)$  given by Eq. 4 i.e.

$$Q'' = f''(H) = 0$$

$$\therefore H = H_i \tag{6}$$

This procedure is valid only when there is one point of inflection and the function is continuously increasing in  $0 \leq H \leq \infty$  which is always true for any discharge-head function.

Step 2: Consider the straight line given by Eq. 3 to be tangential to the two binding curves  $f_1(H)$  and  $f_2(H)$  at the points  $T_1$  and  $T_2$ , respectively. Let it cut the curves  $f_2(H)$  and  $f_1(H)$  at the points  $A$  and  $B$  as shown in Fig. 3. Let us assume that with a slight variation of 'm', we get a straight line longer than the one given by Eq. 3 (AB), with in the binding limits.

$$Q_L = (m \pm \Delta m)H + C \tag{7}$$

But  $(m + \Delta m)$  shifts the line beyond the boundary  $f_1(H)$  at  $T_1$  and  $(m - \Delta m)$  for  $f_2(H)$  at  $T_2$ . Hence the values of  $m$  is the optimum value which yields the straight line of maximum length.

Now let us consider a small variation in 'C' as  $(C \pm \Delta C)$  with which we get a longer straight line than the one given by Eq. 3.

$$Q_L = mH + (C \pm \Delta C) \quad (8)$$

But  $(C + \Delta C)$  shifts the line beyond the boundary  $f_2(H)$  at  $T_2$  and  $T_2$  and  $(C - \Delta C)$  for  $f_1(H)$  at  $T_1$ . Hence the value of  $C$  is the optimum value along with  $m$  which yields the straight line of maximum length. Thus, the Eq. 3, gives us the straight line of maximum length and hence the maximum linearity range.

Step 3: The equation of the straight line passing through  $T_1$  and  $T_2$  is

$$\left[ \frac{Q_L - f_1(H_1)}{H - H_1} \right] = \left[ \frac{f_2(H_2) - f_1(H_1)}{H_2 - H_1} \right] \quad (9)$$

$$\therefore Q_L = \left[ \frac{f_2(H_2) - f_1(H_1)}{H_2 - H_1} \right] (H - H_1) + f_1(H_1) \quad (10)$$

or

$$Q_L = \left[ \frac{f_2(H_2) - f_1(H_1)}{H_2 - H_1} \right] H - \left[ \frac{f_2(H_2) - f_1(H_1)}{H_2 - H_1} \right] H_1 + f_1(H_1) \quad (11)$$

Comparing Eqs 3 and 11,

$$m = \left[ \frac{f_2(H_2) - f_1(H_1)}{H_2 - H_1} \right]$$

and

$$C = f_1(H_1) - H_1 \left[ \frac{f_2(H_2) - f_1(H_1)}{H_2 - H_1} \right]$$

or

$$C = f_1(H_1) - mH_1 \quad (12)$$

Equation 4 can be written in the implicit forms as

$$f(Q, H) = Q - f(H) \quad (13)$$

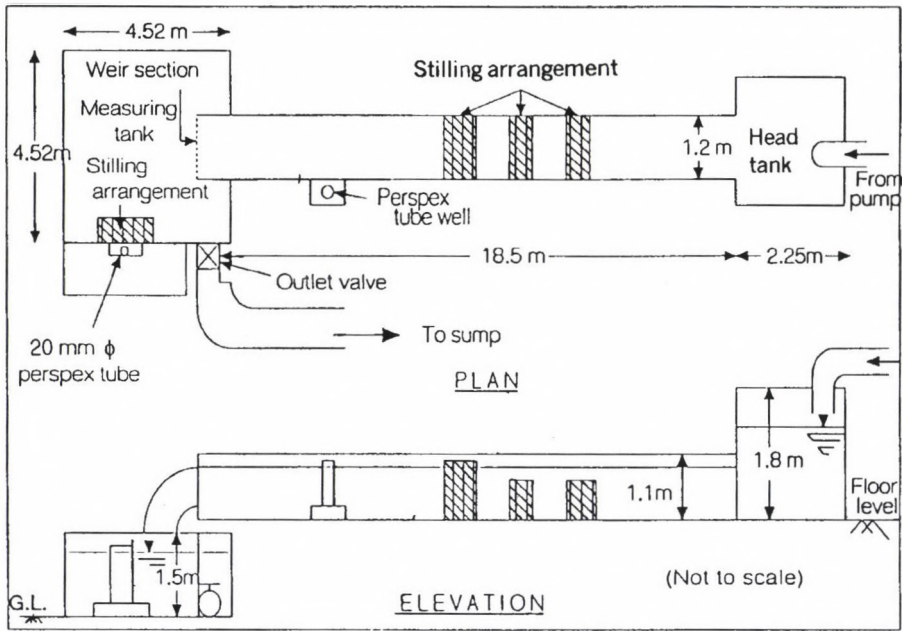


Fig. 4. Experimental setup

For any point on the curve  $f(H)$

$$f(Q, H) = 0 \tag{14}$$

Similarly,

$$f_1[H_1, f_1(H_1)] = 0 \tag{15}$$

or

$$f_2[H_2, f_2(H_2)] = 0 \tag{16}$$

$$\therefore f_1[H_1, f_1(H_1)] = f_2[H_2, f_2(H_2)] \tag{17}$$

further

$$f_1'[H_1, f_1(H_1)] = f_2'[H_2, f_2(H_2)] \tag{18}$$

From Eqs 17 and 18, we can solve for  $H_1$  and  $H_2$  and then from Eq. 12 we get  $m$  and  $C$ . The fact that the Eqs 15 and 16 have the real roots, proves that there exists a common tangent between  $f_1(H)$  and  $f_2(H)$  curves.

Step 4: now considering the extreme points,  $A$  and  $B$ .



$$(Q_L)_A = mH_A + C = K_d f(H_A) \quad (19)$$

and

$$(Q_L)_B = mH_B + C = K_u f(H_B) \quad (20)$$

From Eqs 19 and 20 we get  $H_A$  and  $H_B$  and

$$LR = H_A - H_B \quad (21)$$

where  $H_A$  is the baseflow depth or the minimum depth above which discharge has a near linear relationship with head and  $H_B$  is the terminal point of the linearity range ( $LR$ ).

Expressing  $f(H)$  as

$$Q = f(H) = K_1 + K_2H + K_3H^2 + K_4H^3 \quad (22)$$

a third degree polynomial,  $K_1$ ,  $K_2$ ,  $K_3$  and  $K_4$  are evaluated by regression analysis as  $K_1 = -0.0032833886$ ,  $K_2 = 0.1831059819$ ,  $K_3 = 0.726040637$  and  $K_4 = -0.3301597981$ . The point of inflection  $H_i = 0.733$  and with  $E = 1.5$ ,  $T_1 = 0.50428$ ,  $T_2 = 0.91168$ ,  $H_A = 0.376$ ,  $H_B = 1.190$ ,  $m = 0.673437$ , and  $C = -0.104785$ .

Hence the proposed linear discharge-head relationship to replace the theoretical one given by Eq. 20 is

$$Q_L = (0.6734H - 0.1048) \quad 0.376 \leq H \leq 1.190 \quad (23a)$$

or

$$Q_L = 0.67(H - 0.16) \quad 0.38 \leq H \leq 1.19 \quad (23b)$$

or dimensionally

$$q_L = 0.67KR^{\frac{3}{2}}(h - 0.16R) \quad 0.38R \leq H \leq 1.19R \quad (23c)$$

#### 4. Experiments

The weirs chosen for the experiment have the radii  $R = 1$  cm and  $R = 30$  cm. The weirs were cut from 6.5 mm M.S. plate accurately using a nibbling machine and edges were chamfered to 45%. The weir was placed at the end of the channel as shown in the experimental setup (Fig. 5). An electronic point gauge with a least count of 0.01 mm was used at 4 m upstream of the weir section to measure the flow depth. The time required to collect a fixed volume of water in the measuring tank (4.52 m × 4.52 m × 1.5 m) was computed from an electronic timer triggered automatically by signals from an electronic switch attached to the level indicators. Figure 6 shows the nappe of the discharging weir.

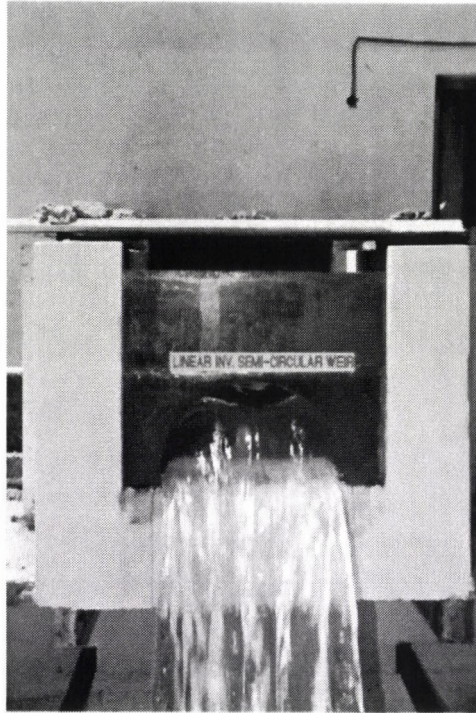


Fig. 5

### *Analysis of results*

Figure 7 shows the plot of  $Q_{\text{actual}}$  vs head measured from the datum situated at  $0.1556 R$  above the crest. The straight line portion of the plot beyond a certain head, confirms the theory. It is also seen that the coefficient of discharge  $C_d$ , for any head within the fixed range does not vary by more than  $\pm 1\%$  of the average coefficient of discharge (0.620–0.623) which supports the assumption of constant  $C_d$  in our analysis.

### **5. Concluding remarks**

A new general rational algebraic optimization procedure is presented to obtain the maximum linearity range for flow through weirs exhibiting an extended  $S$  curve type of head-discharge relationship, subject to a maximum permissible error. It is shown that with this technique it is possible to fix the extreme points of linearity range exactly and thus obtain the maximum linearity range. It is further shown that this procedure is superior to the earlier methods used such as range of points or tangent point or graphical method, which are essentially search methods or a trial and error procedures, which not only give varying results but also consume large CPU time on computers and hence are

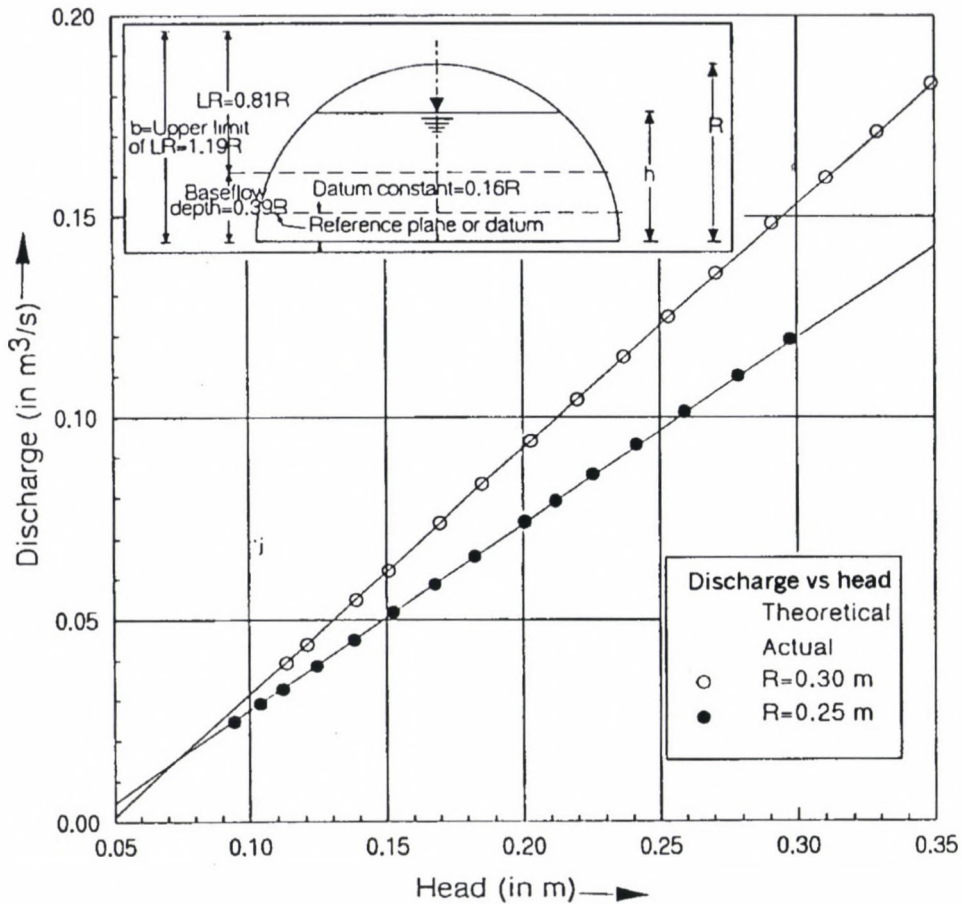


Fig. 6. Discharge-head variation (experimental)

uneconomical. As an example the technique is applied to analyse the flow through a semicircular weir to show its effectiveness. It is shown that the discharge through the semicircular weir with a diameter of  $2R$  as its crest, is proportional to the linear power of head measured from a reference plane located at  $0.156R$  above the crest for all flows in the range of  $0.376R \leq h \leq 1.190R$  within a prefixed maximum permissible deviation of  $\pm 1.5\%$  from the theoretical discharge. The ratio  $Q_{\max}/Q_{\min}$  which is the measure of discharging capacity of the weir is 4.7. Experiments with two weirs show a constant average coefficient of discharge varying from 0.620–0.623. The presented technique can also be successfully used for the analysis of inverted V-notch or bell-mouth weir, etc. It is hoped that the weir presented here of conventional geometrical shape with simple linear head-discharge characteristics would be practically useful as a simple flow recorder in minor irrigation, hydraulic engineering and open drainage.



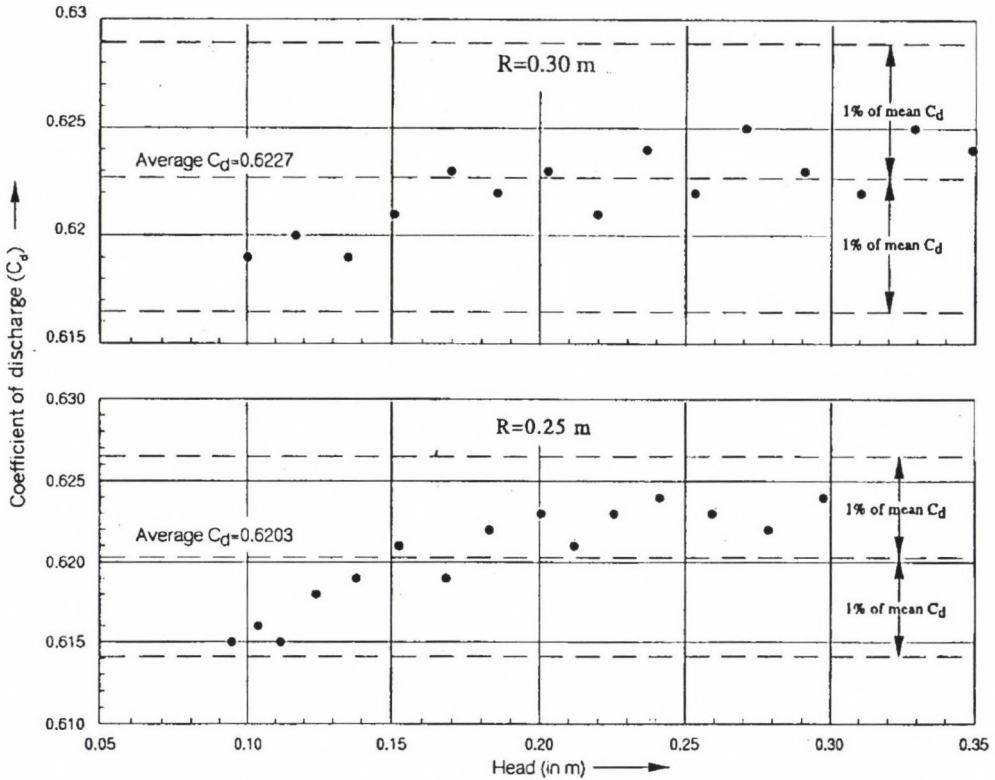


Fig. 7. Variation of coefficient of discharge with head

### Acknowledgements

The authors are grateful to the authorities of the Indian Institute of Science, for providing the necessary facilities for conducting this work. M. N. Shesha Prakash, Asst. Prof., J. N. N. College of Engineering, is indebted to the authorities of the National Education Society, Shimoga for granting permission to work on deputation at I.I.Sc., Bangalore.

### References

1. Keshava Murthy, K.-Giridhar, D. P.: Inverted V-notch: A practical proportional weir. J. Irrig. and Drain. Engrg. Div., ASCE, 115 (1989), 1035-1050
2. Keshava Murthy, K.-Giridhar, D. P.: Improved inverted V-notch or chimney weir. J. Irrig. and Drain. Engrg. Div., ASCE, 116 (1990), 374-386
3. Keshava Murthy, K.-Giridhar, D. P.: Geometrically simple linear weirs using circular quadrants: bell mouth weirs. J. Hydr. Res., IAHR 29 (1991), 497-508

4. Keshava Murthy, K.–Shesha Prakash, M. N.: Practical constant accuracy linear weir. J. Irri. and Dra. Engrg. Div., ASCA, 120 (1994), 526–538
5. Keshava Murthy, K.–Shesha Prakash, M. N.: Constant accuracy chimney weir. Accepted for presentation at IX. Congress of Asian Pacific Division of IAHR to be held at Singapore during Aug. 24–26, 1994
6. Ramamurthy, A. S.–Subramanya, K.–Pani, B. S.: Quadrant-Plate weirs. J. Hydr. Div., Proc., ASCE, 103 (1977), 1431–1441
7. Troskolanski, A. T.: Hydrometry: Theory and practice of hydraulic measurements. Pergamon Press, New York (1960), 301–302
8. Venkataraman, P.–Subramanya, K.: A practical proportional weir. Water Power 25 (1973), 189–190





# NUMERICAL EXAMINATION OF THE STIFFENING SYSTEM OF TALL BUILDINGS BY USING OBJECT-ORIENTED PROGRAMMING

Klopka, Z.\*—Ládi, N.\*\*

*\*Technical University of Budapest, Department of Reinforced Concrete Structures,  
Bertalan Lajos u. 2, H-1521 Budapest, Hungary*

*\*\*Zepter International Ungarn, Asztalos S. u. 9/12, H-1087 Budapest, Hungary*

(Received: 17 April 1996)

In order to design efficiently the shear walls of tall buildings their influences are to be determined accurately. This paper shows a numerical method of shear wall analysis written in an object-oriented environment. Numerical examinations have been carried out to emphasize the importance of the own warping rigidity of wall-groups.

## 1. Introduction

Architectural solutions and modern construction methods presently used with tall buildings require slimmer, more efficient structures, therefore various elements of the structure are more sharply differentiated by their roles; their dimensions are also approaching the minimal sizes required by the laws of force. We must define loads and influences more precisely when designing these structural elements, therefore more accurate, computerized procedures involving a greater amount of calculations are replacing traditional, manual methods.

The load bearing structure of tall buildings has a twofold role: the task of slabs is to take and pass over vertical loads to the vertical load-bearing structures, while shear walls take over and pass horizontal influences and vertical loads to the foundation. Forces distributed to certain shear walls and groups of shear walls ("cores") may be hard to determine by intuition. The traditional calculation method does not derive the accurate result due to complex shear-wall arrangements on one side and because of some neglected rigidities on the other.

In this study we will present a modern computer application in which horizontal loads acting on the building are distributed over shear walls; the method is not only increasing the accuracy of the calculation model but also provides a user-friendly, graphics-oriented programming environment for the planning engineer.

## 2. Problem description

The stiffening model of tall buildings has been an interesting topic of many studies, the principle solution of the problem can be found in works listed in the bibliography section under numbers [2, 3] and [5]. The procedure presented here is also based on concentrating the shear walls of a building into an equivalent stiffening core. The strength of the core is calculated from the appropriate rigidities of walls and groups of walls (hereinafter also “wall-groups”). Horizontal forces acting on the building are affecting the cantilever restrained in foundation at bottom level – the stiffening core. By calculating the deformation of the cantilever we can define forces acting on individual walls – as a function of the rigidities of these walls.

Within the model we have to account for the shear, flexural and torsional rigidity of walls based on the flexibility theory of thin-walled, open-sectioned beams. Closed sections cannot be used in calculations and the model is limited to stable walls which do not change their longitudinal cross-section (Fig. 2.1).

The basic assumption of the method used in the model is that slabs – as infinitely rigid in their plane, and infinitely flexible plates perpendicular to their middle surfaces – are able to secure the structure’s geometrical properties concerning its height.

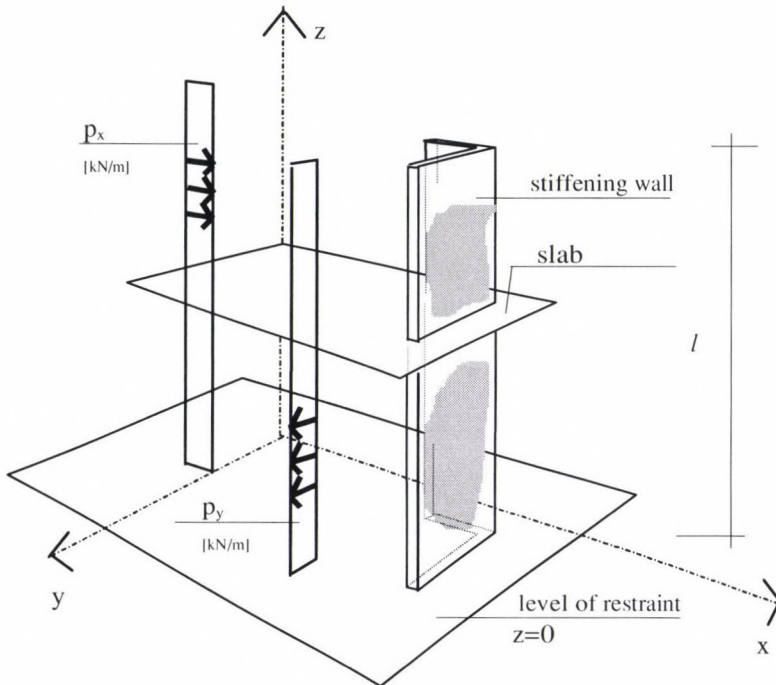


Fig. 2.1. The view of the model

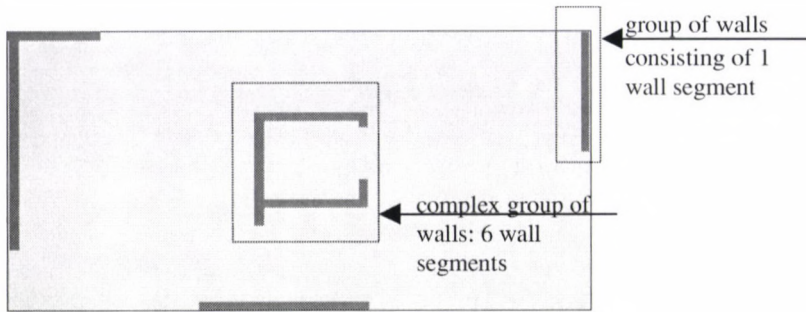


Fig. 2.2. Examples of groups of shear walls

We have differentiated between walls and wall-groups in the developed computer program. Walls are wall segments with a straight axis, of constant width, and having two characteristic joints (start and end joints). All walls which do not deform independently of each other due to their connection compose a group of walls (Fig. 2.2).

### 3. Differential equations of thin-walled, open-sectioned beams

The characteristic of thin-walled, open-sectioned beams is that one dimension of their cross-section – width – is significantly smaller than the other dimension. At the same time, the dimensions of the cross-section are small when compared to beam length. The joints of a general beam in 3-D space have a degree of freedom of seven, as compared to solid beams with a degree of freedom of six. With solid beams the “missing” degree of freedom is actually describing the warping effect. The warping effect means that due to the influence of a given torsional moment, particles of the beam’s cross-section do not suffer torsion only (simple, free or St. Venant type torsion) but they also deplane, i.e. displace in the direction of the beam’s axis. To formulate differently, thin-walled sections are resisting the torsion not only with their torsional rigidity but also with flexural rigidity. As a result of torsion, stress in the direction of the axis (normal stress) is generated.

Thin-walled sections may be open or closed. Closed sections are “rigid to torsion”, i.e. only influences of simple torsion are dominant. The simple torsional rigidity of open sections is so small that it may be neglected when compared to warping rigidity.

The differential equations of a thin-walled, open-sectioned cantilever fixed on one end with loads uniformly distributed along and perpendicular to its  $z$  axis are as follows [4]:

$$EFw'' = 0 \quad (3.1)$$

$$EI_y u'''' = p_x \quad (3.2)$$



$$EI_x v'''' = p_y \quad (3.3)$$

$$EI_{\Omega\Omega} \varphi'''' - GI_t \varphi'' = m_D \quad (3.4)$$

where  $x, y$  are the principal axes,

$E$  is the flexibility modulus,

$G$  is the shear flexibility modulus,

$F$  is the area of the beam's cross-section,

$I_x, I_y$  are moments of inertia about the  $x$  and  $y$  axis of the section, respectively

$I_{\Omega\Omega}, I_t$  are warping and free torsional moments of inertia of the beam, respectively,

$u, v, w$  are displacements of the section's cross-section area in directions  $x, y$  and  $z$ , respectively,

$\varphi$  is the rotation of the cross-section about axis  $z$ , the positive direction is clockwise in a left-handed coordinate system,

$p_x, p_y$  are distributed forces acting on the section in directions  $x$  and  $y$ , respectively,

$m_D$  is the distributed moment of torsion acting on the section, and

$( )'', ( )''''$  indicate the second and fourth derivatives with respect to  $z$ .

Differential equation (3.1) describes influences in the direction of the axis, equations (3.2) and (3.3) describe inclined bending and equation (3.4) describes torsion. The axial compression/tension and inclined bending have been discussed in numerous studies, several computerized procedures are available, therefore we did not want to work out these two cases in more detail (for further data see [3, 4]).

Equation (3.4) is known as the differential equation of warping torsion. Boundary conditions of the equation are the following:

$$z = 0 \quad \varphi = 0, \varphi' = 0 \quad (\text{rotation and warping are restrained for } z = 0),$$

$$z = l \quad T = 0, W = 0 \quad (\text{moment of torsion and bimoment are zero on the nonfixed end of the cantilever})$$

We solve equation (3.4) by taking into consideration boundary conditions according to [1]:

$$\varphi(z) = -\frac{m_D}{GI_t \operatorname{ch}(k)} \left[ -\frac{l^2}{k^2} - \frac{l^2}{k} \operatorname{sh}(k) + z \left( l - \frac{z}{2} \right) \operatorname{ch}(k) + \frac{l^2}{k^2} \operatorname{ch} \left( \frac{k}{l} z \right) + \frac{l^2}{k} \operatorname{sh} \left( \frac{k}{l} (l - z) \right) \right] \quad (3.5)$$

where  $k = l \sqrt{\frac{GI_t}{EI_{\Omega\Omega}}}$  is the Timoshenko parameter and

$l$  is the length of the beam.

The layout of shear walls in tall buildings usually follows open sections. The simple torsional rigidity of open wall-groups may be neglected when compared to warping rigidity, therefore  $k \approx 0$ . Substituting this and analyzing the cross-section at the level of restraint ( $z = 0$ ), the fourth derivative of the rotation can be obtained:

$$\varphi'''' = (z = 0) = \frac{m_D}{EI_{\Omega\Omega}} \quad (3.6)$$

By using equations (3.2), (3.3) and (3.6) we can determine forces acting on wall-groups.

#### 4. Examining the equivalent core

In chapter 4 we will briefly survey the more important interdependencies of the torsion of thin-walled open sections.

##### 4.1. Shear center of the group of walls

After having defined the center of the group of walls, we calculate the moment of inertia with respect to axes  $x$ ,  $y$  ( $I_x$ ,  $I_y$ ) and the centrifugal moment of inertia ( $I_{xy}$ ). During the next step we have to define the sectorial coordinates of the walls' endpoints by choosing any point  $P$  for the pole and any point as the zero point. Sectorial coordinate  $\omega_P$  is the characteristic of points in the cross-section, similarly to the coordinates. We calculate it by using the following formula:

$$\omega_P = \int_S h_P ds \quad (4.1)$$

where  $S$  is the mid-line of the cross-section,  
 $h_P$  is the perpendicular distance of the section from the pole, and  
 index  $P$  indicates the pole.

Since we have calculated the sectorial coordinates for all characteristic points (joints), we define the sectorial centrifugal moment of the cross-section from the following formulae:

$$I_{x\omega_P} = \int_F y \omega_P dF, \quad (4.2)$$

$$I_{y\omega_P} = \int_F x \omega_P dF. \quad (4.3)$$

We obtain the shear center of the cross-section from the formulae given below:

$$x_C = x_P - \frac{I_{x\omega_P}}{I_x}, \quad (4.4)$$

and

$$y_C = y_P - \frac{I_{y\omega_P}}{I_y}. \quad (4.5)$$

By taking the shear center of the group of walls as the pole we calculate the sectorial coordinates of the characteristic points – we indicated these as  $\omega_P$ .

When deriving differential equations (3.1) and (3.4) (see [2, 4]) we have assumed that the sectorial static moment of the cross-section is equal to zero:

$$S_{\Omega} = \int_F \Omega dF = 0, \quad (4.6)$$

where  $\Omega$  are sectorial coordinates for the shear center – their integral value with respect to the surface is zero.

This condition is still not satisfied in our case since the zero point – from which we started measuring our sectorial coordinates – has been arbitrarily chosen. Therefore, we have to find a new zero point for which equation (4.6) holds true, i.e. sectorial coordinates have to be standardized:

$$\Omega = \omega_C + \omega_0, \quad (4.7)$$

where

$$\omega_0 = -\frac{S_{\omega}}{F}, \quad (4.8)$$

and

$$S_{\omega} = \int_F \omega_C dF. \quad (4.9)$$

Afterwards only the group's sectorial moment of inertia should be calculated.

$$I_{\Omega\Omega} = \int_F \Omega^2 dF. \quad (4.10)$$

#### 4.2. Shear center of the equivalent core

Inertia of wall-groups ( $I_x$ ,  $I_y$ ,  $I_{xy}$ ) as well as their distance from each other influences the position of the level's shear center. Derivation can be found in [5] – we will only print the formulae here:



$$x_C = \frac{I_{xy}(\sum I_{y,i} y_{C,i} - \sum I_{xy,i} x_{C,i}) - I_y(\sum I_{xy,i} y_{C,i} - \sum I_{x,i} x_{C,i})}{I_x I_y - I_{xy}^2} \tag{4.11}$$

$$y_C = \frac{I_x(\sum I_{y,i} y_{C,i} - \sum I_{xy,i} x_{C,i}) - I_{xy}(\sum I_{xy,i} y_{C,i} - \sum I_{x,y} x_{C,i})}{I_x I_y - I_{xy}^2} \tag{4.12}$$

where  $x_{C,i}, y_{C,i}$  is the distance of the shear center of the  $i^{th}$  group of walls from the origin (see Fig. 4.1),  $\Sigma$  is summing up with respect to  $i$  (by wall-groups).

4.3. Rigidity of the equivalent core

The flexural or centrifugal inertia of the core equals the sum of inertia of wall-groups:

$$\begin{aligned} I_{xx} &= \sum I_{x,i} , \\ I_{yy} &= \sum I_{y,i} , \\ I_{xy} &= \sum I_{xy,i} . \end{aligned} \tag{4.13}$$

The torsional inertia of the core is equal to zero since the simple torsional rigidity of wall-groups has been neglected. We calculate the sectorial inertia of the core (warping rigidity) by using the following formula:

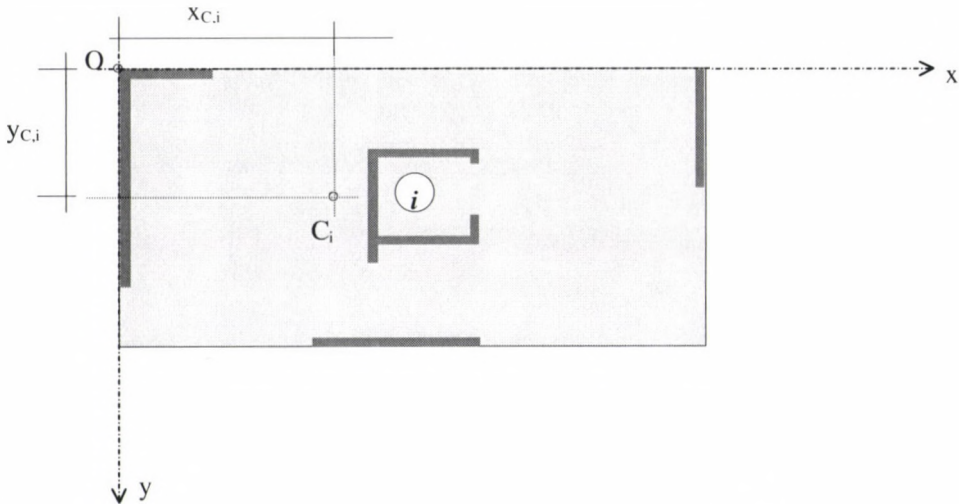


Fig. 4.1. The graphical interpretation of distances between origin and the wall-group's shear center

$$I_{\Omega\Omega} = \Sigma I_{\Omega,i} + \Sigma I_{x,i} (\bar{x}_{c,i})^2 - \Sigma I_{xy,i} \bar{x}_{c,i} \bar{y}_{c,i} + \Sigma I_{y,i} (\bar{y}_{c,i})^2, \quad (4.14)$$

where  $\bar{x}_{c,i}$ ,  $\bar{y}_{c,i}$  is the distance between the equivalent core and the shear centers of the  $i^{\text{th}}$  group of walls (see Fig. 4.2),  $\Sigma$  is summing up with respect to  $i$  (by wall-groups).

The first element of formula (4.14) is the sum of sectorial inertia and the other components have their origin in the impact of flexural rigidity used for warping the group of walls (compare to the Steiner's component of flexural inertia). Thus we have calculated figures necessary for solving the differential equation of the thin-walled, open-sectioned beams.

#### 4.4. Forces distributed over wall-groups

Let us place the coordinate system into the shear center. Since the  $x$  and  $y$  are no longer principal axes, equations (3.1) and (3.2) are modified in the following way:

$$\begin{aligned} EI_{yy}u_C'''' + EI_{xy}v_C'''' &= p_x \\ EI_{xx}v_C'''' + EI_{xy}u_C'''' &= p_y. \end{aligned} \quad (4.15)$$

where  $u_C$  and  $v_C$  are displacement components of the equivalent core's shear center.

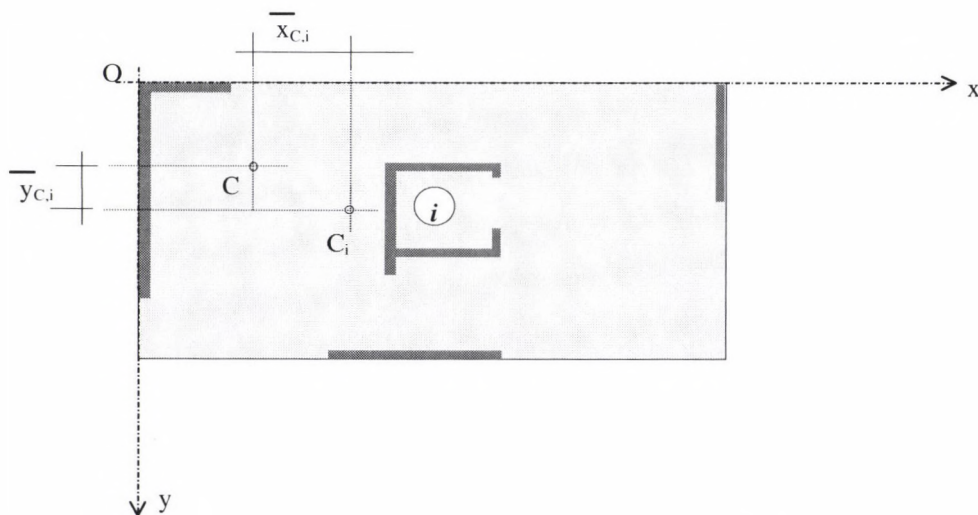


Fig. 4.2. The interpretation of distances between the equivalent core's and the wall-group's shear center

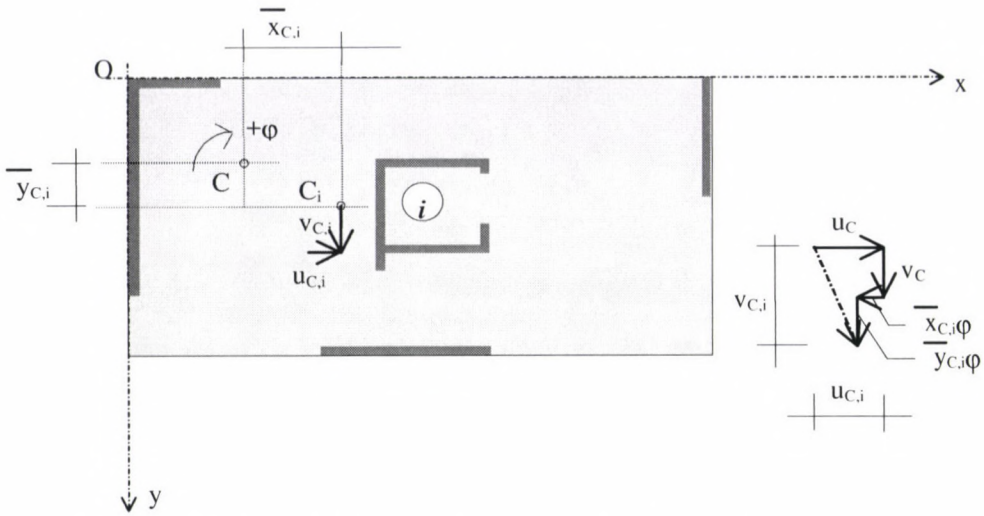


Fig. 4.3. The interpretation of the shear wall-group's displacement

We can express the fourth derivatives of displacements from the two equations shown above:

$$Eu_C'''' = \frac{I_{yy}P_x - I_{xy}P_y}{I_{xx}I_{yy} - I_{xy}^2}, \tag{4.16}$$

and

$$Ev_C'''' = \frac{I_{xx}P_y - I_{xy}P_x}{I_{xx}I_{yy} - I_{xy}^2}. \tag{4.17}$$

We can obtain the fourth derivative of the core's rotation from equation (3.6):

$$E\phi'''' = \frac{m_D}{I_{\Omega\Omega}}. \tag{4.18}$$

In formulating the problem we assumed that horizontal forces acting on the building are constant, i.e. their intensity does not change along the height of the building. Based on this we can conclude that the fourth derivative of the equivalent core's displacement or rotation also remains unchanged along the height.

The following are translatory displacements of the wall-groups indicated by *i* (see Fig. 4.3):

$$\begin{aligned} u_i &= u_C - \phi \bar{y}_{C,i} \\ v_i &= v_C - \phi \bar{x}_{C,i} \end{aligned} \tag{4.19}$$



By substituting the fourth derivatives of the above given statements into equations (4.16) and (4.17) we can determine the forces in the shear centers of wall-groups (for more details see [2]):

$$p_{x,i} = EI_{yy}(u_C'''' - \varphi'''' \bar{y}_{C,i}) + EI_{xy}(v_C'''' + \varphi'''' \bar{x}_{C,i}), \quad (4.20)$$

$$p_{y,i} = EI_{xy}(u_C'''' - \varphi'''' \bar{y}_{C,i}) + EI_{xx}(v_C'''' + \varphi'''' \bar{x}_{C,i}), \quad (4.21)$$

In equations (4.20) and (4.21) we can differentiate between forces acting on the group of walls: forces which appear due to the flexural rigidity of the group of walls and forces which result from the warping rigidity of the group of walls. We transform equations (4.20) and (4.21):

$$p_{x,i} = EI_{yy}u_C'''' + EI_{xy}v_C'''' - EI_{yy}\varphi''''\bar{y}_{C,i} + EI_{xy}\varphi''''\bar{x}_{C,i}, \quad (4.22)$$

$$p_{y,i} = EI_{xy}u_C'''' + EI_{xx}v_C'''' - EI_{xy}\varphi''''\bar{y}_{C,i} + EI_{xx}\varphi''''\bar{x}_{C,i}, \quad (4.23)$$

In the above shown formulae the first two elements result from flexure and the second two elements represent components resulting from warping.

## 5. Object-oriented programming applied in the calculation of influences of shear walls [7]

### 5.1. Object-oriented approach

Object-oriented programs represent the leading software methodology of the '90s. They provide a higher level of abstraction, which allows us to write programs that are much easier to read, debug, correct and maintain. Data types built in such programs are called *classes* with *objects* as class instances.

Our guideline was to build the shear wall program in an object-oriented manner. Entities arising in the engineering calculation were formulated as classes. As a result, we have developed the following classes: Joint class, which describes a joint ( $x$ ,  $y$ , sectorial coordinate), Wall class describes straight wall sections (starting, ending joint, thickness, centroid, area), WallGroup class describes walls built together (centroid, shear center).

As input data for the calculation we have a list of walls, where each wall is related to the coordinates of its ends and the thickness. Walls are grouped in a function where the Joint and WallGroup arrays are established. This function determines which walls are connected and also extends to those special cases in which two walls intersect (a new joint and two more walls are created) or in which two walls lay on the same line and have a mutual joint (they are merged); the function also labels internally the wall-groups. This function is called whenever a change to a wall occurs such as wall translation, deletion, etc.

The main part of the program is the part with the actual engineering calculation. Common functions for different classes are defined within each class independently, but in the main program their call is the same, only arguments are different. For instance, functions that calculate the coordinates of the centroid are defined both in the Wall and WallGroup classes having the same name, and in the main program, only the function name is called. The entity characteristics related to a specific class are being calculated within that specific class, and what is even more important, they can be accessed only through the functions of this class. Building the code in this manner, the source is not a single procedure including a lot of lines in which every variable has to be calculated one after the other, but a union of independent procedures.

Classes also include functions that handle the mouse and control the graphical output. The Joint, Wall and WallGroup classes contain the Draw and DrawPr functions, which draw to the screen or to the printer joints, walls, wall-groups including their labels, centroids, shear centers, etc. in a chosen enlargement and color. The MouseNear function returns the entity characteristics if the cursor nears a joint or wall, and if one of the entities (joint or wall) is double-clicked, it will be selected.

The fact that classes can be accessed only through functions ensures that class variables can be reached in a controlled and foreseen way. As a result, classes already constructed and procedures already tested become a closed structure, and in a further program development their behavior can not be interfered with by chance (unlike the classical programming technique, where public variables and arrays can be accessed from anywhere within the program).

### *5.2. The Microsoft Windows™ as program environment\**

When we started the development of the program, we had the intention of creating a user-friendly environment as well as facilitating engineering calculations. A nowadays commonly used operating system, the Microsoft Windows, offers us a wide variety of tools such as graphical windows, menus, mouse handling, graphic printout, standard dialogues, etc. Programs written under Windows may run independently from the amount of RAM and hardware available (various kinds of monitors, printers); precisely, the operating system and not the programmer should take care of compatibility problems.

Our program is written in Borland C++™, using the zApp™ 2.2 as graphical interface. The program is multilingual, enables the zoom control, setting the colors of entities, and printing of numerical and graphical outputs.

The Windows has introduced to PC users a graphical environment, which tends to become a standard. The written program follows this guideline: the program window contains the menu, the toolbar and the central window.

\*Microsoft, Microsoft Windows, Microsoft Word are registered trademarks of the Microsoft Corporation, Borland C++ is a registered trademark of Borland International, Inc., zApp is a registered trademark of Inmark Development Corporation.



Besides using hot-keys, commands can be accessed through menus and/or through toolbars. The menu commands are organized into groups: *main* – connected to file and printing functions, *project* – functions associated with calculations, *options* – related to the screen display, and the usual *help* group. Commonly used functions, like file opening, saving, printing functions are identical to other Windows applications (like Notepad or even Word™), so the user is not burdened with additional help information concerning these commands.

The central window contains the floor plan, on top and on the left are rulers which give us an impression of the zoom magnitude. Since the real drawing can be much larger than the virtual one drawn on the screen, scrollbars define the portion of the real drawing depicted in the central window. The status bar is placed at the lower part of the window. This line contains some useful information, like on-line help, the joint or wall nearest to the mouse, the current grid and the actual mouse coordinates.

It is interesting to mention that at the time the program is being developed, the display area (screen resolution, window size, printing size) is not known. This means that general functions, which determine display parameters at every run, are to be used. For every entity we have to determine whether the entity is within the display size – in which case it really has to be drawn – or whether it is placed somewhere outside.

The program is multilingual, i.e. menus, messages, data and results can be written out in four languages. The language, entity colors chosen and other settings are written in a start-up file, which is read at every program start.

The Windows operational system operates in *preemptive multitasking*, which means that there has to be a co-operation between the programs in order for them to run simultaneously without error. The running time of every function is of crucial importance, because all other tasks have to be paused for that time. While the program runs, it receives and sends information via the operating system, and upon receiving information it decides about further steps. Such messages are, for instance, pressing of a key, moving or clicking of the mouse, closing or pulling a window “to back” or changing of a field content in a dialogue box. Thanks to this information exchange system, the program gains clarity, since it contains event handling functions rather than one endless loop. Every object such as the menu, the toolbar or the main window utilizes its own message-analyzing function. The mouse click message is always transferred by the operation system to the actual object under the mouse cursor.

An example of message-sending or message-receiving is the way in which screens are rendered. When screen repainting is wanted (after data input or calculations), the program sends a message to the operational system stating that some of the main window’s area is dirty, i.e., some particular window area should be redrawn. Windows then calls the drawing functions of objects currently placed on the affected area and these objects are redrawn.

Summarizing, our experience is that the programming of an object-orientated Window application is paying off – after the difficult first steps have been made – in easy readability and maintenance of the program code on one side, and user-friendliness on the other.



## 6. Sample problems for examining the effect of warping rigidity

### Problem 6.1

In this problem we will present steps applied in the program for determining forces affecting wall-groups.

Determine the forces arising in each shear wall shown in Fig. 6.1.

The detailed calculations of warping characteristics for group of walls No. 2 is shown below.

Let us assume that we have calculated the centroid of the group, flexural inertias, principal axes and the angle between axis (1) and axis  $x$ .

$$x_s = 1.265 \text{ m}$$

$$y_s = 1.38 \text{ m}$$

$$I_x = 2.5491 \text{ m}^4$$

$$I_y = 2.5598 \text{ m}^4$$

$$I_{xy} = -0.0876 \text{ m}^4$$

$$\alpha = -43.252^\circ$$

$$I_1 = 2.642 \text{ m}^4$$

$$I_2 = 2.467 \text{ m}^4$$

$$A = 2.28 \text{ m}^2$$

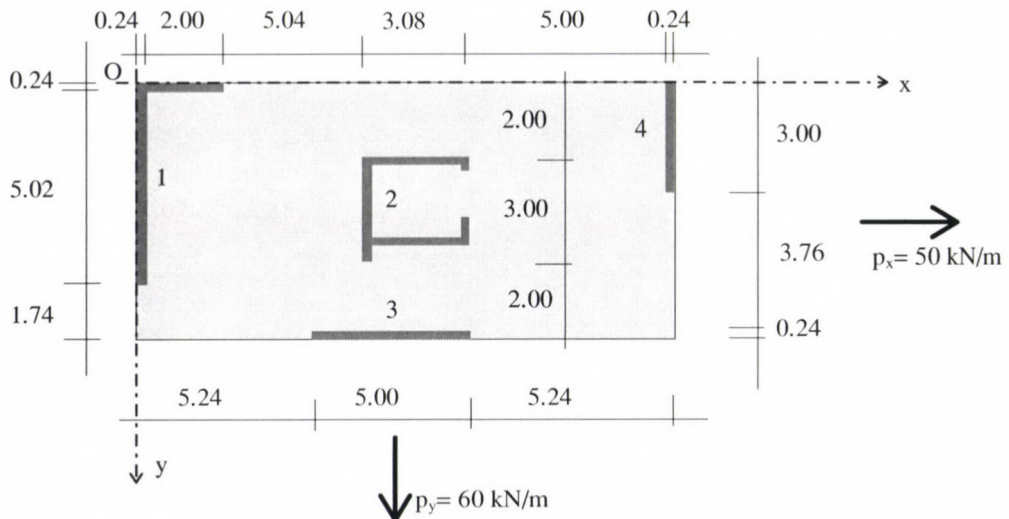


Fig. 6.1. Arrangement of the shear walls

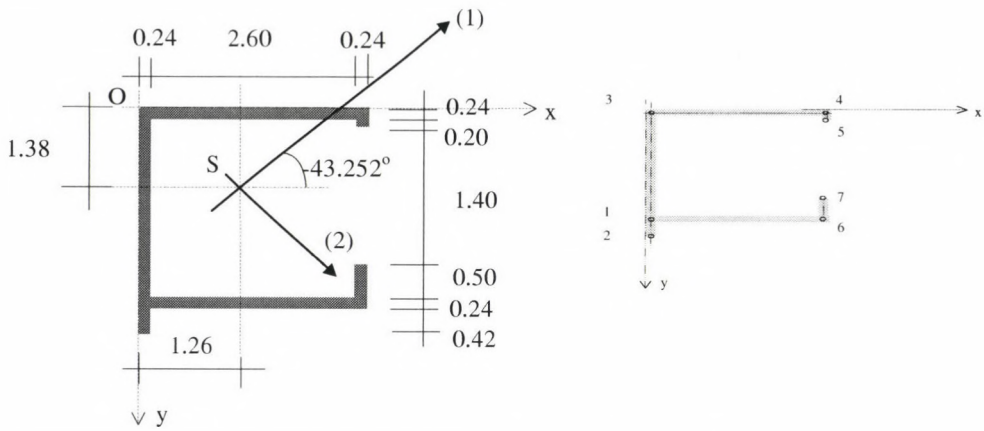


Fig. 6.2. Geometry of the wall-group and the joint numbering

6.1.1. Transforming coordinates

Since formulae used for determining sectorial coordinates are valid in the coordinate system of the principal axes, coordinates have to be transformed in the following way:

$$x = x \cos\alpha + y \sin\alpha - x_s$$

$$y = -x \sin\alpha + y \cos\alpha - y_s .$$

Original and transformed coordinates are summarised in the following table:

	$x$ [m]	$y$ [m]	$x$ [m]	$x$ [m]
Group 1	0	2.34	-1.57	0.002
Group 2	0	2.88	-1.94	0.39
Group 3	0	0	0.03	-1.70
Group 4	2.84	0	2.1	0.24
Group 5	2.84	0.32	1.88	0.47
Group 6	2.84	2.34	0.49	1.95
Group 7	2.84	1.72	0.92	1.50

6.1.2. Calculating sectorial coordinates

We randomly chose the position of the pole and the position of the zero point. Let that be point 1.

$$(4.1) \quad \omega_1 = \omega_2 = \omega_3 = \omega_4 = 0,$$

$$\omega_5 = -2.84 \cdot 2.34 = -6.65 \text{ m}^2,$$

$$\omega_6 = -6.65 - 2.84 \cdot 0.32 = -7.55 \text{ m}^2$$

$$\omega_7 = 2.84 \cdot 0.62 = 1.76 \text{ m}^2$$

Sectorial centrifugal moment:

$$(4.2) \quad I_{x\omega} = (0.24 \cdot 2.84)/6 \cdot [-6.65(2 \cdot 2.10 + 0.03)] + (0.24 \cdot 0.32)/6 \cdot [-6.65(2 \cdot 2.1 + 1.88) - 7.55(2 \cdot 1.88 + 2.1)] + (0.24 \cdot 0.62)/6 \cdot [1.72(1.76(2 \cdot 0.92 + 0.49))] = -4.177 \text{ m}^5$$

$$(4.3) \quad I_{y\omega} = (0.24 \cdot 2.84)/6 \cdot [-6.65(2 \cdot 0.24 - 1.7)] + (0.24 \cdot 0.32)/6 \cdot [-6.65(2 \cdot 0.24 + 0.47) - 7.55(2 \cdot 0.47 + 0.24)] + (0.24 \cdot 0.62)/6 \cdot [1.76(2 \cdot 1.5 + 1.95)] = 0.943 \text{ m}^5$$

The above shown integration have been calculated by using commonly known formulae of numeric integration.

### 6.1.3. Shear center of the group of walls

In the coordinate system of the principal axes:

$$(4.4) \quad x_C = -0.943/2.467 = -0.3822 \text{ m}$$

$$(4.5) \quad y_C = -4.177/2.642 = -1.581 \text{ m}$$

in the  $x, y$  coordinate system that we assumed:

$$x_C = -0.3822 \cos(-43.252^\circ) - (-1.581) \sin(-43.252^\circ) + 0 = -1.36 \text{ m}$$

$$y_C = -0.3822 \sin(-43.252^\circ) + (-1.581) \cos(-43.252^\circ) + 2.34 = 1.45 \text{ m}$$

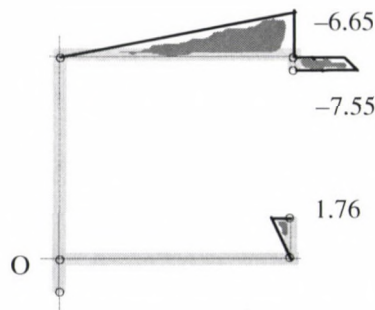


Fig. 6.3. Sectorial coordinates  $\omega_i$  [m<sup>5</sup>]



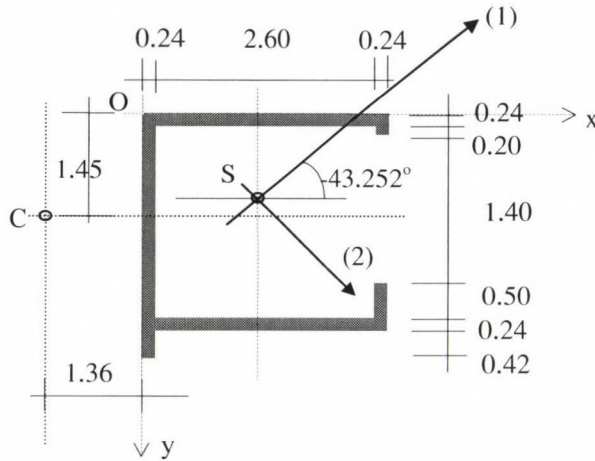


Fig. 6.4. The principal axes and the shear center

#### 6.1.4. Standardizing sectorial coordinates

We have to recalculate sectorial coordinates by taking the shear center as the new pole:

$$(4.1) \quad \omega_1 = 0$$

$$\omega_2 = 0 - 0.54 \cdot 1.36 = -0.734 \text{ m}^2$$

$$\omega_3 = 0 + 2.34 \cdot 1.36 = 3.18 \text{ m}^2$$

$$\omega_4 = 3.18 - 2.84 \cdot 1.45 = -0.938 \text{ m}^2$$

$$\omega_5 = -0.938 - 0.32 \cdot 4.2 = -2.28 \text{ m}^2$$

$$\omega_6 = 0 + 2.84 \cdot 0.89 = 2.53 \text{ m}^2$$

$$\omega_7 = 2.53 + 2.84 \cdot 0.89 + 0.62 \cdot 4.2 = 5.13 \text{ m}^2$$

$$(4.9) \quad S_\omega = \frac{(2.34 \cdot 0.24)}{2} \cdot 3.18 + \frac{(0.54 \cdot 0.24)}{2} \cdot (-0.734) + \frac{(2.84 \cdot 0.24)}{2} \cdot 2.53 + \frac{(0.62 \cdot 0.24)}{2} \cdot (2.53 + 5.13) + \frac{(2.84 \cdot 0.24)}{2} \cdot (3.18 - 0.938) + \frac{(0.32 \cdot 0.24)}{2} \cdot (-2.28 - 0.938) = 2.918 \text{ m}^4$$

$$(4.8) \quad \omega_0 = -2.918 / 2.28 = -1.279 \text{ m}^2$$

$$(4.7) \quad \omega_1 = 0 - 1.279 = -1.279 \text{ m}^2$$

$$\omega_2 = -0.734 - 1.279 = -2.013 \text{ m}^2$$

$$\omega_3 = 3.18 - 1.279 = 1.9 \text{ m}^2$$

$$\omega_4 = -0.938 - 1.279 = -2.217 \text{ m}^2$$

$$\omega_5 = -2.28 - 1.279 = -3.559 \text{ m}^2$$

$$\omega_6 = 2.53 - 1.279 = 1.251 \text{ m}^2$$

$$\omega_7 = 5.13 - 1.279 = 3.85 \text{ m}^2$$

6.1.5. Calculating sectorial moment of inertia

$$(4.10) \quad I_{\Omega\Omega} = (2.84 \cdot 0.24) / 6 \cdot [1.9 \cdot (2 \cdot 1.9 - 2.217) - 2.217(2 \cdot -2.217 + 1.9)] + \\ (2.84 \cdot 0.24) / 6 \cdot [-1.279(2 \cdot -1.279 + 1.251) + 1.251(2 \cdot 1.251 - 1.279)] + \\ (2.88 \cdot 0.24) / 6 \cdot [1.9(2 \cdot 1.9 - 2.013) - 2.013(2 \cdot -2.103 + 1.9)] + \\ (0.32 \cdot 0.24) / 6 \cdot [-2.217(2 \cdot -2.217 - 3.559) - 3.559 \cdot (2 \cdot -3.559 - 2.217)] + \\ (0.62 \cdot 0.24) / 6 \cdot [1.251(2 \cdot 1.251 + 3.85) + 3.85(2 \cdot 3.85 + 1.251)] = 3.93 \text{ m}^6$$

Thus, we have defined the warping characteristics of group No. 2.

Let us assume that we have already calculated the geometrical characteristics of the other wall-groups and we summarise those in the following table:

	$x_s$ [m]	$y_s$ [m]	$A$ [m <sup>2</sup> ]	$I_x$ [m <sup>4</sup> ]	$I_y$ [m <sup>4</sup> ]	$I_{xy}$ [m <sup>4</sup> ]	$x_C$ [m]	$y_C$ [m]
Group 1	0.43	1.94	1.74	5.104	0.602	-0.978	0.12	0.12
Group 2	8.53	3.39	2.28	2.55	2.56	-0.09	6.04	3.57
Group 3	7.86	6.88	1.2	0	2.5	0	7.86	6.88
Group 4	14.58	1.50	0.72	0.54	0	0	15.48	1.5
				8.194	5.662	-1.065		

Calculate the following values:

$$\sum y_{s,i} \cdot y_{C,i} = 0.072 + 9.1392 + 17.2 + 0 = 26.41 \text{ m}^5, \\ \sum x_{s,i} \cdot x_{C,i} = 0.612 + 15.4 + 0 + 8.36 = 24.37 \text{ m}^5, \\ \sum x_{y,i} \cdot x_{C,i} = -0.1176 - 0.5436 + 0 + 0 = -0.6612 \text{ m}^5, \\ \sum x_{y,i} \cdot y_{C,i} = -0.1176 - 0.3213 + 0 + 0 = -0.4389 \text{ m}^5.$$

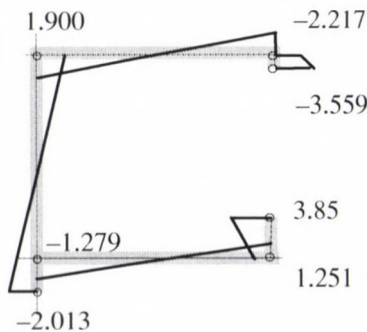


Fig. 6.5. Sectorial coordinates  $\Omega$  [m<sup>5</sup>]

6.1.6. Shear center of the equivalent core:

$$(4.11) \quad x_C = \frac{-1.065[26.41 - (-0.6612)] - 5.662[-0.4389 - 24.37]}{8.194 * 5.662 - 1.065^2} = 2.47 \text{ m},$$

$$(4.12) \quad y_C = \frac{8.194[26.41 - (-0.6612)] - (-1.065)[-0.4389 - 24.37]}{8.194 * 5.662 - 1.065^2} = 4.32 \text{ m}.$$

6.1.7. Rigidities of the equivalent core:

$$(4.13) \quad \begin{aligned} I_{xx} &= 8.194 \text{ m}^4, \\ I_{yy} &= 5.662 \text{ m}^4, \\ I_{xy} &= -1.065 \text{ m}^4. \end{aligned}$$

We use the following table for calculating sectorial inertia (warping rigidity):

	$\bar{x}_{C,i}$ [m]	$\bar{y}_{C,i}$ [m]	$I_{yy,i} * \bar{y}_{C,i}^2$ [m <sup>6</sup> ]	$I_{xy,i} * \bar{x}_{C,i} * \bar{y}_{C,i}$ [m <sup>6</sup> ]	$I_{xx,i} * \bar{x}_{C,i}^2$ [m <sup>6</sup> ]	$I_{\Omega\Omega,i}$ [m <sup>6</sup> ]
Group 1	-2.35	-4.2	10.62	-9.673	28.19	0
Group 2	3.57	-0.75	1.44	0.24	32.50	3.93
Group 3	5.39	2.56	16.384	0	0	0
Group 4	13.01	-2.82	0	0	91.4	0
			28.44	-9.433	152.09	3.93

$$(4.14) \quad I_{\Omega\Omega} = 3.93 + 28.44 - 2 * (-9.433) + 152.09 = 203.33 \text{ m}^6.$$

Loads affecting the building:

$$p_x = 50 \text{ kN/m}$$

$$p_y = 60 \text{ kN/m}$$

$$m_D = -60 * 5.33 - 50 * 0.82 = -360.8 \text{ kNm/m}$$

6.1.8. Displacement of the equivalent core:

$$(4.16) \quad Eu''' = \frac{5.662 * 60 - (-1.065) * 50}{8.194 * 5.662 - 1.065^2} = 8.682 \text{ kN/m}^5$$

$$(4.17) \quad Ev''' = \frac{8.194 * 50 - (-1.065) * 60}{8.194 * 5.662 - 1.065^2} = 10.464 \text{ kN/m}^5$$

$$(4.18) \quad E\phi''' = -\frac{360.8}{203.33} = -1.774 \text{ kN/m}^5 * \text{m}.$$



6.1.9. Distribution of forces acting on the wall-groups

Load of the first group:

$$(4.20) \quad p_{x,1} = 0.602 * [10.464 + (-4.2) * (-1.774)] + (-0.98) * [8.682 - (-2.35) * (-1.774)] = 362 \text{ kN/m}$$

$$(4.21) \quad p_{y,1} = (-0.98) * [10.464 + (-4.2) * (-1.774)] + 5.104 * [8.682 - (-2.35) * (-1.774)] = 5.478 \text{ kN/m}$$

In a similar way for the other wall-groups:

$$p_{x,2} = 28.61 \text{ kN/m}$$

$$p_{y,2} = 37.3 \text{ kN/m}$$

$$p_{x,3} = 15.3 \text{ kN/m}$$

$$p_{y,3} = 0 \text{ kN/m}$$

$$p_{x,4} = 0 \text{ kN/m}$$

$$p_{y,4} = 17.22 \text{ kN/m}$$

Problem 6.2

In this example we analyze the importance of the group's own warping rigidity and the ways in which rigidity influences results.

6.2.1. Let us annualise the problem shown in [6]:

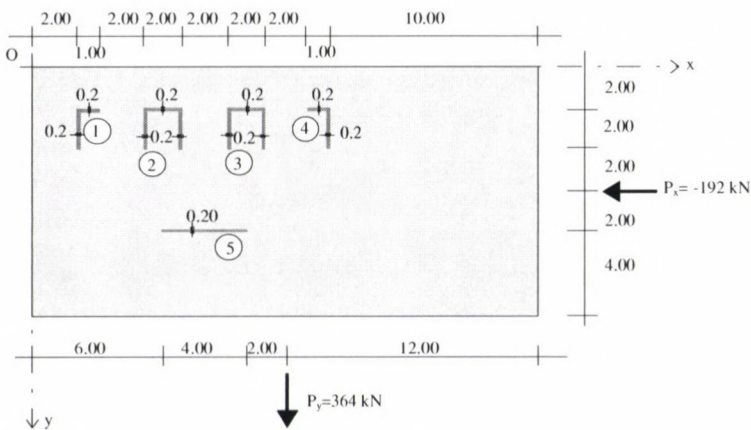


Fig. 6.6

We have calculated forces acting on wall-groups by taking into account their own warping rigidity or by neglecting that; results are summarised in the following table:

Table 6.1

	Warping rigidity of wall groups				Maximum difference $\Delta$ [%]
	Neglected		Taken into account		
	$P_{x,i}$ [kN]	$P_{y,i}$ [kN]	$P_{x,i}$ [kN]	$P_{y,i}$ [kN]	
Group 1	-3.1	12.6	-2.9	12.1	6.4
Group 2	18.3	93.0	19.1	92.6	4.4
Group 3	18.3	163.0	19.1	163.4	4.4
Group 4	26.7	115.4	26.8	115.9	0.4
Group 5	-252.2	0	-254.1	0	0.7
	-192.0	384.0	-192.0	384.0	

6.2.2. Let us analyse the example given above with the difference that we reduce the width of walls No. 1, 4, 5 to 10 cm and increase the width of walls No. 2, 3 to 30 cm.

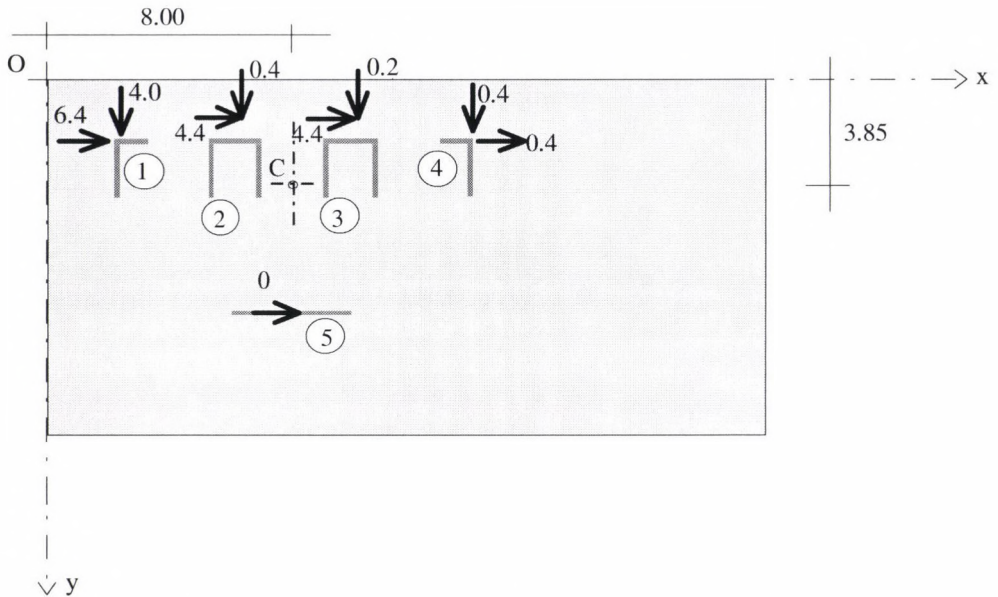


Fig. 6.7. Differences between forces acting on shear walls (in  $|\Delta|$  %)

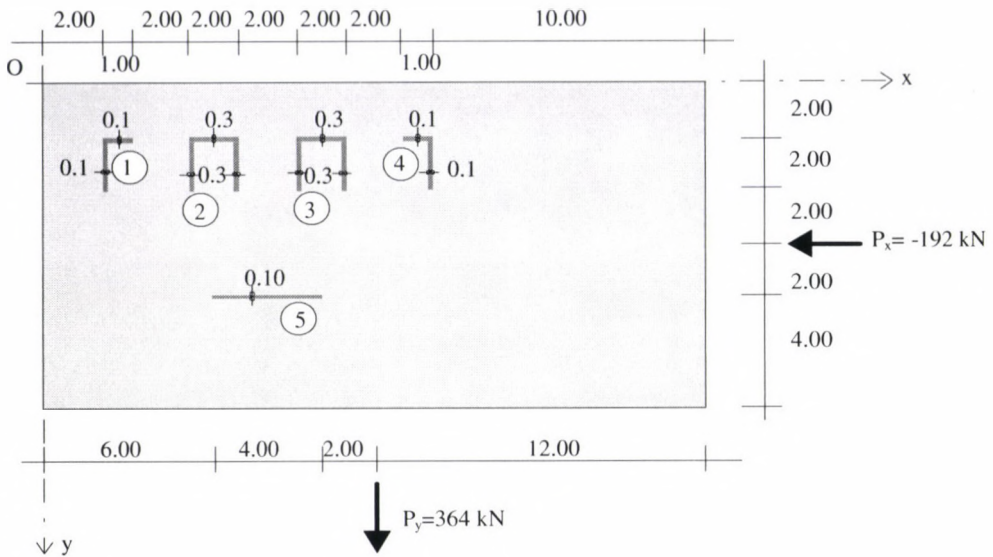


Fig. 6.8

We have calculated forces acting on wall-groups by taking into account their own warping rigidity or by neglecting that; results are summarized in the following table:

Table 6.2

	Warping rigidity of wall groups				Maximum difference
	Neglected		Taken into account		
	$P_{x,i}$ [kN]	$P_{y,i}$ [kN]	$P_{x,i}$ [kN]	$P_{y,i}$ [kN]	$\Delta$ [%]
Group 1	2.9	-15	3.2	-16	10.3
Group 2	10.1	75.9	12.1	73.8	19.8
Group 3	10.1	253.2	12.1	255.3	19.8
Group 4	15.7	69.8	15.9	70.9	1.6
Group 5	-230.8	0	-235.4	0	2
	-192.0	384.0	-192.0	384.0	



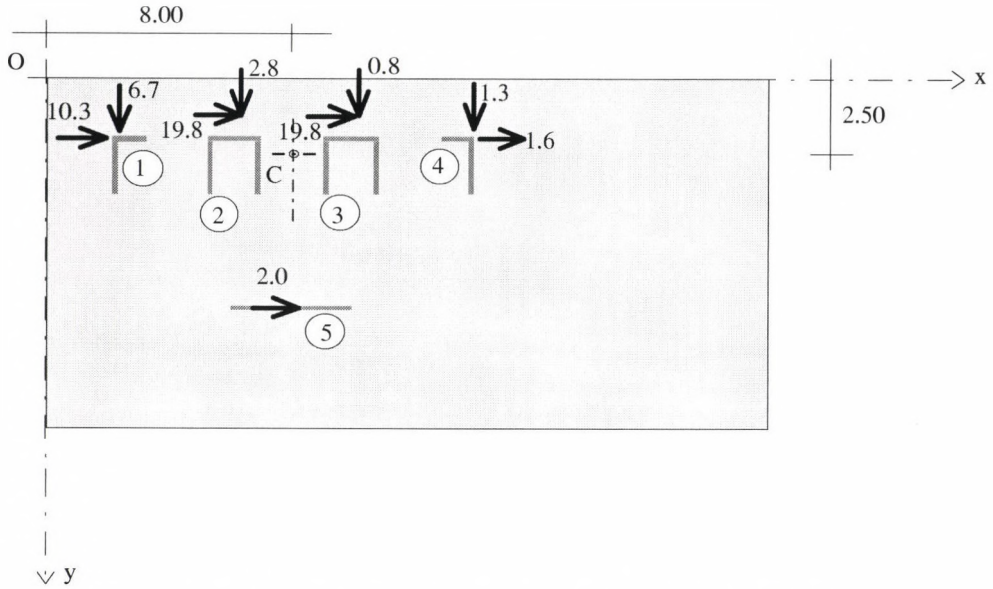


Fig. 6.9. Differences between forces acting on shear walls (in  $|\Delta|$  %)

6.2.3. Let us analyse the example given above with the difference that we reduce the length of group 5 to half of the original length.

We have calculated forces acting on wall-groups by taking into account their own warping rigidity or by neglecting that; results are summarized in the following table:

Table 6.3

	Warping rigidity of wall groups				Maximum difference
	Neglected		Taken into account		
	$P_{x,j}$ [kN]	$P_{y,j}$ [kN]	$P_{x,j}$ [kN]	$P_{y,j}$ [kN]	$\Delta$ [%]
Group 1	12.7	-64.9	13.7	-69.6	7.9
Group 2	-79.6	-33.6	-79	-43.5	29.5
Group 3	-79.6	362.7	-79	372.6	4.3
Group 4	25.4	119.8	26.5	124.5	1.6
Group 5	-70.9	0	-74.1	0	4.5
	-192.0	384.0	-192.0	384.0	

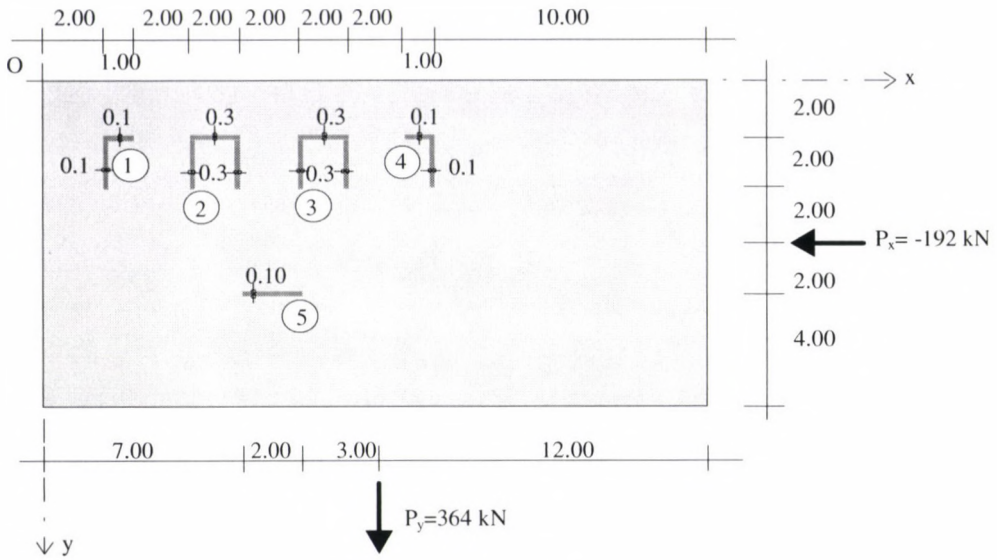


Fig. 6.10

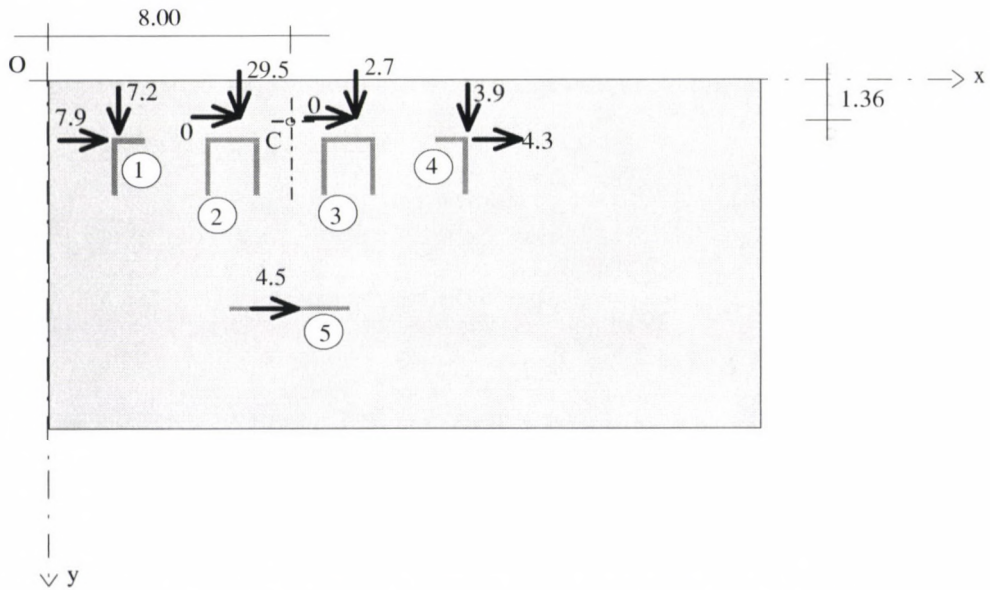


Fig. 6.11. Differences between forces acting on shear walls (in  $|\Delta| \%$ )

### 6.3. Evaluation of numerical examinations

By comparing tables 6.1, 6.2 and 6.3 we can make the following conclusions:

- neglecting the warping rigidity of wall-groups does not cause a significant change ( $\Delta < 10\%$ ) in the results in case of a geometrically “balanced” system of shear walls with the necessary number of walls and with proper dimensions;
- if the geometrical layout (shear walls concentrated around the shear center, see example 6.2.3), and dimensions of walls are such that the magnitude of warping rigidity approaches the flexural rigidity of the group of walls (short, thick walls, see wall-groups No. 2 and 3 in example 6.2.2, 6.2.3), than neglecting the wall-groups’ own warping rigidity might cause significant errors ( $\Delta$  even 30%);
- due to a complex geometric layout neither the magnitude of the error nor its position can be determined (in example 6.2.2 the greatest difference occurred in force  $p_x$  with wall-groups No. 2 and 3 and in example 6.2.3 the greatest difference was spotted in  $p_y$  of the group of walls No. 2).

Results shown above can be explained by formula (4.14). The condition under which the wall-groups’ own warping rigidity may be neglected is when the second, third and fourth component of the formula is several orders of magnitudes larger than the first component (compare the magnitude of warping rigidity noted with the group of walls in example 6.1 with the rigidity of the equivalent core). We can achieve this by evenly distributing walls, placing them as far from each other as possible within the plan (in the formula we are using the squared value of the distance between the wall-groups’ and the level’s shear center) and leaving the inertia of the walls at a maximum level (the length of the wall is raised to the third power).

## 7. Conclusion

In the study we have presented a numerical analysis of influences acting on the shear walls of tall buildings. By using the presented method we can calculate shear wall loads in tall buildings with greater accuracy than by using the manual method, thus wall dimensions and loads acting on foundations are more reliable. During our numeric examinations we have emphasised the importance of the wall-groups’ own rigidity and errors due to neglecting them. We have also presented the use of applied object-oriented programming.

## Acknowledgement

This paper presents some results of the research project sponsored by OTKA (contract No. F-015862). Further the authors are grateful to Dr. István Bódi for his help and assistance.



### References

1. W. S. Wlassow: Dünnwändige elastische Stäbe. Band I. Berlin. VEB Verlag für Bauwesen. 1964. pp. 49–113.
2. Szmodits, Kázmér: Útmutató panelépületek statikai tervezéséhez. Budapest. Építéstudományi Intézet. 1975. (in Hungarian)
3. Kollár, Lajos–Póth, László: Több falelemmel merevített épület igénybevételei. Építés-, Építészettudomány. XXIV (1994) 3–36. (in Hungarian)
4. Dorde Ladinovic: Prilog statickoj i dinamickej analizi visokih zgrada ukrucenih vertikalnim tankozidnim elementima. MSci Thesis. Technical University of Novi Sad. 1994.
5. B. Brandt–H. Schäfer–H. Reeh: Zum Stabilitätsnachweis von Hochhäusern. Beton- und Stahlbetonbau. 70 (1975) 211–223.
6. Massányi, Tibor–Dulácska, Endre: Statikusok könyve. Budapest. Műszaki Könyvkiadó. 1989. pp. 107–114. (in Hungarian)
7. Ládi, Nándor–Klopka, Zoltán: Gaven – Users Manual. 1996.



## CONTINUUM MECHANICS

### (Summary and Comparison of Solid and Fluid Mechanics)

Kozák, I.

*University of Miskolc, Department of Mechanics, H-3515 Miskolc, Hungary*

(Received: 18 March 1996)

The first part of the paper has a qualitative character, it emphasises the main features (according to the description methods, the strain and strain rate, the system of internal forces and the influence of the environment) of three typical mechanical models such as rigid body, solid body and fluid. The second part has a quantitative character, it summarises the kinematics, the general laws and the constitutive equations of continua. Particular attention is devoted to the comparison of solid and fluid mechanics.

#### 1. Introduction

At the meeting held on November 17th, 1994, the Hungarian National Committee of the International Union of Theoretical and Applied Mechanics (IUTAM) has entrusted the author with delivering a lecture about continuum mechanics at the opening plenary session of the 7th Hungarian Conference of Mechanics. According to the requirements, the lecture was presented for scientists of both fields (solid and fluid mechanics), demonstrating generality and validity of mechanics of continua both for solid bodies and for fluids.

Preparing for the lecture, the first guess of the author was about the theory of constitutive equations, later on he considered to give a historical view of the development of continuum mechanics, but he had to realise that these topics would exceed the limits of a one hour lecture. In the end he delivered his lecture, according to his knowledge based on his previous activity in mechanics of solid bodies and gave an overview of the subject, in some cases presenting special features and underlining similarities and differences between the models of solids and fluids.

After the lecture the author was asked by some of his colleagues to publish his lecture. The present paper is an attempt to meet these demands. The author has tried to overcome the formal difficulties caused by the technical differences between an oral lecture with projected transparencies and a published paper. This circumstance has to be emphasised, because the author has not attempted to write a detailed paper on the subject. Instead he intended, after an appropriate rearrangement, to supplement the projected materials with additional explanations.





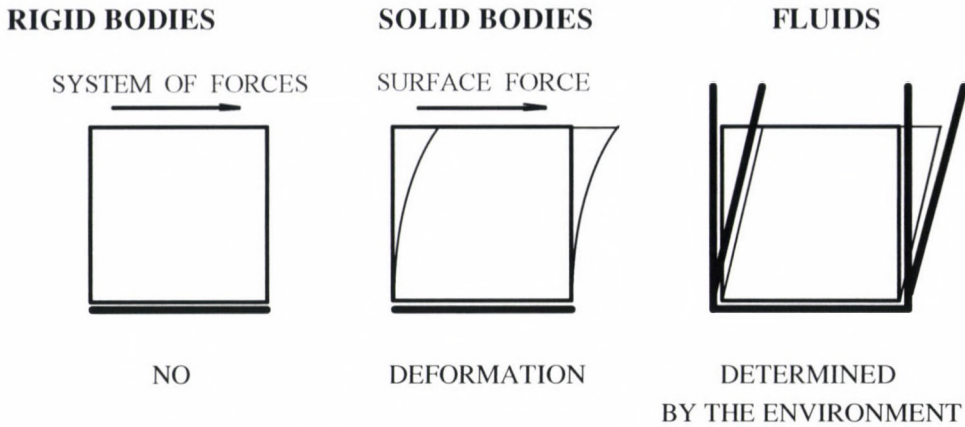


Fig. 2. Characteristic forms of motion in comparison with environment (in static state)

In Fig. 2 the heavy lines represent the environment. In the first two cases, it is assumed that the points on the bottom of each body are not able to move with respect to the environment.

Figures 3 and 4 demonstrate (a) how to identify the points of a body and (b) the related description methods. It is worth mentioning that for rigid bodies and fluids, the present configuration is preferred. For rigid bodies, however, the material description, for fluids the spatial description are used.

For solid bodies in the initial configuration the Lagrangian description, in the present configuration the spatial or the material description are employed. It can also be mentioned that for the solution of some problems of fluids, the spatial-Lagrangian method is useful. In this case, for a short period of time, the velocity of material points is defined by the spatial velocity field.

In Fig. 5 and Fig. 6 the displacement and the velocity field in Fig. 7 and Fig. 8 the strain and strain rate fields are compared for the three basic models.

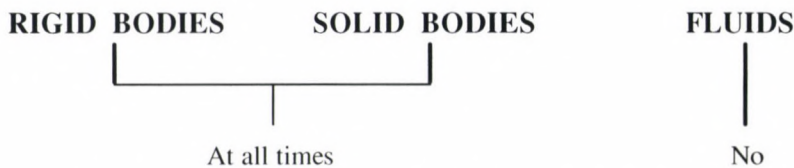


Fig. 3. Possibility of the identification of material points during the motion

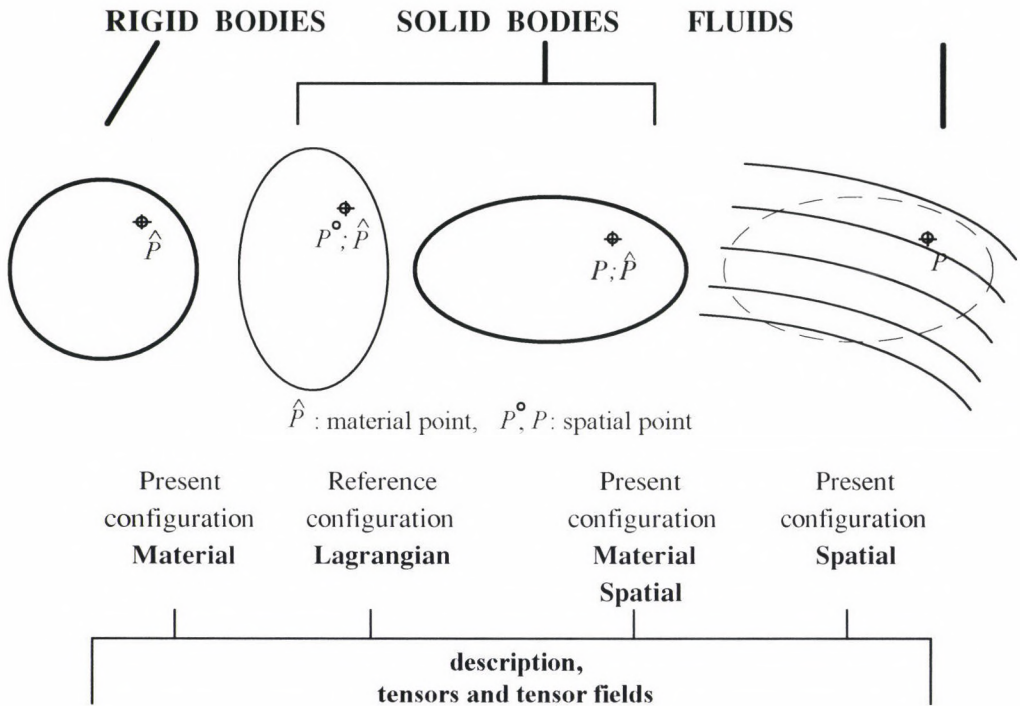


Fig. 4. Description methods

Figure 9 emphasises the Cauchy assumption and shows the structures of the Cauchy stress tensor  $T$  defined in the present configuration.  $T_d$  denotes the stress deviator,  $p$  is the hydrodynamic pressure,  $T^V$  is the viscous stress tensor,  $T^{tur}$  is the turbulent stress tensor. Figure 9 also illustrates the dependency of  $T$  (without the heat effect) and the characteristic stress tensors.

In Fig. 10 the method of calculation of stress power is presented (including that that for rigid bodies there is not stress power at all) with an emphasis on the connection of the stress power to the strain rates.

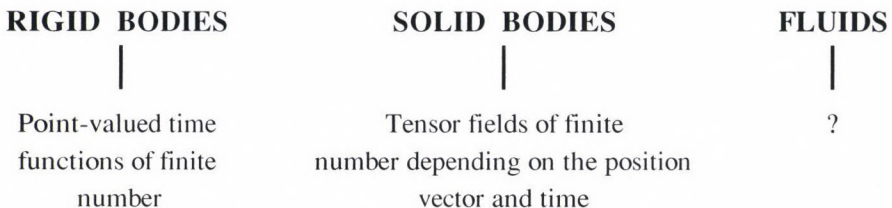


Fig. 5. Displacements



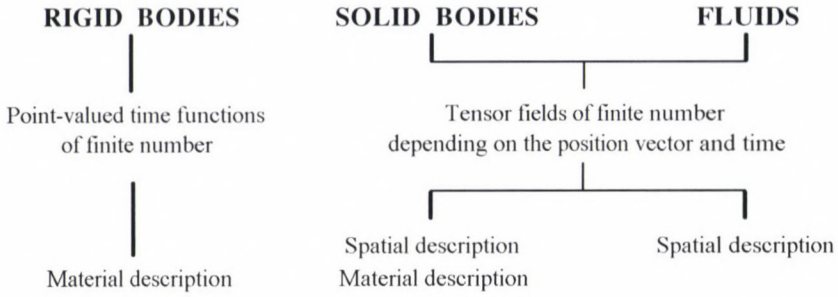


Fig. 6. Velocities

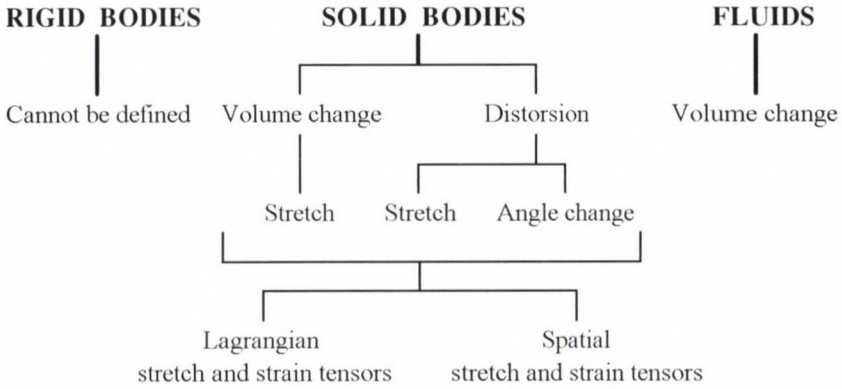


Fig. 7. Deformation

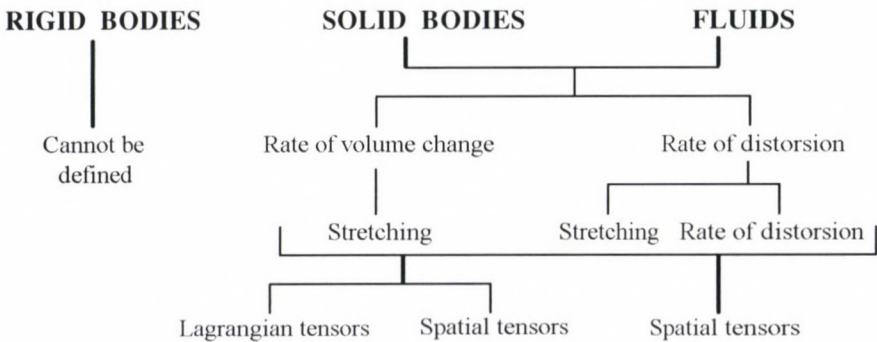
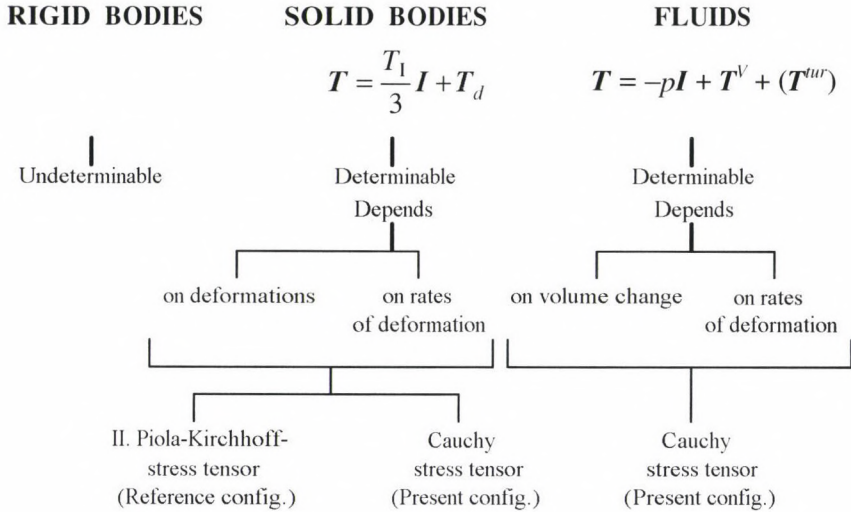


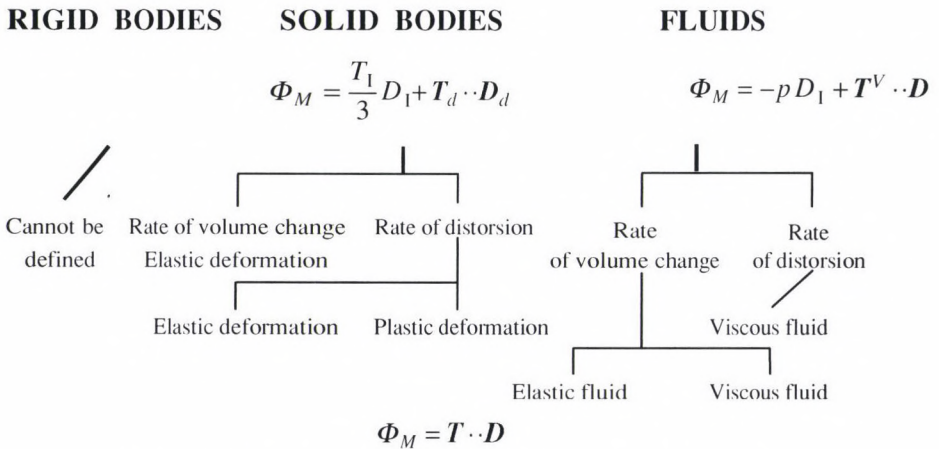
Fig. 8. Rates of deformation



Cauchy's hypothesis:  $\rho_n = \rho_n(\mathbf{r}, \mathbf{n}; t) = \mathbf{T}(\mathbf{r}; t) \cdot \mathbf{n}$

$\mathbf{T}$  is the Cauchy stress tensor

Fig. 9. Internal forces. Stress tensors



$\mathbf{T}$  is the Cauchy stress tensor,  $\mathbf{D}$  is the strain rate tensor

Fig. 10. Stress power per unit volume  
(Negative value of internal force)

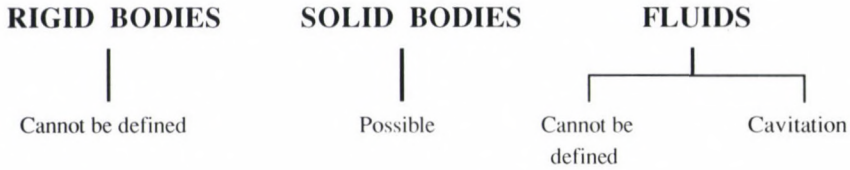


Fig. 11. Fracture. Plastic state

Figure 11 represents the possibility of fracture and elastic-plastic deformation. From this point of view cavitation can be called the fracture of fluids.

Figure 12 shows the effect of environment for the three basic models.

In Fig. 13 the structure of continuum mechanics (mechanics of solid bodies and fluids) is described without detailing the special theories of continuum mechanics (for example: elasticity, shell theory, etc.) The underlined expressions are valid equally for solid bodies and fluids. It follows from the nature of the matter that the kinematics of continua and the laws of classical physics are related to all continua, however, depending on the what the material properties are the constitutive equations differ from each other (for example: elastic body, thermoviscous fluid).

### 3. Kinematics of continua

#### 3.1. Nonlinear theory of deformation

Let the co-ordinates of the *spatial points*  $P^0$  and  $P$  be

$x^{01}, x^{02}, x^{03}$  in the reference configuration

and

$x^1, x^2, x^3$  in the present configuration.

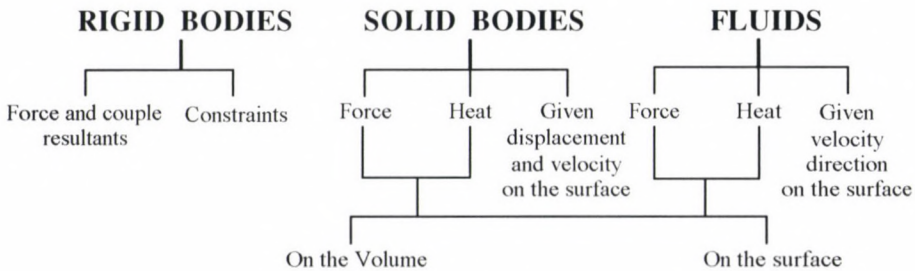


Fig. 12. Influence of environment



CONTINUUM MECHANICS

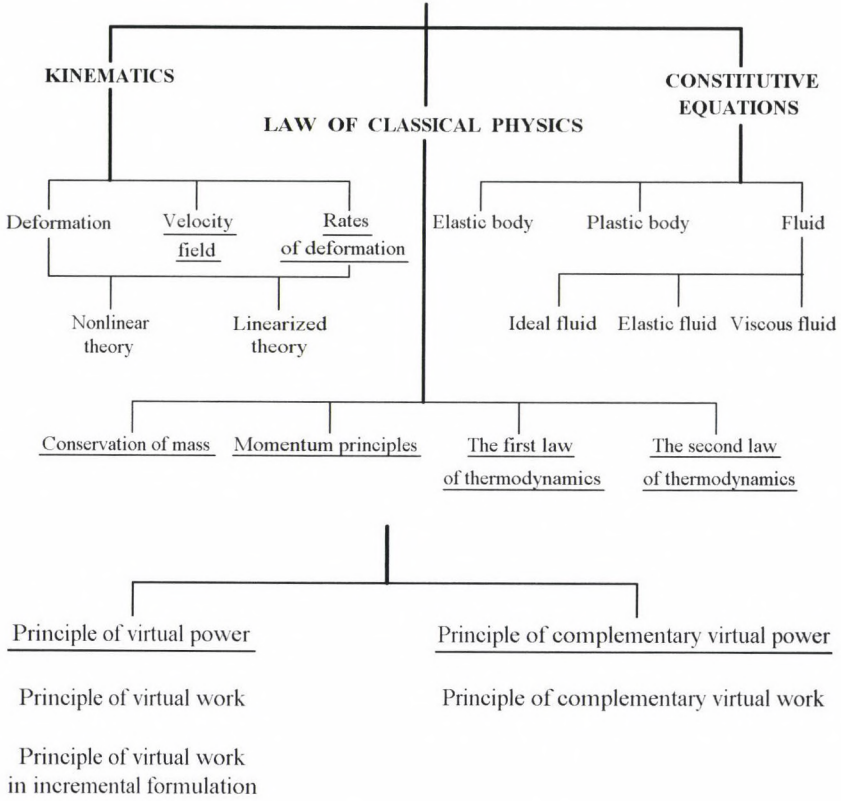


Fig. 13. The Continuum Mechanics

Co-ordinates of the point  $\hat{P}$  of continuum (convective co-ordinates) are denoted by  $X^1, X^2, X^3$  both in the reference and present configurations.

If the position vectors of  $\hat{P}$  are  $\mathbf{r}^o$  and  $\mathbf{r}$  in the reference and present configuration then the motion of continua is given by

$$\mathbf{r} = \mathbf{r}(\mathbf{r}^o; t); \quad x^p = x^p(x^{o1}, x^{o2}, x^{o3}; t); \quad J = \left| \frac{\partial x^p}{\partial x^{ok}} \right| > 0.$$

The points of continua are identified by the position co-ordinates in the reference configuration:

$$x^{ok} = k^{ok}(X^1, X^2, X^3).$$

The motion of the solid body can also be given by *the displacement field*:

$$\mathbf{r} = \mathbf{r}^\circ + \mathbf{u}^\circ \quad \text{or} \quad \mathbf{r}^\circ = \mathbf{r} + (-\mathbf{u}),$$

in which

- the displacement field  $\mathbf{u}^\circ = \mathbf{u}^\circ(x^{\circ 2}, x^{\circ 2}, x^{\circ 3}; t)$  is related to the reference configuration,
  - the displacement field  $(-\mathbf{u}) = -\mathbf{u}(x^1, x^2, x^3; t)$  is related to the present configuration
- and it holds that  $\mathbf{u}^\circ = \mathbf{u}$ .

*Geometrical measures of deformation* are understood as relations between the geometrical characteristics of continuum elements. These are

$$d\mathbf{r}^\circ, ds^\circ, \alpha^\circ, d\mathbf{A}^\circ, dA^\circ \text{ and } dV^\circ$$

in the reference configuration and

$$d\mathbf{r}, ds, \alpha, d\mathbf{A}, dA \text{ and } dV$$

in the present configuration, where

$d\mathbf{r}^\circ$  and  $d\mathbf{r}$  are material line element vectors,  
 $ds^\circ$  and  $ds$  are scalar material line elements,  
 $\alpha^\circ$  and  $\alpha$  } and { the angles formed } by two material  
 $d\mathbf{A}^\circ$  and  $d\mathbf{A}$  } surface element vectors determined } line  
 $dA^\circ$  and  $dA$  } scalar surface elements determined } element vectors,  
 $dV^\circ$  and  $dV$  are the volume elements determined by three material line element vectors.

Two material line element vectors are related to each other by

$$d\mathbf{r} = \mathbf{F} \cdot d\mathbf{r}^\circ; \quad J = \det[\mathbf{F}] > 0; \quad d\mathbf{r}^\circ = \mathbf{F}^{-1} \cdot d\mathbf{r}.$$

The deformation gradient  $\mathbf{F}$  and the inverse deformation gradient  $\mathbf{F}^{-1}$  can be given in terms of the displacement gradient both in the reference and in the present configurations ( $\nabla^\circ$  and  $\nabla$  are the Hamilton operators):

$$\mathbf{F} = \mathbf{I} + \mathbf{u}^\circ \circ \nabla^\circ; \quad \mathbf{F}^{-1} = \mathbf{I} + (-\mathbf{u}) \circ \nabla,$$

$$\mathbf{F}^{-1} \cdot \mathbf{F} = (\mathbf{I} - \mathbf{u} \circ \nabla) \cdot (\mathbf{I} + \mathbf{u}^\circ \circ \nabla^\circ) = \mathbf{I}.$$

The polar decomposition of the deformation gradient is of the form

$$\mathbf{F} = \mathbf{R} \cdot \mathbf{U}^\circ = \mathbf{V} \cdot \mathbf{R},$$

in which

$$\mathbf{R}^T = \mathbf{R}^{-1}; \quad (\det|\mathbf{R}| = 1) \quad \text{is the rotation tensor,}$$

$\mathbf{U}^\circ; ((\mathbf{U}^\circ)^T = \mathbf{U}^\circ)$  is the right stretch tensor,  
 $\mathbf{V}; (\mathbf{V}^T = \mathbf{V})$  is the left stretch tensor.

The tensors  $\mathbf{U}^\circ$  and  $\mathbf{V}$  are positive definite.

Introducing

the right Cauchy–Green strain tensor  $\mathbf{C}^\circ = \mathbf{F}^T \cdot \mathbf{F} = (\mathbf{U}^\circ)^2$  and

the left Cauchy–Green strain tensor  $\mathbf{B} = \mathbf{F} \cdot \mathbf{F}^T = \mathbf{V}^2$

for the material line element vectors we can write:

$$\underline{d\mathbf{r}} = \mathbf{R} \cdot (\mathbf{C}^\circ)^{1/2} \cdot d\mathbf{r}^\circ; \quad d\mathbf{r}^\circ = \mathbf{R}^{-1} \cdot \mathbf{B}^{-1/2} \cdot d\mathbf{r}.$$

If  $d\mathbf{r}^\circ = \mathbf{e}^\circ ds$ ,  $d\mathbf{r} = \mathbf{e} ds$  and  $|\mathbf{e}^\circ| = |\mathbf{e}| = 1$ , then the stretch is

$$\lambda_s = \frac{ds}{ds^\circ} = \frac{\sqrt{\mathbf{e}^\circ \cdot \mathbf{C}^\circ \cdot \mathbf{e}^\circ}}{\sqrt{\mathbf{e} \cdot \mathbf{B}^{-1} \cdot \mathbf{e}}}.$$

The extension assumes the form

$$\epsilon_s = \frac{ds - ds^\circ}{ds^\circ} = \frac{ds}{ds^\circ} - 1 = \lambda_s - 1.$$

Assume that  $\alpha = \alpha^\circ - \gamma_{12}$ ;  $\alpha^\circ = \frac{\pi}{2}$  and  $\mathbf{e}_1^\circ \cdot \mathbf{e}_2^\circ = \mathbf{e}_1 \cdot \mathbf{e}_2 = 0$ . Then the angle change is

$$\gamma_{12} = \frac{\mathbf{e}_1^\circ \cdot (\mathbf{C}^\circ - \mathbf{I}) \cdot \mathbf{e}_2^\circ}{\lambda_{s1} \lambda_{s2}} = \mathbf{e}_1 \cdot (\mathbf{I} - \mathbf{B}^{-1}) \cdot \mathbf{e}_2.$$

The change of a material surface element is given by

$$d\mathbf{A} = J \sqrt{\frac{g}{g^\circ}} (\mathbf{F}^{-1})^T \cdot d\mathbf{A}^\circ = J \sqrt{\frac{g}{g^\circ}} \mathbf{R} \cdot (\mathbf{C})^{1/2} \cdot d\mathbf{A},$$

$$d\mathbf{A}^\circ = \frac{1}{J} \sqrt{\frac{g^\circ}{g}} \mathbf{R}^{-1} \cdot \mathbf{B}^{1/2} \cdot d\mathbf{A}$$

where  $g^\circ$  and  $g$  are the determinant of the metric tensor in the reference and the present configuration, respectively.

If  $d\mathbf{A}^\circ = \mathbf{n}^\circ dA^\circ$ ;  $d\mathbf{A} = \mathbf{n} dA$  and  $|\mathbf{n}^\circ| = |\mathbf{n}| = 1$ , then for the ratio of scalar surface elements we obtain

$$\lambda_A = \frac{dA}{dA^\circ} = J \sqrt{\frac{g}{g^\circ}} \sqrt{\mathbf{n}^\circ \cdot (\mathbf{C})^{-1} \cdot \mathbf{n}^\circ} = \frac{1}{J} \sqrt{\frac{g}{g^\circ}} \frac{1}{\sqrt{\mathbf{n} \cdot \mathbf{B} \cdot \mathbf{n}}}.$$



The ratio of volume elements is

$$\lambda_V = \frac{dV}{dV^0} = J \sqrt{\frac{g}{g^0}} = \sqrt{C_{III}^0} = \frac{1}{\sqrt{(B^{-1})_{III}}}$$

The volume change is defined by

$$\varepsilon_V = \frac{dV - dD^0}{dV^0} = \frac{dV}{dV^0} - 1 = \lambda_V - 1.$$

It is also customary to use  
the *Green–Lagrange strain tensor*

$$\mathbf{E} = \frac{1}{2}(\mathbf{C}^0 - \mathbf{I}) = \frac{1}{2}[\mathbf{u} \circ \nabla^0 + \nabla^0 \mathbf{u}^0 + (\nabla^0 \circ \mathbf{u}^0) \cdot (\mathbf{u}^0 \circ \nabla^0)]$$

and the *Almansi–Euler strain tensor*

$$\mathbf{E} = \frac{1}{2}(\mathbf{I} - \mathbf{B}^{-1}) = \frac{1}{2}[\mathbf{u} \circ \nabla + \nabla \circ \mathbf{u} - (\nabla \circ \mathbf{u}) \cdot (\mathbf{u} \circ \nabla)].$$

It holds that

$$(\mathbf{u} \circ \nabla) = (\mathbf{u}^0 \circ \nabla^0) \cdot \mathbf{F},$$

$$\mathbf{E} = (\mathbf{F})^T \cdot \mathbf{E}^0 \cdot \mathbf{F}; \quad \mathbf{E}^0 = \mathbf{F}^T \cdot \mathbf{E} \cdot \mathbf{F}.$$

It can be pointed out that the scalar measures of deformation ( $\lambda_s$ ,  $\varepsilon_s$ ,  $\gamma_{12}$ ,  $\lambda_A$ ,  $\lambda_V$ ,  $\varepsilon_V$ ) can be calculated with the strain tensors (in the reference configuration using the tensor  $\mathbf{C}^0$  and in the present configuration using the tensor  $\mathbf{B}$ ), but for the computation of vectorial measure of deformation ( $d\mathbf{r}^0 \Leftrightarrow d\mathbf{r}$ ,  $d\mathbf{A}^0 \Leftrightarrow d\mathbf{A}$ ) both the rotation tensor and strain tensors are necessary. In the corresponding formulae the tensors  $\mathbf{C}^0$  and  $\mathbf{B}$  can be replaced by another strain tensors.

### 3.2. linearized theory of deformation

(In the reference configuration)

In the linearized theory of deformation we neglect the squares and products of the components of  $\mathbf{u}^0 \circ \nabla^0$  and  $\mathbf{u}^0 \nabla^0$  when they are compared with their first power. In this

case all quantities are regarded in the points of the reference configuration and the expressions of the nonlinear theory are linearized. The displacement field  $\mathbf{u}^\circ = \mathbf{u}^\circ(x^{\circ 1}, x^{\circ 2}, x^{\circ 3}, t)$  is sufficient for the description of the motion of continuum.

The additive decomposition of the displacement gradient takes the form

$$\mathbf{u}^\circ \circ \nabla^\circ = \boldsymbol{\varepsilon}^\circ + \boldsymbol{\Psi}^\circ,$$

$$\boldsymbol{\varepsilon}^\circ = \frac{1}{2}(\mathbf{u}^\circ \circ \nabla^\circ + \circ \mathbf{u}^\circ); \quad \boldsymbol{\Psi}^\circ = \frac{1}{2}(\mathbf{u}^\circ \circ \nabla^\circ - \nabla^\circ \circ \mathbf{u}^\circ),$$

in which

$$\boldsymbol{\varepsilon}^\circ; \left( (\boldsymbol{\varepsilon}^\circ)^T = \boldsymbol{\varepsilon}^\circ \right) \quad \text{is the infinitesimal and symmetric strain tensor}$$

and

$$\boldsymbol{\Psi}^\circ; \left( (\boldsymbol{\Psi}^\circ)^T = -\boldsymbol{\Psi}^\circ; \mathbf{R} = \mathbf{I} + \boldsymbol{\Psi}^\circ \right) \quad \text{is the infinitesimal and skew symmetric rotation tensor.}$$

The geometrical measure of deformation for a material line element is

$$\underline{d\mathbf{r}} = d\mathbf{r}^\circ + d\mathbf{u}^\circ = (\mathbf{I} + \mathbf{u}^\circ \circ \nabla^\circ) \cdot d\mathbf{r}^\circ = d\mathbf{r}^\circ + \boldsymbol{\varepsilon}^\circ \cdot d\mathbf{r}^\circ + \boldsymbol{\Psi}^\circ \cdot d\mathbf{r}^\circ = \underline{\underline{\mathbf{R} \cdot d\mathbf{r}^\circ + \boldsymbol{\varepsilon}^\circ \cdot d\mathbf{r}^\circ}}.$$

It is worth comparing the underlined expression above and the underlined expression of Section 3.1. According to the above formulae the rotation and the deformation of a continuum element are added up in the linearized theory of deformation. On the contrary the deformed continuum element takes part in the rotation in the nonlinear theory.

In the linearized theory of deformation the other geometrical measures can be calculated as follows:

$$\text{the stretch is} \quad \lambda_s = 1 + \mathbf{e}^\circ \cdot \boldsymbol{\Psi}^\circ \cdot \mathbf{e}^\circ,$$

$$\text{the extension is} \quad \varepsilon_s = \mathbf{e}^\circ \cdot \boldsymbol{\varepsilon}^\circ \cdot \mathbf{e}^\circ,$$

$$\text{the change of angle is} \quad \gamma_{12} = 2\mathbf{e}_1^\circ \cdot \boldsymbol{\varepsilon}^\circ \cdot \mathbf{e}_2^\circ, \quad \text{when } \mathbf{e}_1^\circ \cdot \mathbf{e}_2^\circ = 0,$$

$$\text{the ratio of volume elements is} \quad \lambda_V = 1 + \varepsilon_1^\circ,$$

$$\text{the change of volume is} \quad \varepsilon_V = \varepsilon_1^\circ = \mathbf{u}^\circ \circ \nabla^\circ.$$

### 3.3. Velocity field. Rates of deformation

(In the present configuration)

The spatial velocity field and its gradient are

$$\mathbf{r}' = \mathbf{v}(x, y, z, t); \quad \mathbf{L} = \mathbf{v} \circ \nabla; \quad (d\mathbf{r})' = d\mathbf{v} = \mathbf{L} \cdot d\mathbf{r}.$$

The additive decomposition for the velocity gradient is of the form

$$\mathbf{L} = \mathbf{v} \circ \nabla = \mathbf{D} + \mathbf{W},$$

$$\mathbf{D} = \frac{1}{2}(\mathbf{v} \circ \nabla + \nabla \circ \mathbf{v}); \quad \mathbf{W} = \frac{1}{2}(\mathbf{v} \circ \nabla - \nabla \circ \mathbf{v}) = \mathbf{I} \times \boldsymbol{\omega},$$

where

$\mathbf{D}$ ; ( $\mathbf{D}^T = \mathbf{D}$ ) is the symmetric strain rate tensor,

$\mathbf{W}$ ; ( $\mathbf{W}^T = -\mathbf{W}$ ) is the skew symmetric spin tensor,

$\boldsymbol{\omega} = \frac{1}{2} \nabla \times \mathbf{v}$  is the spin vector.

For the various strain rates we may write

$$\underline{\underline{(d\mathbf{r})'}} = d\mathbf{v} = \mathbf{L} \cdot d\mathbf{r} = \underline{\underline{\mathbf{W} \cdot d\mathbf{r}}} + \underline{\underline{\mathbf{D} \cdot d\mathbf{r}}} = \boldsymbol{\omega} \times d\mathbf{r} + \mathbf{D} \cdot d\mathbf{r},$$

$$(d\mathbf{A}) = \boldsymbol{\omega} \times d\mathbf{A} - (\mathbf{D} - \mathbf{I}D_1) \cdot d\mathbf{A},$$

$$(\ln \lambda_s)' = \frac{(ds)'}{ds} = \mathbf{e} \cdot \mathbf{D} \cdot \mathbf{e}; \quad (\varepsilon_s)' = (1 + \varepsilon_s) \mathbf{e} \cdot \mathbf{D} \cdot \mathbf{e},$$

$$(\gamma_{12})' = 2\mathbf{e}_1 \cdot \mathbf{D} \cdot \mathbf{e}_2, \quad \text{if } \mathbf{e}_1 \cdot \mathbf{e}_2 = 0,$$

$$(\ln \lambda_A)' = \frac{(dA)'}{dA} = D_1 - \mathbf{n} \cdot \mathbf{D} \cdot \mathbf{n},$$

$$(\ln \lambda_V)' = \frac{(dV)'}{dV} = D_1 = \mathbf{v} \cdot \nabla; \quad (\varepsilon_V)' = (1 + \varepsilon_V) D_1.$$



Comparing the underlined expression written above and the underlined expression of Section 3.2, we can conclude that, from mathematical point of view, the linearized theory of deformation is similar to the theory of rates of deformation.

It is also obvious that the scalar rates of deformation  $\left[ (\ln \lambda_s), (\varepsilon_s), (\gamma_{12}), (\ln \lambda_A), (\ln \lambda_V), (\varepsilon_V) \right]$  can be calculated using the strain rate tensor  $\mathbf{D}$ , while for the computation of vectorial strain rates the vector  $\boldsymbol{\omega}$  is also needed.

### 3.4. Compatibility conditions

In the nonlinear theory of deformation the metric tensor, that has changed due to the deformation, is the right Cauchy–Green strain tensor

$$\mathbf{I} + \mathbf{u}^\circ \circ \nabla^\circ + \nabla^\circ \circ \mathbf{u}^\circ + (\nabla^\circ \circ \mathbf{u}^\circ) \cdot (\mathbf{u}^\circ \circ \nabla^\circ) = \mathbf{C}^\circ.$$

It is easy to recognise that tensor equation above is overdetermined for the three components of  $\mathbf{u}^\circ$  since it is equivalent to six scalar equations which are, therefore, not independent.

For the uniqueness of the displacement field calculated from the tensor field  $\mathbf{C}^\circ$ , up to the rigid body movement, the strain tensor field has to satisfy the compatibility condition.

In the nonlinear theory of deformation this condition means that the Riemann–Christoffel curvature tensor, based on the tensor field  $\mathbf{C}^\circ$  as the deformed metric tensor, has to be zero, i.e., the space of continua remains Euclidean space during the deformation. In the linearized theory of deformation the system of partial differential equations

$$\underline{\underline{\mathbf{u}^\circ \circ \nabla^\circ + \nabla^\circ \circ \mathbf{u}^\circ = 2\boldsymbol{\varepsilon}^\circ}}$$

is also overdetermined for the three components of  $\mathbf{u}^\circ$ . In this case the Saint-Venant compatibility condition is

$$\underline{\underline{\nabla^\circ \times \boldsymbol{\varepsilon}^\circ \times \nabla^\circ = \mathbf{0}}}.$$

For the strain rates

$$\underline{\underline{\mathbf{v} \circ \nabla + \nabla \circ \mathbf{v} = 2\mathbf{D}}}$$

is the overdetermined system of partial differential equations for the three components of the velocity  $\mathbf{v}$ . The corresponding compatibility condition assumes the form

$$\underline{\underline{\nabla \times \mathbf{v} \times \nabla = 2\mathbf{D}}}.$$

The underlined expressions clearly show the mathematical analogy between the linearized theory of deformation and the theory of strain rates.

The problem of *independent compatibility conditions* originates from the fact that the six equations of the above-mentioned compatibility condition are not independent of each other. It can be shown that three compatibility field equations and the compatibility boundary conditions are necessary and sufficient conditions for the compatibility of strain (strain rate) field. The three independent compatibility field equations can be chosen in various ways according to a mathematical rule, i.e., they are not arbitrary.

3.5. Time derivatives of tensors

In accordance with the three remarkable description methods the Lagrangian tensor  $\mathbf{B}^\circ$ , the Eulerian tensor  $\mathbf{H}$  and the material tensor  $\hat{\mathbf{G}}$  are all functions of the corresponding co-ordinates and the time:

Lagrangian tensor	Eulerian tensor	Material tensor
$\mathbf{B}^\circ(x^{01}, x^{02}, x^{03}; t)$	$\mathbf{H}(x^1, x^2, x^3; t)$	$\hat{\mathbf{G}}(X^1, X^2, X^3; t)$

Partial derivatives with respect to time are denoted by

$\frac{\partial \mathbf{B}^\circ}{\partial t}$	$\frac{\partial \mathbf{H}}{\partial t}$	$\frac{\partial \hat{\mathbf{G}}}{\partial t}$
--	--	--

Material time derivatives are denoted and given by

$(\mathbf{B}^\circ)^\cdot = \frac{\partial \mathbf{B}^\circ}{\partial t}$	$\mathbf{H}^\cdot = \frac{\partial \mathbf{H}}{\partial t} + (\mathbf{H} \circ \nabla) \cdot \mathbf{v}$	$(\hat{\mathbf{G}})^\cdot = \frac{\partial \hat{\mathbf{G}}}{\partial t}$
---	--	---

Some remarkable material time derivatives are:

$$\begin{aligned} \mathbf{r}^\cdot &= \mathbf{v}; & (d\mathbf{r})^\cdot &= \mathbf{L} \cdot d\mathbf{r}, \\ \mathbf{F}^\cdot &= \mathbf{L} \cdot \mathbf{F}; & (\mathbf{F}^{-1})^\cdot &= -\mathbf{F}^{-1} \cdot \mathbf{L}, \\ \mathbf{E}^\cdot &= \mathbf{D} - \mathbf{E} \cdot \mathbf{L} - \mathbf{L}^T \cdot \mathbf{E} = \mathbf{D} - \mathbf{E} \cdot \mathbf{D} - \mathbf{D} \cdot \mathbf{E} - (\mathbf{E} \times \boldsymbol{\omega} - \boldsymbol{\omega} \times \mathbf{E}), \\ (\mathbf{E}^\circ)^\cdot &= \mathbf{F}^T \cdot \mathbf{D} \cdot \mathbf{F}. \end{aligned}$$

Physically (or materially) objective time derivatives (in the present configuration) are classified as follows:

A. The basic system of physically objective time derivatives, which is *invariant under arbitrary time dependent transformations*, is given below for second order tensors and for various index positions being marked by asterisks in the right hand side:

- I.  $(\mathbf{H}_{**})^\nabla = \mathbf{H} + \mathbf{L}^T + \mathbf{H} + \mathbf{H} \cdot \mathbf{L}$  (Cotter–Rivlin rate),
- II.  $(\mathbf{H}^*)^\nabla = \mathbf{H} + \mathbf{L} \cdot \mathbf{H} + \mathbf{H} \cdot \mathbf{L}$ ,
- III.  $(\mathbf{H}^*)^\nabla = \mathbf{H} + \mathbf{L}^T \cdot \mathbf{H} - \mathbf{H} \cdot \mathbf{L}^T$ ,
- IV.  $(\mathbf{H}^{**})^\nabla = \mathbf{H} - \mathbf{L} \cdot \mathbf{H} - \mathbf{H} \cdot \mathbf{L}^T$  (Oldroyd rate).

B. The physically objective time derivatives, which are *invariant under time dependent orthogonal transformations only*, are also given for second order tensors in symbolic notations; this results is independent of the index positions:

$$\mathbf{H}^\nabla = \mathbf{H} - (\boldsymbol{\omega} \times \mathbf{H} - \mathbf{H} - \boldsymbol{\omega}) \quad (\text{Jaumann rate}).$$

For  $\mathbf{E}_{**}$  it holds that

$$(\mathbf{E}_{**})^\nabla = \mathbf{E} + \mathbf{L}^T \cdot \mathbf{E} + \mathbf{E} \cdot \mathbf{L} = \mathbf{D} = (\mathbf{F}^{-1}) \cdot (\mathbf{E}^\circ) \cdot \mathbf{F}^{-1}.$$

For  $\mathbf{E}$  it holds that

$$\mathbf{E}^\nabla = \mathbf{D} - \mathbf{E} \cdot \mathbf{D} - \mathbf{D} \cdot \mathbf{E}.$$

### 3.6. Material time derivatives of integrals of tensor fields

Let the integral of the tensor field  $\mathbf{H}(x^1, x^2, x^3; t)$  be

$$\mathbf{J} = \int_{(V)} \mathbf{H} dV.$$

For the time derivative of the above integral we have

$$\begin{aligned} \left( \int_{(V)} \mathbf{H} dV \right) &= \int_{(V)} [\mathbf{H} + \mathbf{H}(\mathbf{v} \cdot \nabla)] dV = \int_{(V)} \left[ \frac{\partial \mathbf{H}}{\partial t} + \frac{\downarrow}{(\mathbf{H} \circ \mathbf{v}) \cdot \nabla} \right] dV = \\ &= \int_{(B)} \frac{\partial \mathbf{H}}{\partial t} dV + \int_{(S)} (\mathbf{H} \circ \mathbf{v}) \cdot d\mathbf{A}. \end{aligned}$$

Let the integral of the tensor field  $\mathbf{G}(x^1, x^2, x^3; t)$  be



$$\mathbf{J} = \int_{(V)} \mathbf{G} dm = \int_{(V')} \mathbf{G} \rho dV.$$

The material time derivative of this integral is as follows

$$\left( \int_{(V)} \mathbf{G} dm \right)' = \int_{(V)} \mathbf{G}' dm = \int_{(V')} \mathbf{G}' \rho dV.$$

### 3.7. Relative motion of continua

The co-ordinate systems and the co-ordinates used to describe the relative motion of a continuum are

$$\begin{array}{lll} \text{the fixed co-ordinate system} & (x) & (x) \stackrel{\text{df}}{=} (x^1, x^2, x^3), \\ \text{the grid co-ordinate system} & (\xi) & (\xi) \stackrel{\text{df}}{=} (\xi^1, \xi^2, \xi^3) \text{ and} \\ \text{the material co-ordinate system} & (X) & (X) \stackrel{\text{df}}{=} (X^1, X^2, X^3). \end{array}$$

It is assumed that the grid co-ordinate system  $(\xi)$  moves with respect to the fixed co-ordinate system  $(x)$  and the continuum itself moves with respect both to the grid co-ordinate system  $(\xi)$  and to the fixed co-ordinate system  $(x)$ . The material co-ordinate system  $(X)$  moves together with the continuum.

The corresponding motions are called and denoted as follows:

– motion of  $(\xi)$  in  $(x)$ , i.e., *the motion of grid*

$$x^p = x_G^p(\xi^1, \xi^2, \xi^3; t); \quad J_G^0,$$

– motion of  $(X)$  in  $(x)$ , i.e., *the relative or apparent motion of continuum*

$$\xi^k = \xi_C^k(X^1, X^2, X^3; t) \quad J_R^0,$$

– motion of  $(X)$  in  $(x)$ , i.e., *the absolute motion of continuum*

$$x^p = x_C^p(X^1, X^2, X^3; t) = x_G^p(\xi_C^p(X^1, X^2, X^3; t), \xi_C^2(\dots), \xi_C^3(\dots); t); \quad J = J_G^0 J_R^0.$$

If the positions of a material point  $\hat{P}$  are denoted by  $P^0$ ,  $P'$ , and  $P$  in the reference, intermediate and present configurations, respectively (see Fig. 14), then

– in the reference configuration,

$$\xi^{\circ k} = \xi^k_C(X^1, X^2, X^3; t); \quad x^{\circ p} = x^p_G(\xi^1, \xi^2, \xi^3; t_0),$$

– in the intermediate configuration

$$\xi^{i k} = \xi^{\circ k}; \quad x^{i p} = x^p_G(\xi^{\circ 1}, \xi^{\circ 2}, \xi^{\circ 3}; t),$$

– and in the present configuration

$$\xi^k = \xi^k_C(X^1, X^2, X^3; t); \quad x^p = x^p_G(\xi^1, \xi^2, \xi^3; t).$$

For the corresponding position vectors and displacements we write

$$\mathbf{r} = \mathbf{r}^{\circ} + \mathbf{u} = \mathbf{r}^{\circ} + \mathbf{u}_G + \mathbf{u}_R = \mathbf{r}' + \mathbf{u}; \quad \mathbf{r}' = \mathbf{r}^{\circ} + \mathbf{u}_G; \quad \mathbf{u} = \mathbf{u}_G + \mathbf{u}_R,$$

where

$\mathbf{u}_C$  is the absolute displacement of the material point,

$\mathbf{u}_G$  is the displacement of the corresponding grid point and

$\mathbf{u}_R$  is the relative displacement of the material point.

If

$\mathbf{F}_G$  is the deformation gradient for the motion of the grid and

$\mathbf{F}_R$  is the deformation gradient for the relative motion of continuum, then

the geometrical measure of deformation for a material line element vector is

$$d\mathbf{r} = \mathbf{F} \cdot d\mathbf{r}^{\circ} = \mathbf{F}_R \cdot \mathbf{F}_G \cdot d\mathbf{r}^{\circ}; \quad \mathbf{F} = \mathbf{F}_R \cdot \mathbf{F}_G.$$

Velocity fields in the present configuration are (the index put in parentheses on the left side denotes the co-ordinate system in which the quantity is defined)

$${}^{(x)}\mathbf{v} = {}^{(x)}\mathbf{v}_G + {}^{(\xi)}\mathbf{v}; \quad {}^{(x)}\mathbf{L} = {}^{(x)}\mathbf{L}_G + {}^{(\xi)}\mathbf{L}.$$

In the neighbourhood of  $P$  it holds that

$$d{}^{(x)}\mathbf{v} = {}^{(x)}\mathbf{L}_P \cdot d\mathbf{r} = {}^{(x)}\mathbf{L}_G \cdot d\mathbf{r} + {}^{(\xi)}\mathbf{L}_P \cdot d\mathbf{r} = {}^{(x)}\mathbf{D}_G \cdot d\mathbf{r} + {}^{(x)}\boldsymbol{\omega}_G \times d\mathbf{r} + d{}^{(\xi)}\mathbf{v}.$$

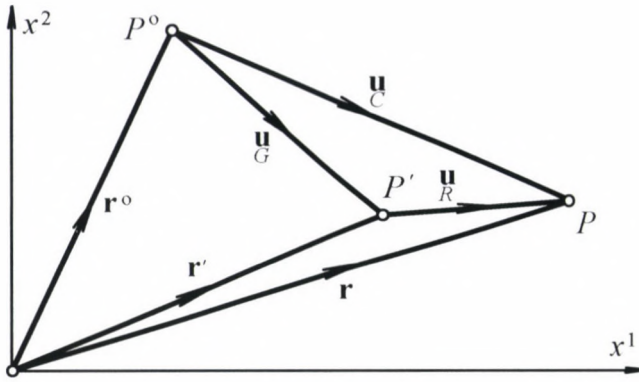


Fig. 14

The acceleration field in the present configuration is

$${}^{(x)}\mathbf{a} = {}^{(x)}\mathbf{a}_G + {}^{(\xi)}\mathbf{a} + 2 {}^{(x)}\mathbf{L}_G^{(\xi)} \mathbf{v}.$$

In the neighbourhood of  $P$ , if  ${}^{(x)}\mathbf{D}_P = \mathbf{0}$ , it holds that

$$d^{(x)}\mathbf{a} = \left[ {}^{(x)}\boldsymbol{\omega}_G \right]_P \times d\mathbf{r} + {}^{(x)}\boldsymbol{\omega}_G \times \left( {}^{(x)}\boldsymbol{\omega}_G \times d\mathbf{r} \right) + 2 {}^{(x)}\boldsymbol{\omega}_G \times d^{(\xi)}\mathbf{v} + d^{(\xi)}\mathbf{a}.$$

The mathematical form of those equations relating the material time derivatives to each other depends on the position of indices:

I. 
$$\left[ {}^{(\xi)}\mathbf{H}_{**} \right] = \left[ {}^{(x)}\mathbf{H} \right] + {}^{(x)}\mathbf{L}_G^T \cdot \mathbf{H} + \mathbf{H} \cdot {}^{(x)}\mathbf{L}_G,$$

II. 
$$\left[ {}^{(\xi)}\mathbf{H}_*^* \right] = \left[ {}^{(x)}\mathbf{H} \right] - {}^{(x)}\mathbf{L}_G \cdot \mathbf{H} + \mathbf{H} \cdot {}^{(x)}\mathbf{L}_G,$$

III. 
$$\left[ {}^{(\xi)}\mathbf{H}_*^* \right] = \left[ {}^{(x)}\mathbf{H} \right] + {}^{(x)}\mathbf{L}_G^T \cdot \mathbf{H} - \mathbf{H} \cdot {}^{(x)}\mathbf{L}_G^T,$$

IV. 
$$\left[ {}^{(\xi)}\mathbf{H}^{**} \right] = \left[ {}^{(x)}\mathbf{H} \right] - {}^{(x)}\mathbf{L}_G \cdot \mathbf{H} - \mathbf{H} \cdot {}^{(x)}\mathbf{L}_G^T.$$

If  ${}^{(x)}\mathbf{D}_G = \mathbf{0}$  then it holds independently of the position of indices that

$$\left[ {}^{(\xi)}\mathbf{H} \right] = \left[ {}^{(x)}\mathbf{H} \right] - {}^{(x)}\boldsymbol{\omega}_G \times \mathbf{H} + \mathbf{H} \times {}^{(x)}\boldsymbol{\omega}_G.$$



#### 4. General principles of continuum mechanics. Laws of classical physics

##### 4.1. Internal forces. Stress tensors

The followings are based on the *Cauchy hypothesis* (see in Fig. 9).  
*Stress tensors:*

$\mathbf{T}$  is the *Cauchy stress tensor* in the present configuration,  
 $\mathbf{K} = \lambda_V \mathbf{T}$  is the *Kirchhoff (or weighted Cauchy) stress tensor* in the present configuration,  
 $\mathbf{T}^\circ$  is the *II. Piola–Kirchhoff stress tensor* in the reference configuration and  
 $\boldsymbol{\sigma}^\circ$  is the *stress tensor in the linearized theory of deformation* (in the reference configuration).

For  $\mathbf{T}$ ,  $\mathbf{T}^\circ$  and  $\mathbf{K}$  it holds that

$$\mathbf{T} \cdot d\mathbf{A} = \mathbf{T}^\circ \cdot d\mathbf{A}^\circ; \quad \mathbf{T}^\circ = \lambda_V \mathbf{F}^{-1} \cdot \mathbf{T} \cdot (\mathbf{F}^{-1})^T = \mathbf{F}^{-1} \cdot \mathbf{K} \cdot (\mathbf{F}^{-1})^T.$$

Physically objective time derivatives:

$$\begin{aligned} (\mathbf{K}^{**})^\nabla &= \dot{\mathbf{K}} - \mathbf{L} \cdot \mathbf{K} - \mathbf{K} \cdot \mathbf{L}^T && \text{is the Oldroyd time rate,} \\ (\mathbf{T}^{**})^\nabla &= (\lambda_V \dot{\mathbf{T}}) + (\mathbf{K}^{**})^\nabla && \text{is the Truesdell time rate.} \end{aligned}$$

##### 4.2. Boundary conditions

In the nonlinear theory of deformation in the present configuration and in the linearized theory of deformation in the reference configuration

$$\begin{aligned} \mathbf{v} &= \tilde{\mathbf{v}} && x \in (A_V), \\ \mathbf{v}^\circ &= \tilde{\mathbf{v}}^\circ && x \in (A_V^\circ), \text{ are the velocity boundary conditions,} \\ \mathbf{u} &= \tilde{\mathbf{u}} && x \in (A_U), \\ \mathbf{u}^\circ &= \tilde{\mathbf{u}}^\circ && x \in (A_U^\circ), \text{ are the displacement boundary conditions and} \\ \mathbf{T} \cdot \mathbf{n} &= \tilde{\mathbf{p}} && x \in (A_T), \\ \boldsymbol{\sigma}^\circ \cdot \mathbf{n}^\circ &= \tilde{\mathbf{p}}^\circ && x \in (A_T^\circ), \text{ are the stress boundary conditions.} \end{aligned}$$

In the present configuration the general forms of the above mechanical boundary conditions are

$$(\mathbf{v} - \tilde{\mathbf{v}}) \cdot (\mathbf{T} \cdot \mathbf{n} - \tilde{\mathbf{p}}) \quad x \in (A_V) \quad (\mathbf{u} - \tilde{\mathbf{u}}) \cdot (\mathbf{T} \cdot \mathbf{n} - \tilde{\mathbf{p}}) \quad x \in (A_V).$$

For the strain rate tensor  $\mathbf{D}$  (in an arbitrary case) and for the strain tensor  $\boldsymbol{\varepsilon}^\circ$  (in the linearized theory of deformation) the boundary conditions may be expressed as

the compatibility boundary condition:  $(\nabla \times \mathbf{D} \times \nabla) \cdot \mathbf{n} = \mathbf{0}; x \in (A_t),$   
 $(\nabla^\circ \times \boldsymbol{\varepsilon}^\circ \times \nabla^\circ) \cdot \mathbf{n}^\circ = \mathbf{0}; x \in (A^\circ_t),$

the strain rate boundary condition for tensor  $\mathbf{D}$  on  $(A_v),$

the strain boundary condition for tensor  $\boldsymbol{\varepsilon}^\circ$  on  $(A^\circ_u),$

the connecting boundary condition on the common curve of  $(A_t),$  and  $(A_v)$  and on the common curve of  $(A^\circ_t)$  and  $(A^\circ_u).$

For the heat conduction problem the following boundary conditions can be prescribed:

boundary condition for the temperature:  $\Theta = \tilde{\Theta} \quad x \in (A_\Theta)$  and

boundary condition for the temperature gradient:  $(\Theta \nabla) \cdot \mathbf{n} = \tilde{\vartheta} \quad x \in (A_\vartheta)$

Here  $\Theta$  is the temperature field.

4.3. Conservation of mass. Continuity equation

In the present configuration:

$$\left. \begin{array}{l} \left( \int_{(V)} \rho dv \right)' = 0 \\ (V) \text{ is arbitrary} \\ \int_{(V)} \frac{\partial \rho}{\partial t} dV + \int_{(A)} (\rho \mathbf{v} \cdot d\mathbf{A}) = 0 \\ (V) \text{ is arbitrary} \end{array} \right\} \Rightarrow \frac{\partial \rho}{\partial t} + \downarrow (\rho \mathbf{v}) \cdot \nabla = 0.$$

$$(dm)' = 0; \Rightarrow \rho^\circ dV^\circ = \rho dV; \Rightarrow \frac{\rho^\circ}{\rho} = \frac{dV}{dV^\circ} = \lambda_V.$$

For incompressible continua:

$$\rho' = 0; \Rightarrow \mathbf{v} \cdot \nabla = D_1 = 0; \quad \lambda_V = 1.$$

4.4. Momentum balance laws

According to the *basic principles of dynamics*, the kinetic vector system [of continuum] {of an arbitrary part of continuum} is equivalent to the system of external forces acting [on continuum] {on the part considered}, i.e.,

$$\mathbf{a}\rho dV \quad x \in (V)^m = \left\{ \begin{array}{ll} \mathbf{q}dV & x \in (V) \\ \mathbf{T} \cdot d\mathbf{A} & x \in (A) \end{array} \right\} \quad (V) \text{ is arbitrary,}$$

in which  $\mathbf{a}\rho$  is the density of the kinetic vector system.

The conditions of equivalence and the equations of motion that follow from them are

$$\int_{(V)} \mathbf{a}\rho dV = \int_{(V)} \mathbf{q}dV + \int_{(A)} \mathbf{T} \cdot d\mathbf{A} \quad \Rightarrow \quad \rho\mathbf{a} = \mathbf{T} \cdot \nabla + \mathbf{q} \quad x \in (V),$$

I. Cauchy equation of motion,

$$\int_{(V)} \mathbf{r} \times \mathbf{a}\rho dV = \int_{(V)} \mathbf{r} \times \mathbf{q}dV + \int_{(A)} \mathbf{r} \times \mathbf{T} \cdot d\mathbf{A} \quad \Rightarrow \quad \mathbf{T}^T = \mathbf{T} \quad x \in (V),$$

II. Cauchy equation of motion.

The balance of mechanical energy can be obtained by transforming I. Cauchy equation of motion and then integrating the equation resulted

$$\begin{aligned} \rho \mathbf{v} \cdot \dot{\mathbf{v}} &= \mathbf{v} \cdot \mathbf{v} \cdot \overset{\downarrow}{\mathbf{T}} \cdot \nabla + \mathbf{q} \quad \Rightarrow \quad \frac{1}{2} \rho (\dot{v}^2) = -\Phi_M + \overset{\downarrow}{\mathbf{v}} \cdot \mathbf{T} \cdot \nabla + \mathbf{v} \cdot \mathbf{q} \quad \Rightarrow \\ &\Rightarrow \quad T = P_B + P_K, \end{aligned}$$

where

$$\Phi_M = \mathbf{T} \cdot \cdot \mathbf{D} = \frac{I}{\lambda_v} \mathbf{T}^\circ \cdot \cdot (\mathbf{E}^\circ) \quad \text{is the stress power per unit volume,}$$

$$\Phi_M = \frac{T_1 D_1}{3} + T_d \cdot \cdot D_d \quad \text{is the stress power in terms of the stress and strain rate deviator,}$$

$$T = \frac{1}{2} \int_{(V)} v^2 \rho dV \quad \text{is the kinetic energy of continuum,}$$

$$P_B = - \int_{(V)} \Phi_M dV \quad \text{is the power of the internal forces and}$$

$$P_K = \int_{(V)} \mathbf{v} \cdot \mathbf{q} dV + \int_{(A)} \mathbf{v} \cdot \mathbf{T} \cdot d\mathbf{A} \quad \text{is the power input (the power of the tractions and body forces).}$$



#### 4.5. The first law of thermodynamics. The energy equation

Here and in the sequel we write

$e_B \rho$  for the internal energy per unit volume,

$E_B = \int_{(V)} e_B \rho dV$  for the internal energy of continuum

$\mathbf{h}$  for the density of heat flow,

$\Phi_Q = \rho r - \mathbf{h} \cdot \nabla$  for the heat power per unit volume and

$P_Q = \int_{(V)} r \rho dA - \int_{(A)} \mathbf{h} \cdot d\mathbf{A} = \int_{(V)} \Phi_Q dV$  is the heat input.

The first law of thermodynamics and the energy equation are read

$$(T + E_B)' = P_K + P_Q \Rightarrow \rho(e_B)' = \Phi_Q + \Phi_M,$$

$$\rho(e_B)' = \mathbf{T} \cdot \cdot \mathbf{D} + \rho r - \mathbf{h} \cdot \nabla.$$

#### 4.6. The second law of thermodynamics. The entropy inequality

Let

$\Theta$  the temperature field,

$\rho s$  be the entropy per unit volume and

$S = \int_{(V)} s \rho dV$  the entropy of continuum,

respectively. Further let

$$\Phi_D = \Phi_M - \rho[(e_B)' - \Theta s] = \mathbf{T} \cdot \cdot \mathbf{D} - \rho[(e_B)' - \Theta s] \geq 0$$

be the dissipation power per unit volume.

The Clausius-Duhem inequality and their local form are as follows

$$\begin{aligned} S' - \int_{(V)} \frac{r}{\Theta} \rho dV - \int_{(A)} \frac{\mathbf{h} \Theta}{\Theta} \cdot d\mathbf{A} &\Rightarrow \\ \Rightarrow \rho \Theta s' - \rho r - \mathbf{h} \cdot \nabla + \frac{\mathbf{h} \cdot (\Theta \nabla)}{\Theta} &= \Phi_Q + \frac{\mathbf{h} \cdot (\Theta \nabla)}{\Theta}. \end{aligned}$$

In another form

$$\Phi_D - \frac{\mathbf{h} \cdot (\Theta \nabla)}{\Theta} = \mathbf{T} \cdot \cdot \mathbf{D} - \rho \left[ (e_B) \cdot - \Theta s \right] - \frac{\mathbf{h} \cdot (\Theta \nabla)}{\Theta} > 0.$$

It holds that  $\Phi_D \geq 0$  and  $\mathbf{h} \cdot (\Theta \nabla) < 0$ .

The energy equation can also be written as  $\rho \Theta s = \Phi_D + \Phi_Q$ .

#### 4.7. The energy equation and the entropy inequality in the reference configuration

$$\begin{aligned} & \rho^\circ (e_B) \cdot \mathbf{T}^\circ \cdot \cdot (\mathbf{E}^\circ) \cdot + \rho^\circ r - \mathbf{h}^\circ \cdot \nabla^\circ, \\ & \mathbf{T}^\circ \cdot \cdot (\mathbf{E}^\circ) \cdot - \rho^\circ \left[ (e_B) \cdot - \Theta s \cdot \right] - \frac{\mathbf{h}^\circ \cdot (\Theta \nabla^\circ)}{\Theta} > 0, \end{aligned}$$

in which  $\mathbf{h}^\circ$  is defined by

$$\mathbf{h} \cdot d\mathbf{A} = \lambda_V \mathbf{h} \cdot (\mathbf{F}^{-1})^T \cdot d\mathbf{A}^\circ = \mathbf{h}^\circ \cdot d\mathbf{A}^\circ.$$

#### 4.8. The principle of virtual power

In the present configuration the principle takes the form

$$\int_{(V)} \mathbf{T} \cdot \cdot \delta \mathbf{D} dV = \int_{(V)} (\mathbf{q} - \rho \mathbf{a}) \cdot \delta \mathbf{v} dV + \int_{(A_t)} \tilde{\mathbf{p}} \cdot \delta \mathbf{v} dA.$$

The corresponding side conditions are as follows:  $\delta \mathbf{v} = \mathbf{0} \quad x \in (A_v),$

$$\delta \mathbf{D} = \frac{1}{2} (\delta \mathbf{v} \circ \nabla + \nabla \circ \delta \mathbf{v}) \quad x \in (V) \text{ and}$$

$$\mathbf{T}^T = \mathbf{T} \quad x \in (V).$$

Consequently  $\mathbf{T} \cdot \nabla + \mathbf{q} - \rho \mathbf{a} = \mathbf{0} \quad x \in (V),$

$$\mathbf{T} \cdot \mathbf{n} = \tilde{\mathbf{p}} \quad x \in (A_t).$$

In total Lagrangian formulation

$$\int_{(V^\circ)} \mathbf{T}^\circ \cdot \cdot \delta (\mathbf{E}^\circ) dV^\circ = \int_{(V^\circ)} \lambda_V (\mathbf{q} - \rho \mathbf{a}) \cdot \delta \mathbf{v} dV^\circ + \int_{A_t^\circ} \lambda_{A_t} \tilde{\mathbf{p}} \cdot \delta \mathbf{v} dA^\circ$$

is the form of the principle in the reference configuration.

#### 4.9. The principle of complementary virtual power

In the present configuration the principle takes the form

$$\int_{(V)} \mathbf{D} \cdot \delta \mathbf{T} dV = \int_{(A_v)} \tilde{\mathbf{v}} \cdot \delta \mathbf{T} \cdot d\mathbf{A}.$$

The corresponding side conditions are as follows:  $\delta \mathbf{T} \cdot \mathbf{n} = \mathbf{0}$   $x \in (A_t)$   
 $\delta \mathbf{T} \cdot \nabla = \mathbf{0}$   $x \in (V)$  and  
 $\delta \mathbf{T}^T = \delta \mathbf{T}$   $x \in (A_t)$ .

From the principle it follows 3 compatibility equations on  $(V)$ ,  
 the compatibility boundary conditions on  $(A_t)$ ,  
 the strain rate boundary conditions on  $(A_v)$  and  
 the connecting boundary condition on the common curve  
 of  $(A_t)$  and  $(A_v)$ .

#### 4.10. The principle of virtual work

In the present configuration the principle has the form

$$\int_{(V)} \mathbf{T} \cdot \delta \mathbf{E}^L dV - \int_{(V)} (\mathbf{q} - \rho \mathbf{a}) \cdot \delta \mathbf{u} dV + \int_{(A_t)} \tilde{\mathbf{p}} \cdot \delta \mathbf{u} dA.$$

The corresponding side conditions are:  $\delta \mathbf{u} = \mathbf{0}$   $x \in (A_u)$ ,  
 $\delta \mathbf{E}^L = \frac{1}{2} (\delta \mathbf{u} \circ \nabla + \nabla \circ \delta \mathbf{u})$   $x \in (V)$  and  
 $\mathbf{T}^T = \mathbf{T}$   $x \in (V)$ .

The equations that follows from the principle are the those following from the principle of virtual power.

In total Lagrangian formulation

$$\int_{(V^0)} \mathbf{T}^\circ \cdot \delta \mathbf{E}^\circ dV^\circ - \int_{(V^0)} \lambda_V (\mathbf{q} - \rho \mathbf{a}) \cdot \delta \mathbf{u} dV^\circ + \int_{A_t^0} \lambda_{A_t} \tilde{\mathbf{p}} \cdot \delta \mathbf{u} dA^\circ$$

is the form of the principle in the reference configuration.

In the linearized theory of deformation we have

$$\int_{(V^0)} \boldsymbol{\sigma}^\circ \cdot \delta \boldsymbol{\epsilon}^\circ dV^\circ = \int_{(V^0)} (\mathbf{q}^\circ - \rho^\circ \mathbf{a}^\circ) \cdot \delta \mathbf{u}^\circ dV^\circ + \int_{A_t^0} \tilde{\mathbf{p}}^\circ \cdot \delta \mathbf{u}^\circ dA^\circ$$



for which

$$\begin{aligned} \delta \mathbf{u}^\circ &= \mathbf{0} \quad x \in (A_u^\circ), \\ \delta \boldsymbol{\varepsilon}^\circ &= \frac{1}{2} (\delta \mathbf{u}^\circ \circ \nabla^\circ + \nabla^\circ \circ \delta \mathbf{u}^\circ) \quad x \in (V^\circ) \text{ and} \\ \boldsymbol{\sigma}^{\circ T} &= \boldsymbol{\sigma}^\circ \quad x \in (V^\circ). \end{aligned}$$

are the side conditions and from which it follows

$$\begin{aligned} \boldsymbol{\sigma}^\circ \cdot \nabla^\circ + \mathbf{q}^\circ - \rho^\circ \mathbf{a}^\circ &= \mathbf{0} \quad x \in (V^\circ), \\ \boldsymbol{\sigma}^\circ \cdot \mathbf{n}^\circ &= \tilde{\mathbf{p}}^\circ \quad x \in (A_t^\circ). \end{aligned}$$

4.11. *The principle of complementary virtual work in the linearized theory of deformation*

In the reference configuration the principle is of the form

$$\int_{(V^\circ)} \boldsymbol{\varepsilon}^\circ \cdot \delta \boldsymbol{\sigma}^\circ dV^\circ = \int_{A_u^\circ} \tilde{\mathbf{u}}^\circ \cdot \delta \boldsymbol{\sigma}^\circ \cdot d\mathbf{A}^\circ.$$

The corresponding side conditions are  $\delta \boldsymbol{\sigma}^\circ \cdot \mathbf{n}^\circ = \mathbf{0} \quad x \in (A_t^\circ)$ ,  
 $\delta \boldsymbol{\sigma}^\circ \cdot \nabla^\circ = \mathbf{0} \quad x \in (V^\circ)$ , and  
 $\delta \boldsymbol{\sigma}^{\circ T} = \delta \boldsymbol{\sigma}^\circ \quad x \in (V^\circ)$ ,

From the principle it follows 3 compatibility equations on  $(V^\circ)$ ,  
 the compatibility boundary conditions on  $(A_t^\circ)$ ,  
 the strain boundary conditions on  $(A_u^\circ)$  and  
 the connecting boundary condition on the common curve  
 of  $(A_t^\circ)$  and  $(A_u^\circ)$ .

4.12. *The principle of virtual work in incremental total Lagrangian formulation in the reference configuration*

In accordance with Fig. 5 we shall assume that from the  $k$ -th state of continuum characterised by the displacement field  ${}^{(k)}\mathbf{u}^\circ$ , we can achieve the next state, for which there is no a separate sign, using the increment of the displacement field  $\Delta {}^{(k)}\mathbf{u}^\circ$ .

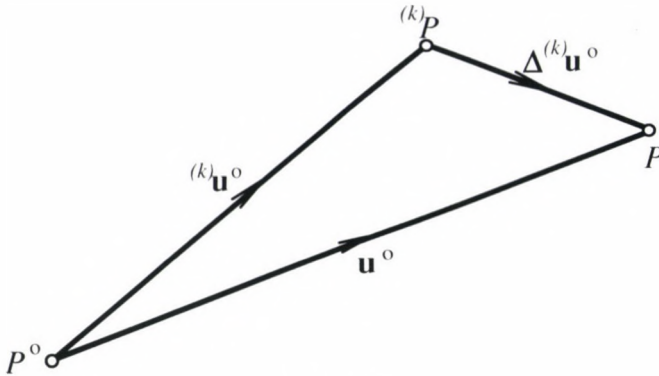


Fig. 15

It holds

– for the displacement field that

$$\mathbf{u}^\circ = {}^{(k)}\mathbf{u}^\circ + \Delta {}^{(k)}\mathbf{u}^\circ,$$

– for the Green–Lagrange strain field that

$$\begin{aligned} \mathbf{E}^\circ &= {}^{(k)}\mathbf{E}^\circ + \Delta {}^{(k)}\mathbf{E}^\circ = \frac{1}{2} \left\{ \left[ {}^{(k)}\mathbf{u}^\circ + \left( \Delta {}^{(k)}\mathbf{u}^\circ \right) \right] \circ \nabla^\circ + \nabla^\circ \circ \left[ {}^{(k)}\mathbf{u}^\circ + \left( \Delta {}^{(k)}\mathbf{u}^\circ \right) \right] \right\} + \\ &+ \frac{1}{2} \left\{ \nabla^\circ \circ \left[ {}^{(k)}\mathbf{u}^\circ + \left( \Delta {}^{(k)}\mathbf{u}^\circ \right) \right] \right\} \cdot \left\{ \left[ {}^{(k)}\mathbf{u}^\circ + \left( \Delta {}^{(k)}\mathbf{u}^\circ \right) \right] \circ \nabla^\circ \right\} = {}^{(k)}\mathbf{E}^\circ + \Delta {}^{(k)}\mathbf{E}^{\circ L} + \Delta {}^{(k)}\mathbf{E}^{\circ N}, \end{aligned}$$

$$\begin{aligned} \Delta {}^{(k)}\mathbf{E}^{\circ L} &= \frac{1}{2} \left[ \left( \Delta {}^{(k)}\mathbf{u}^\circ \right) \circ \nabla^\circ + \nabla^\circ \circ \left( \Delta {}^{(k)}\mathbf{u}^\circ \right) \right] + \\ &+ \frac{1}{2} \left\{ \left( \nabla^\circ \circ {}^{(k)}\mathbf{u}^\circ \right) \cdot \left[ \left( \Delta {}^{(k)}\mathbf{u}^\circ \right) \circ \nabla^\circ \right] + \left[ \nabla^\circ \circ \left( \Delta {}^{(k)}\mathbf{u}^\circ \right) \right] \cdot \left( {}^{(k)}\mathbf{u}^\circ \circ \nabla^\circ \right) \right\}, \end{aligned}$$

$$\Delta {}^{(k)}\mathbf{E}^{\circ N} = \frac{1}{2} \left[ \nabla^\circ \circ \left( \Delta {}^{(k)}\mathbf{u}^\circ \right) \right] \cdot \left[ \left( \Delta {}^{(k)}\mathbf{u}^\circ \right) \circ \nabla^\circ \right],$$

and for the II. Piola–Kirchhoff stress field that

$$\mathbf{T}^\circ = {}^{(k)}\mathbf{T}^\circ + \Delta {}^{(k)}\mathbf{T}^\circ.$$

It is assumed that the virtual displacement field is

$$\delta \mathbf{u}^\circ = \delta \left( \Delta {}^{(k)}\mathbf{u}^\circ \right),$$

i.e., we do not change the displacement field  ${}^{(k)}\mathbf{u}^\circ$ . If this is the case for the virtual strain field we can write

$$\delta \mathbf{E}^\circ = \delta \left( \Delta^{(k)} \mathbf{E}^\circ \right) = \delta \left( \Delta^{(k)} \mathbf{E}^{\circ L} \right) + \delta \left( \Delta^{(k)} \mathbf{E}^{\circ N} \right),$$

where

$$\begin{aligned} \delta \left( \Delta^{(k)} \mathbf{E}^{\circ L} \right) &= \frac{1}{2} \left[ \delta \left( \Delta^{(k)} \mathbf{u}^\circ \right) \circ \nabla^\circ + \nabla^\circ \circ \delta \left( \Delta^{(k)} \mathbf{u}^\circ \right) \right] + \\ &+ \frac{1}{2} \left\{ \left( \nabla^\circ \circ \Delta^{(k)} \mathbf{u}^\circ \right) \cdot \left[ \delta \left( \Delta^{(k)} \mathbf{u}^\circ \right) \circ \nabla^\circ \right] + \left[ \nabla^\circ \circ \delta \left( \Delta^{(k)} \mathbf{u}^\circ \right) \right] \cdot \left( \Delta^{(k)} \mathbf{u}^\circ \circ \nabla^\circ \right) \right\}, \\ \delta \left( \Delta^{(k)} \mathbf{E}^{\circ N} \right) &= \frac{1}{2} \left\{ \left[ \nabla^\circ \circ \left( \Delta^{(k)} \mathbf{u}^\circ \right) \right] \cdot \left[ \delta \left( \Delta^{(k)} \mathbf{u}^\circ \right) \circ \nabla^\circ \right] + \left[ \nabla^\circ \circ \delta \left( \Delta^{(k)} \mathbf{u}^\circ \right) \right] \cdot \left[ \left( \Delta^{(k)} \mathbf{u}^\circ \right) \circ \nabla^\circ \right] \right\} \end{aligned}$$

The form for the principle of virtual work we have sought is

$$\begin{aligned} &\int_{(V^\circ)} \left[ {}^{(k)}\mathbf{T}^\circ + \Delta^{(k)} \mathbf{T}^\circ \right] \cdot \delta \mathbf{E}^\circ dV^\circ = \\ &= \int_{(V^\circ)} \left[ {}^{(k)}\lambda_V + \Delta^{(k)}\lambda_V \right] \mathbf{q} \cdot \delta \mathbf{u}^\circ dV^\circ + \int_{(A_i^\circ)} \left[ {}^{(k)}\lambda_A + \Delta^{(k)}\lambda_A \right] \tilde{\mathbf{p}} \cdot \delta \mathbf{u}^\circ dA^\circ. \end{aligned}$$

Since the product  $\Delta^{(k)} \mathbf{T}^\circ \cdot \delta \left( \Delta^{(k)} \mathbf{E}^{\circ N} \right)$  is not linear in  $\Delta^{(k)} \mathbf{T}^\circ$ ,  $\left( \Delta^{(k)} \mathbf{u}^\circ \right) \circ \nabla^\circ$  we neglect it. We also neglect the terms  $\Delta^{(k)} \lambda_V \mathbf{q}$  and  $\Delta^{(k)} \lambda_A \tilde{\mathbf{p}}$ . In this way we obtain

$$\begin{aligned} &\int_{(V^\circ)} \Delta^{(k)} \mathbf{T}^\circ \cdot \delta \left( \Delta^{(k)} \mathbf{E}^{\circ L} \right) dV^\circ + \int_{(V^\circ)} {}^{(k)}\mathbf{T}^\circ \cdot \delta \left( \Delta^{(k)} \mathbf{E}^{\circ N} \right) dV^\circ = \\ &= - \int_{(V^\circ)} {}^{(k)}\mathbf{T}^\circ \cdot \delta \left( \Delta^{(k)} \mathbf{E}^{\circ L} \right) dV^\circ + \\ &+ \int_{(V^\circ)} {}^{(k)}\lambda_V \mathbf{q} \cdot \delta \left( \Delta^{(k)} \mathbf{u}^\circ \right) dV^\circ + \int_{(A_p^\circ)} {}^{(k)}\lambda_A \tilde{\mathbf{p}} \cdot \delta \left( \Delta^{(k)} \mathbf{u}^\circ \right) dA^\circ. \end{aligned}$$

Due to the neglects we have made the latter form of the principle of virtual work applies to an unbalanced state of the body and for this reason the solution (the equilibrium state of the body) can be sought, starting from a given state, by a series of subse-



quent iteration steps. We assume that the  $k$ -th state precedes the  $k$ -th iteration step. For the iteration solution an error limit should be defined and an appropriate form of the constitutive equations should also be chosen for the body under consideration.

In the first step, the field  ${}^{(1)}\mathbf{u}^\circ$ ,  ${}^{(1)}\mathbf{E}^\circ$  and  ${}^{(1)}\mathbf{T}^\circ$  are given, the geometrical measures of deformation are units ( ${}^{(1)}\lambda_{V'} = {}^{(1)}\lambda_{A'} = 1$ ) and the displacement increment  $\Delta {}^{(1)}\mathbf{u}^\circ$  is the unknown.

In the  $k$ -th step of the iteration the displacement increment  $\Delta {}^{(k)}\mathbf{u}^\circ$  is the unknown while the increment of the stress tensor  $\Delta {}^{(k)}\mathbf{T}^\circ$  is calculated from the increment of the strain tensor  $\Delta {}^{(k)}\mathbf{E}^{\circ L}$  using a linear approximation .

After the  $k$ -th iteration step we have

$$\begin{aligned} {}^{(k+1)}\mathbf{u}^\circ &= {}^{(k)}\mathbf{u}^\circ + \Delta {}^{(k)}\mathbf{u}^\circ, \\ {}^{(k+1)}\mathbf{E}^\circ &= {}^{(k)}\mathbf{E}^\circ + \Delta {}^{(k)}\mathbf{E}^\circ = {}^{(1)}\mathbf{E}^\circ + \sum_{i=1}^k \left( \Delta {}^{(i)}\mathbf{E}^{\circ L} + \Delta {}^{(i)}\mathbf{E}^{\circ N} \right), \\ {}^{(k+1)}\mathbf{T}^\circ &= {}^{(k)}\mathbf{T}^\circ + \Delta {}^{(k)}\mathbf{T}^\circ. \end{aligned}$$

In the last expression the total increment  $\Delta {}^{(k)}\mathbf{T}^\circ$  (not only the linear part of it) has to be calculated. The geometrical measures of deformation  ${}^{(k+1)}\lambda_{V'}$  and  ${}^{(k+1)}\lambda_{A'}$  have to be determined from  ${}^{(k+1)}\mathbf{u}^\circ$ .

## 5. Constitutive equations

Constitutive equations relate to each other the variables describing the macroscopic behaviour of a continuum. These are the so-called phenomenological variables which depend on the material.

In the sequel we are concerned with thermomechanical models of materials.

At the point of continuum, identified by the position vector  $\mathbf{r}^\circ$  in the reference configuration, the independent variables of the constitutive equations are the motion  $\mathbf{r} = \mathbf{r}(\mathbf{r}^\circ; t)$ , the temperature field  $\Theta = \Theta(\mathbf{r}^\circ; t)$  and those quantities being calculated from these variables (for instance the strain tensors, the strain rate tensor, and the gradient of the temperature field).

The dependent variables of the constitutive equations are the stress tensor  $\mathbf{T}$  (or  $\mathbf{T}^\circ$ ), the heat flow per unit area  $\mathbf{h}$ , the free energy  $f$  and the entropy  $s$  (or the internal energy  $e_B$  and the entropy  $s$ ).

The free energy per unit volume related to  $e_B$  and  $s$ :  $f = e_B - \Theta s$ .

It also holds that

$$\Phi_D = \Phi_M - \rho \left[ (e_B)^\cdot - \Theta s^\cdot \right] = \Phi_M - \rho (f^\cdot + s \Theta^\cdot) = \mathbf{T} \cdot \mathbf{D} - \rho (f^\cdot + s \Theta^\cdot),$$

$$\Phi_M - \rho (f^\cdot + s \Theta^\cdot) - \frac{\mathbf{h} \cdot (\Theta \nabla)}{\Theta} > 0.$$

### 5.1. Thermoelastic body

The dissipative power is zero, i.e.,

$$\Phi_D = 0,$$

$$\mathbf{T} \cdot \mathbf{D} - \rho (f^\cdot + s \Theta^\cdot) = 0 \quad \mathbf{T}^\circ \cdot \cdot (\mathbf{E}^\circ)^\cdot - \rho^\circ (f^\cdot + s \Theta^\cdot) = 0.$$

The free energy depends on  $\mathbf{E}^\circ$  and  $\Theta$  (or  $\mathbf{E}$  and  $\Theta$ ) only:

$$f = f(\mathbf{E}^\circ, \Theta) \quad \text{or} \quad f = f(\mathbf{E}, \Theta).$$

The constitutive equations (11 scalar equations) assume the form:

$$\mathbf{T}^\circ = \rho^\circ \frac{\partial f}{\partial \mathbf{E}^\circ} = \rho^\circ \left( \frac{\partial e_B}{\partial \mathbf{E}^\circ} - \Theta \frac{\partial s}{\partial \mathbf{E}^\circ} \right) \quad \text{or} \quad \mathbf{T} = \rho \frac{\partial f}{\partial \mathbf{E}} = \rho \left( \frac{\partial e_B}{\partial \mathbf{E}} - \Theta \frac{\partial s}{\partial \mathbf{E}} \right),$$

$$f = f(\mathbf{E}^\circ, \Theta), \quad s = -\frac{\partial f}{\partial \Theta} \quad \text{and} \quad \text{the law of heat conduction for } \mathbf{h}.$$

The equation of heat conduction is

$$\mathbf{h} \cdot \nabla - \rho \Theta \left[ \frac{\partial^2 f}{\partial \Theta \partial \mathbf{E}^\circ} \cdot \cdot (\mathbf{E}^\circ)^\cdot + \frac{\partial^2 f}{\partial \Theta^2} \Theta^\cdot \right] = 0.$$

In the linearized theory of deformation the constitutive equations may be expressed as

$$\sigma^\circ = \rho^\circ \frac{\partial f}{\partial \boldsymbol{\varepsilon}^\circ},$$

$$f = f(\Theta, \rho^{-1}), \quad s = -\frac{\partial f}{\partial \Theta}, \quad \mathbf{h} = -\kappa \cdot (\Theta \nabla).$$

The constitutive equations for a *geometrically and physically linear, anisotropic and thermoelastic body* are of the form

$$\rho^\circ f(\boldsymbol{\varepsilon}^\circ, \Theta) = \frac{1}{2} \boldsymbol{\varepsilon}^\circ \cdot \cdot \cdot {}^{(4)}X^\circ \cdot \cdot \cdot \boldsymbol{\varepsilon}^\circ - \boldsymbol{\beta} \cdot \cdot \boldsymbol{\varepsilon}^\circ (\Theta - \Theta^\circ) - \frac{1}{2} c (\Theta - \Theta^\circ)^2,$$

$$\boldsymbol{\sigma}^\circ = {}^{(4)}C^\circ \cdot \cdot \cdot \boldsymbol{\varepsilon}^\circ - \boldsymbol{\beta} (\Theta - \Theta^\circ), \quad \rho^\circ s = \boldsymbol{\beta} \cdot \cdot \boldsymbol{\varepsilon}^\circ + c (\Theta - \Theta^\circ),$$

$$\mathbf{h} = -\kappa \cdot (\Theta \nabla).$$

Only 21 components of the fourth order tensor  ${}^{(4)}C^\circ$  can be independent.

For *geometrically and physically linear, isotropic and thermoelastic body* the constitutive equations become simplified:

$$\rho^\circ f(\boldsymbol{\varepsilon}^\circ, \Theta) = \mu \boldsymbol{\varepsilon}^\circ \cdot \cdot \boldsymbol{\varepsilon}^\circ + \frac{1}{2} \lambda (\varepsilon_1^\circ)^2 - \beta \varepsilon_1^\circ (\Theta - \Theta^\circ) - \frac{1}{2} c (\Theta - \Theta^\circ)^2,$$

$$\boldsymbol{\sigma}^\circ = 2\mu \boldsymbol{\varepsilon}^\circ + \lambda \varepsilon_1^\circ \mathbf{I} - \beta (\Theta - \Theta^\circ) \mathbf{I}; \quad \rho^\circ s = \beta \varepsilon_1^\circ + c (\Theta - \Theta^\circ),$$

$$\mathbf{h} = -\kappa (\Theta \nabla).$$

$$\boldsymbol{\varepsilon}^\circ = \frac{1}{2\mu} \left( \boldsymbol{\sigma}^\circ - \frac{\lambda}{2\mu + 3\lambda} \sigma_1^\circ \mathbf{I} \right) + \frac{\beta}{2\mu + 3\lambda} (\Theta - \Theta^\circ) \mathbf{I}.$$

After being rewritten in terms of  $G$  and  $\nu$  the *Hooke's law* takes the form

$$\boldsymbol{\sigma}^\circ = 2G \left( \boldsymbol{\varepsilon}^\circ + \frac{\nu}{1-2\nu} \varepsilon_1^\circ \mathbf{I} \right) - \beta (\Theta - \Theta^\circ) \mathbf{I},$$

$$\boldsymbol{\varepsilon}^\circ = \frac{1}{2G} \left( \boldsymbol{\sigma}^\circ - \frac{\nu}{1+\nu} \sigma_1^\circ \mathbf{I} \right) + \alpha (\Theta - \Theta^\circ) \mathbf{I},$$

in which

$$\sigma_1^\circ = 3K \varepsilon_1^\circ - 3\beta (\Theta - \Theta^\circ), \quad \varepsilon_1^\circ = \frac{1}{3K} \sigma_1^\circ + 3\alpha (\Theta - \Theta^\circ).$$



### 5.2. Elastic body

For an elastic body the entropy  $s$  is constant,  $\nabla\Theta = \mathbf{0}$  and  $\mathbf{h} = \mathbf{0}$ .

If  $e_B = e_B(E_{11}^\circ, E_{12}^\circ, \dots, E_{33}^\circ)$ , or  $e_B = e_B(E_{\text{I}}, E_{\text{II}}, \dots, E_{\text{III}})$ , then the constitutive equations (6 scalar equations) may be expressed as

$$\mathbf{T}^\circ = \rho^\circ \frac{\partial e_B}{\partial \mathbf{E}^\circ} \quad \text{or} \quad \mathbf{T} = \rho \frac{\partial e_B}{\partial \mathbf{E}}.$$

In the general case

$$\mathbf{T}^\circ = a_\circ^\circ \mathbf{I} + a_1^\circ \mathbf{E}^\circ + a_2^\circ (\mathbf{E}^\circ)^2 \quad \text{or} \quad \mathbf{T} = a_\circ \mathbf{I} + a_1 \mathbf{E} + a_2 \mathbf{E}^2,$$

where  $a_\circ^\circ$ ,  $a_1^\circ$ ,  $a_2^\circ$  and  $a_\circ$ ,  $a_1$ ,  $a_2$  depend on the scalar invariants.

For an isotropic elastic body  $e_B = e_B(E_{\text{I}}^\circ, E_{\text{II}}^\circ, E_{\text{III}}^\circ)$ , or  $e_B = e_B(E_{\text{I}}, E_{\text{II}}, E_{\text{III}})$ .

For the constitutive equations for a geometrically and physically linear elastic body we can write

$$\boldsymbol{\sigma}^\circ = {}^{(4)}\mathbf{C}^\circ \cdot \boldsymbol{\varepsilon}^\circ; \quad \boldsymbol{\varepsilon}^\circ = ({}^{(4)}\mathbf{C}^\circ)^{-1} \cdot \boldsymbol{\sigma}^\circ,$$

if the body is anisotropic and

$$\boldsymbol{\sigma}^\circ = 2\mu \boldsymbol{\varepsilon}^\circ + \lambda \boldsymbol{\varepsilon}_1^\circ \mathbf{I} = 2G \left( \boldsymbol{\varepsilon}^\circ + \frac{\nu}{1-2\nu} \boldsymbol{\varepsilon}_1^\circ \mathbf{I} \right),$$

$$\boldsymbol{\varepsilon}^\circ = \frac{1}{2\mu} \left( \boldsymbol{\sigma}^\circ - \frac{\lambda}{2\mu + 3\lambda} \boldsymbol{\sigma}_1^\circ \mathbf{I} \right) = \frac{1}{2G} \left( \boldsymbol{\sigma}^\circ - \frac{\nu}{1+\nu} \boldsymbol{\sigma}_1^\circ \mathbf{I} \right).$$

if the body is isotropic (Hooke's law).

### 5.3. Thermoviscous fluids

The stress tensor is  $\mathbf{T} = -p \mathbf{I} + \mathbf{T}^V$ .

The constitutive equations (12 scalar equations) are

$$p = p(\Theta, \rho^{-1}), \quad \mathbf{T}^V = \mathbf{T}^V(\Theta, \nabla\Theta, \mathbf{D}, \rho^{-1}), \quad (\mathbf{T}^V = \mathbf{0}, \text{ if } \mathbf{D} = \mathbf{0}),$$

$$f = f(\Theta, \rho^{-1}), \quad s = -\frac{\partial f}{\partial \Theta} \quad \text{and} \quad \text{the law of heat conduction for } \mathbf{h}.$$

The dissipative power is  $\Phi_D = \mathbf{T}^V \cdot \mathbf{D}$ .

The equation of heat conduction is

$$\mathbf{h} \cdot \nabla + \rho \Theta \left( \frac{1}{\rho} \frac{\partial p}{\partial \Theta} D_1 - \frac{\partial^2 f}{\partial \Theta^2} \Theta \right) - \mathbf{T}^V \cdot \cdot \mathbf{D} - r \rho = 0.$$

*The Stokesian linear thermoviscous fluids:*

*In anisotropic case:*

$$p = p(\Theta, \rho^{-1}), \quad \mathbf{T}^V = {}^{(4)}C^V \cdot \cdot \mathbf{D},$$

$$f = f(\Theta, \rho^{-1}), \quad s = -\frac{\partial f}{\partial \Theta} \quad \text{and} \quad \mathbf{h} = -\boldsymbol{\kappa} \cdot (\nabla \Theta),$$

where  ${}^{(4)}C^V = {}^{(4)}C^V(\Theta, \rho^{-1})$  and  $\boldsymbol{\kappa} = \boldsymbol{\kappa}(\Theta, \rho^{-1})$ .

In isotropic case:

$$p = p(\Theta, \rho^{-1}), \quad \mathbf{T}^V = 2\mu^V \mathbf{D} + \lambda^V D_1 \mathbf{I},$$

$$f = f(\Theta, \rho^{-1}), \quad s = -\frac{\partial f}{\partial \Theta} \quad \text{and} \quad \mathbf{h} = -\boldsymbol{\kappa}(\nabla \Theta),$$

where  $\mu^V = \mu^V(\Theta, \rho^{-1})$ ,  $\lambda^V = \lambda^V(\Theta, \rho^{-1})$  and  $\boldsymbol{\kappa} = \boldsymbol{\kappa}(\Theta, \rho^{-1})$ .

#### 5.4. Isotherm viscous fluids

Constitutive equations for viscous, isotherm, nonlinear model of fluids may be expressed as

$$p = p(\rho^{-1}); \quad \mathbf{T}^V = a_0^V \mathbf{I} + a_1^V \mathbf{D} + a_2^V \mathbf{D}^2,$$

where  $a_i^V = a_i^V(\rho, D_1, D_{II}, D_{III})$ .

Constitutive equation for viscous, isotherm, linear (Newton type) model of fluids are

$$\mathbf{T} = -p \mathbf{I} + \lambda^V \mathbf{D} + 2\mu^V \mathbf{D}; \quad 3\lambda^V + 2\mu^V \geq 0, \quad \mu^V > 0.$$

#### 5.5. Ideal fluids. Elastic fluids

For ideal fluids  $\mathbf{T}^V = \mathbf{0}$ ,

$$\mathbf{T} = -p \mathbf{I}; \quad \rho = \text{constant}.$$

For elastic fluids  $T^v = \mathbf{0}$  and  $\rho \neq \text{constant}$ ,

$$\mathbf{T} = -p\mathbf{I}; \quad p = p(\rho^{-1}).$$

5.6. Elastic, plastic body. Flow theory with isotropic kinematic hardening

It can be assumed that the deformation can be decomposed to a plastic and an elastic part (see Fig. 16).

$$\mathbf{F} = \mathbf{F}_{el} \cdot \mathbf{F}_{pl}; \quad \mathbf{D} = \mathbf{D}_{el} + \mathbf{D}_{pl} = (\mathbf{D}_{del} + \mathbf{D}_{iel}) + \mathbf{D}_{pl}; \quad \mathbf{D}_{ipl} = \mathbf{0}.$$

The yield condition is of the form (see Fig. 17)

$$\mathbf{Q} = \mathbf{K}_d - \mathbf{s} \quad Q_1 = K_{d1} = s_1 = 0,$$

$$f(\mathbf{Q}) = \sqrt{2Q_{II}} = \sqrt{\mathbf{Q} \cdot \mathbf{Q}} = \sqrt{\frac{2}{3}} k(\vartheta),$$

where  $\mathbf{s}$  is the shift of the centre of the yield surface,  $k$  is a material constant (yield stress),  $\vartheta$  is the work of plastic deformation (per unit volume) defined by:

$$\vartheta = \int_0^t \mathbf{Q} \cdot \mathbf{D}_{pl} dt.$$

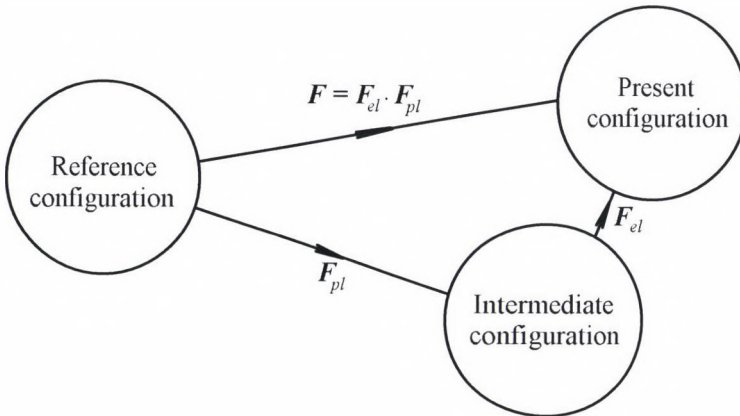


Fig. 16



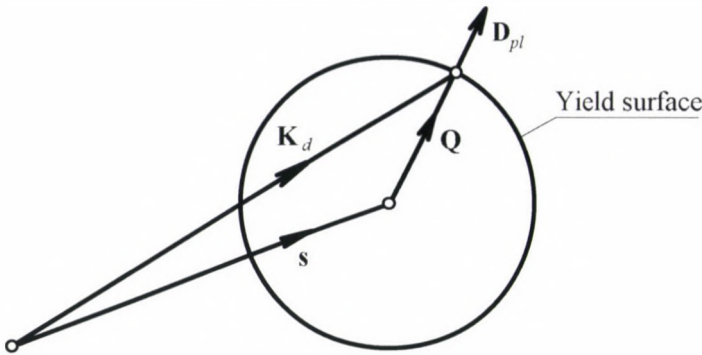


Fig. 17

*Prandtl–Reuss type elastic equations of plasticity are as follows:*

– the elastic part of the strain rate tensor:

$$D_{iel} = D_i = \frac{1}{3K} (\mathbf{K}_i^{**})^\nabla; \quad D_{del} = D_d - D_{pl} = \frac{1}{2G} (\mathbf{K}_d^{**})^\nabla,$$

– the plastic part of the strain tensor:

$$D_{pl} = D_{dpl} = \gamma \cdot \frac{\partial f(\mathbf{Q})}{\partial \mathbf{Q}} = \gamma \cdot \frac{\mathbf{Q}}{f(\mathbf{Q})},$$

– the velocity of the centre of the yield surface:

$$(\mathbf{s}^{**})^\nabla = \frac{2}{3} k(\vartheta) \frac{dh(\vartheta)}{d\vartheta} D_{pl},$$

and the consistency condition:

$$\gamma \cdot = \frac{1}{H} \frac{\partial f(\mathbf{Q})}{\partial \mathbf{Q}} \cdot \cdot (\mathbf{Q}^{**})^\nabla; \quad H = \frac{2}{3} k(\vartheta) \frac{dk(\vartheta)}{d\vartheta}.$$

*Constitutive equation for the Kirchhoff stress tensor in the present configuration:*

$$(\mathbf{K}^{**})^\nabla = {}^{(4)}C_{ep} \cdot \cdot (\mathbf{E}^{**})^\nabla = {}^{(4)}C_{ep} \cdot \cdot \mathbf{D},$$

in which

$${}^{(4)}C_{ep} = K(I \circ I) + 2G \left[ \frac{1}{2} {}^{(4)}I - \frac{1}{3} (I \circ I) \right] - \frac{2G}{\lambda} \left( \frac{\mathbf{Q}}{f(\mathbf{Q})} \circ \frac{\mathbf{Q}}{f(\mathbf{Q})} \right),$$

$${}^{(4)}I = (g^{pr} g^{qs} + g^{ps} g^{qr}) (\mathbf{g}_p \circ \mathbf{g}_q \circ \mathbf{g}_r \circ \mathbf{g}_s),$$

$$\lambda = 1 + \frac{k(\vartheta)}{3G} \left[ \frac{dk(\vartheta)}{d\vartheta} + \frac{dh(\vartheta)}{d\vartheta} \right]$$

and in the reference configuration:

$$(\mathbf{T}^\circ)^\cdot = {}^{(4)}C_{ep}^\circ \cdot \cdot (\mathbf{E}^\circ)^\cdot,$$

in which

$${}^{(4)}C_{ep}^\circ = \left[ (\mathbf{F}^{-1}) \circ (\mathbf{F}^{-1}) \right] \cdot \cdot {}^{(4)}C_{ep} \cdot \cdot \left[ (\mathbf{F}^{-1})^T \circ (\mathbf{F}^{-1})^T \right].$$

Observe that the latter constitutive equations were determined using the approach of *objective time derivatives*.

# TANGENT MODULUS FOR NONSMOOTH MATERIALS

Kurutz, M.

*Technical University of Budapest, Department of Structural Mechanics  
Műegyetem rkp. 3, H-1521 Budapest, Hungary*

(Received: 15 July 1996)

Nonsmooth characteristics of the tangent modulus resulted both from irreversible behaviour of materials and from polygonal approximation of nonlinear material behaviour will be analysed.

Stability analyses of inelastic irreversible systems are based on the concept of tangent modulus. The switch from loading to unloading in the material behaviour results nonsmooth material functions. On the other hand, for a load history analysis instantaneous tangent modulus is needed. Thus, by applying polygonal approximation, we obtain nonsmooth material behaviour again. Naturally, originally polygonal material behaviour can also be happened. Consequently, for global stability analyses related to the total domain of possible deflections, nonsmooth tangent modulus is needed.

In this paper, the concept of the tangent modulus is extended to materials having nonsmooth characteristics due to both inelastic loading-unloading and polygonal form.

## 1. Introduction

The classical elastic stability analyses are based on the convex and smooth elastic potential. The nearly also classical elastic-plastic stability analyses need also smooth and convex potential by using the concept of the linear comparison solid by reducing the problem to a quasi-elastic analysis. The modern inelastic stability analyses including strain softening and damage are extended to nonconvex potentials by the generalization of the tangent modulus. However, the condition of smoothness is further on required by using the concept of linear comparison solid.

The nonsmooth characteristics of strain energy functionals can be derived from two facts. On the one hand, it can be caused by the change of material phases: the switch over of loading-unloading is an original nonsmooth characteristics of any inelastic behaviour. Even to avoid nonsmoothness, the concept of linear comparison solid has been introduced by Hill [10]. On the other hand, the nonsmooth properties of strain energy can be resulted by nonsmooth functions of material laws directly, due to original or approximate polygonal constitutive laws.

By using the tools of nonsmooth analysis, we can get over the difficulties of nonsmoothness of both type. In this paper an extension, the so-called *nonsmooth tangent*



*modulus* is introduced. By the help of it, global stability analyses of nonlinearly inelastic structures can be investigated. Equilibrium paths of structures having arbitrary nonlinear irreversible materials can be analysed by applying polygonal approximation.

## 2. Short history of the tangent modulus

Since the tangent modulus plays a key role in dissipative stability analyses, it seems to be obvious to run through the development of it. Thus we can see how the tangent modulus changed during about a century long time period how it became from a simple material constant to an indicator tensor of dissipative systems.

### 2.1. Appearance of the tangent modulus

The concept of tangent modulus is resulted by the development of plasticity, namely, the plastic bifurcation problems. In the history of the tangent modulus we follow the state of art given by Bruhns in [6]. Plastic column buckling, the possible bifurcation of the structure was in the focus of interest in the past, however, the progress in the subject was not smooth.

The problem to calculate the critical value of the load on the top of the column, when the straight configuration becomes unstable, was first solved by Euler in 1744, by assuming linear elastic material. However if the stress in the column exceeds the yield limit, plastic flow will occur, and the original elastic modulus is not valid any more.

A typical stress-strain curve  $\sigma = \sigma(\varepsilon)$  can be seen in Fig. 1a where the slope

$$E_t = d\sigma / d\varepsilon \quad (1)$$

of the curve beyond the elastic limit  $\sigma_o$  is the tangent modulus. The tangent modulus is still the function of strains, since  $E_t(\varepsilon) = d\sigma(\varepsilon) / d\varepsilon$ , thus for the stress-strain function, a bilinear idealization seen in Fig. 1b is often used. This model seems to be advantageous since  $E_t = const$ , however, the nonsmoothness in  $\sigma(\varepsilon)$ , namely, the jump in the slope passing through the yield limit, causes difficulties.

The first revised formula for the critical load was suggested by Engesser in 1889 by using the tangent modulus  $E_t$  in the formula of Euler. A bit later, in 1891 by Considère and in 1894 by Jasinski, an important observation was made by pointing out that in a buckling mode, in one part of the section a purely elastic strain reversal will occur, while in the remaining part, plastic loading will continue. Considère introduced the so-called reduced modulus  $E_r$  for which  $E \geq E_r \geq E_t$ . Some examples were performed later by Kármán in 1910.

Then, for a long period, the reduced modulus theory was accepted as the exact solution of the problem, while the tangent modulus theory was considered as an approximate approach. Until 1946–47 [39, 40] there was no considerable progress in the subject.

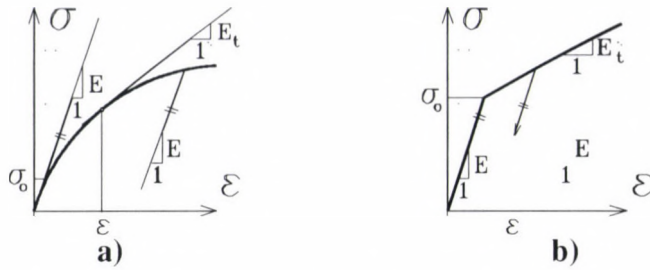


Fig. 1. The tangent modulus and the bilinear idealization

Then, Shanley in [40], by the help of a simple model demonstrated the important distinction between uniqueness and stability. He recognised that the tangent modulus load is the lowest possible bifurcation load. At this load, the straight configuration loses its uniqueness but not its stability. Moreover, Shanley has written the tangent modulus in the form in which it is used in our time too

$$\delta\sigma = \begin{cases} E \delta\varepsilon & \text{for elastic loading or elastic-plastic unloading} \\ E_t \delta\varepsilon & \text{for plastic loading} \end{cases} \quad (2)$$

By an other important observation of Shanley that at the instant of bifurcation there is no change in the load, Shanley quasi-made an advance of the grounds of the concept of linear comparison solid of Hill. However, it took another decade until the continuum theory of bifurcation was laid down by the fundamental paper of Hill in 1958.

2.2. The tangent modulus in the continuum theory

Until the famous paper of Hill in [10], the tangent modulus was considered as the property of a material point only. Hill was who extended the concept of the tangent modulus to the whole body or structure by characterising the “resistance” of the body by the tangent modulus [11–13].

By following Bruhns [6], here we refer only to those results of Hill which are in closed connection with the tangent modulus. Hill suggested for the rate constitutive relations of bodies with elastic-plastic material and finite deformations as follows

$$\dot{\sigma} = E\dot{\varepsilon} - \frac{\alpha}{h}(\lambda\dot{\varepsilon})\lambda = \dot{\sigma}(\dot{\varepsilon}) \quad (3)$$

in the case of smooth yield surface and associated plastic deformations. Here  $\dot{\sigma}$  is the tensor of the so-called objective stress increments and  $E$  is the tensor of the instantaneous elastic moduli while  $\lambda$  represents the normals to the yield hypersurface interfaces separating the domains of elastic and plastic behaviour. Here  $\alpha$  is a positive function of hardening as follows



$$\alpha = \begin{cases} 1 & \text{if } \lambda \dot{\epsilon} \geq 0 \\ 0 & \text{if } \lambda \dot{\epsilon} < 0 \end{cases} \quad (4)$$

as an indicator of the regimes of plastic loading and unloading. When the stress lies within the yield surface, the material is purely elastic thus  $\alpha = 0$ .

On the basis of Hill's tangent modulus, a strain rate potential

$$W(\dot{\epsilon}) = \frac{1}{2} \dot{\epsilon} E \dot{\epsilon} - \frac{1}{2} \frac{\alpha}{h} (\lambda \dot{\epsilon})^2 = W_e(\dot{\epsilon}) + W_p(\dot{\epsilon}) \quad (5)$$

can be introduced as the potential function of the stress rates

$$\dot{\sigma} = \frac{\partial W(\dot{\epsilon})}{\partial \dot{\epsilon}} = E \dot{\epsilon} - \frac{\alpha}{h} \lambda \dot{\epsilon} \lambda \quad (6)$$

However, this function is nonsmooth with respect to the strain rates. It has continuous first derivative and sectionally continuous second derivative. The jump in the second derivative is due to the jump in the indicator  $\alpha$ . Without using the tools of nonsmooth analysis the nonsmooth potential function (5) cannot be handled.

Here we recall the observation of Shanley according to which in column buckling at the instant of bifurcation unloading is absent. In this sense Hill introduces a special material called *linear comparison solid* of the nonlinearly inelastic material with the property that unloading is excluded through  $\alpha = 1$ . Thus, he could avoid the nonsmoothness by obtaining a smooth potential

$$W(\dot{\epsilon}) = \frac{1}{2} \dot{\epsilon} E \dot{\epsilon} - \frac{1}{2h} (\lambda \dot{\epsilon})^2 = W_e(\dot{\epsilon}) + W_p(\dot{\epsilon}) \quad (7)$$

Hill's results have a great importance. Any solution of the rate boundary value problem for the elastic-plastic solid is unique when the uniqueness of the analogous boundary value problem for the comparison solid is assured.

Hill's results are related to elastic-plastic behaviour. However, in modern stability analyses the strain softening and damage, moreover, the strain localization became even more important. Thus, the results of Hill have recently been extended to these cases, on the basis of the thermodynamics and by using the tools of functional analysis.

### 2.3. Thermodynamic generalization of the tangent modulus

The thermodynamic extension of the tangent modulus is the merit of Nguyen [28, 29], Halphen and Nguyen [9] by introducing the *generalized time-independent standard material*. This concept is the basis of the modern bifurcation theories. The most general constitutive relations of strain softening materials is given by the authors Rice, Raniecki, Rudnicki and Bruhns in [34, 35, 36, 37]. The newest modern mathematical description is given by Benallal, Billardon, Geymonat and Doghri [3, 4, 5].



Assuming small isothermal strains according to [3, 4], the behaviour of the generalized standard materials can be characterised by the *free energy* and the *domain of reversibility* and the *consistency condition*.

Generally the *free energy* is a function of the strains  $\varepsilon$ , the internal variables  $\alpha$ , and the temperature  $T$

$$\psi = \psi(\varepsilon, \alpha, T) \quad (8)$$

where function  $\psi$  is the potential function of the thermodynamic state variables, the stresses  $\sigma$ , the thermodynamic forces  $A$ , and the entropy  $s$  as follows

$$\sigma = \rho \frac{\partial \psi}{\partial \varepsilon} \quad A = -\rho \frac{\partial \psi}{\partial \alpha} \quad s = -\rho \frac{\partial \psi}{\partial T} \quad (9)$$

The irreversible behaviour is characterised by the *domain of reversibility*

$$C(\alpha) = \{A \mid f(A, \alpha, T) \leq 0\} \quad (10)$$

and the *potential of dissipation*  $F(A, \dot{\alpha}, T)$  in connection with the normality law

$$\dot{\alpha} = \lambda \frac{\partial F}{\partial A} = N_c(A) \quad (11)$$

where  $N_c(A)$  is the outer normal vector of the convex set  $C(\alpha)$  at  $A$ .

The nonnegative multiplier  $\lambda \geq 0$  results from the *consistency condition*  $\dot{f} = 0$ .

$$\lambda = \left\langle \frac{\frac{\partial f}{\partial A} \circ \Lambda : \dot{\varepsilon}}{h} \right\rangle \quad (12)$$

where

$$h = \frac{\partial f}{\partial A} \circ \Pi \circ \frac{\partial F}{\partial A} - \frac{\partial F}{\partial \alpha} \circ \frac{\partial F}{\partial A} \rangle 0 \quad (13)$$

and

$$\Lambda = -\rho \frac{\partial^2 \psi}{\partial \alpha \partial \varepsilon} \quad \Pi = \rho \frac{\partial^2 \psi}{\partial \alpha \partial \alpha} \quad (14)$$

Here the symbol  $\circ$  denotes the scalar product while  $\otimes$  denotes the tensorial product between tensors. Symbol “:” denotes the double tensor contraction, and  $\langle x \rangle = \max \langle x, 0 \rangle$  ensures the nonnegativity.

Considering an *isothermal* process of the elastic-plastic damaging strain softening material, the temperature and the entropy can be eliminated, thus, the thermodynamic state laws (9) read

$$\psi = \psi(\varepsilon, \alpha) \quad \sigma = \rho \frac{\partial \psi}{\partial \varepsilon} \quad A = -\rho \frac{\partial \psi}{\partial \alpha} \quad (15)$$

In this way, the isothermal mechanical behaviour of the time-independent standard dissipative material can be characterised by three thermodynamic potential functions  $\psi(\varepsilon, \alpha)$ ,  $f(A, \alpha)$  and  $F(A, \dot{\alpha})$ . By the help of these functions the most general thermodynamic form of the tangent modulus can be obtained. Taking the time derivative  $\dot{\sigma} = \partial \sigma / \partial t$  as quasi-static velocity into account, the rate constitutive relation can be written in the form

$$\dot{\sigma} = \mathbf{L}(\varepsilon, \alpha) : \dot{\varepsilon} \quad (16)$$

where the operator  $\mathbf{L}$  is the tangent modulus as follows

$$\mathbf{L} = \begin{cases} \mathbf{E} & \text{if } f(A, \alpha) = 0 \text{ and } \mathbf{b} : \mathbf{E} : \dot{\varepsilon} < 0 \\ \mathbf{E} - \frac{(\mathbf{E} : \mathbf{a}) \otimes (\mathbf{b} : \mathbf{E})}{h} & \text{if } f(A, \alpha) = 0 \text{ and } \mathbf{b} : \mathbf{E} : \dot{\varepsilon} \geq 0 \end{cases} \quad (17)$$

in which

$$\mathbf{E} = \rho \frac{\partial^2 \psi}{\partial \varepsilon \partial \varepsilon} \quad (18)$$

is the tensor of elastic moduli and the tensors

$$\mathbf{a} = \mathbf{E}^{-1} : \Lambda^T \circ \frac{\partial F}{\partial A} \quad \mathbf{b} = \frac{\partial f}{\partial A} \circ \Lambda : \mathbf{E}^{-1} \quad (19)$$

are related to the domain of reversibility.

The most general form (15) of the tangent modulus contains equally any nonlinear inelastic and even strain softening or damaging behaviour of materials. Let us consider now the special cases.

In the case of *elastic-plastic materials*, function  $\psi$  is the Helmholtz free energy. According to [3], for elastic-plastic materials, tensors  $\Lambda = \Pi = \mathbf{E}$ , and functions  $f = F$ , so  $\mathbf{a} = \mathbf{b}$ , consequently, the tangent modulus (17) is simplified to

$$\mathbf{L} = \begin{cases} \mathbf{E} & \text{if } f(A, \alpha) = 0 \text{ and } \mathbf{a} : \mathbf{E} : \dot{\varepsilon} < 0 \\ \mathbf{E} - \frac{(\mathbf{E} : \mathbf{a}) \otimes (\mathbf{a} : \mathbf{E})}{h} & \text{if } f(A, \alpha) = 0 \text{ and } \mathbf{a} : \mathbf{E} : \dot{\varepsilon} \geq 0 \end{cases} \quad (20)$$

Moreover, in the case of *linearly elastic and perfectly plastic materials*,  $\mathbf{a} = \mathbf{b} = \mathbf{1}$  and  $\partial f / \partial \alpha = 0$ , consequently, the tangent modulus reads

$$\mathbf{L} = \begin{cases} \mathbf{E} & \text{for elastic loading, elastic-plastic unloading} \\ \mathbf{0} & \text{for plastic loading} \end{cases} \quad (21)$$

In the case of *damaging materials*, a scalar damage coefficient  $D$  is introduced among the internal variables. According to [4] for *elastic-damaging materials*, let  $D$  be  $0 \leq D \leq D_{cr} \leq +\infty$ , fulfilling the basic condition that for  $D = 0$ , the material is perfectly elastic, and for  $D = D_{cr}$ , the material is perfectly damaged. The stiffness of the material is characterised by the function  $g(D)E$  where  $E$  is the initial elastic modulus. Thus the rate constitutive law can be written in the form

$$\dot{\sigma} = \begin{cases} g(D) E \dot{\varepsilon} & \text{if } Q < a(D) \\ g(D) E \dot{\varepsilon} + \frac{g'(D)}{k'(D)} \langle \varepsilon E \dot{\varepsilon} \rangle E \varepsilon & \text{if } Q = a(D) \end{cases} \quad (22)$$

where  $Q$ ,  $0 \leq Q \leq a(D)$  is the damage internal force, and  $a(D) \geq 0$ ,  $a'(D) \geq 0$  is a given function. Moreover,  $k(D) = -a(D) / g'(D)$ .

For a very simple model of the elastic-damaging material, [3] suggest for  $D$  to be  $0 \leq D \leq 1$ , and for the free energy  $\rho\psi(\varepsilon, D) = \frac{1}{2}(1-D) \varepsilon : E : \varepsilon$ , moreover,  $k(D) = Q_0 + MD$  where  $Q_0$  and  $M$  are material constants. Thus they suggest the tangent modulus as

$$\mathbf{L} = (1-D) \mathbf{E} - \frac{(\mathbf{E} : \varepsilon) \otimes (\varepsilon : \mathbf{E})}{M} \quad (23)$$

for the elastic-damaging material.

Other tangent moduli are obtained for continuum damage material on the basis of fracture mechanics by Janson and Hult [14] and Del Piero and Sampaio [8].

#### 2.4. Application of the tangent modulus to discrete systems

The discrete version of the tangent modulus is analysed in the monograph of Bazant and Cedolin [2]. By applying the concept of the *tangentially equivalent elastic structure* and an incremental quasi-elastic analysis, they can use the basic principles of thermodynamics.

According to the kinematic degree of freedom of the discrete structure, by introducing the generalized coordinates  $q = \{q_i\}$ ,  $i = 1, 2, \dots, n$ , the tangentially equivalent elastic version of the original inelastic structure can be defined by [2, p. 635]:

$$d\mathbf{f}_T = \hat{\mathbf{K}}_T(v) dq \quad \text{and} \quad d\mathbf{f}_S = \hat{\mathbf{K}}_S(v) dq \quad (24)$$

in which  $\mathbf{f}_T$  and  $\mathbf{f}_S$  are the internal forces being work-compatible with the kinematic parameters  $\mathbf{q}$ . Moreover  $\hat{\mathbf{K}}_T$  and  $\hat{\mathbf{K}}_S$  are the matrices of the tangent moduli, the so-called *structural tangent modulus* for isothermal and isentropic conditions, respectively, related to the direction  $v$  of the  $d\mathbf{q}$ .



Thus the concept of the potential energy can be extended to the inelastic problems for isothermal and isentropic cases as follows

$$\pi_T = F - W \quad \text{and} \quad \pi_S = U - W \quad (25)$$

where  $F$  is the Helmholtz free energy,  $U$  is the total energy and  $W$  is the energy of the external forces [2, p. 638].

According to the concept of tangentially equivalent elastic structure the internal forces  $\mathbf{f}_T$  and  $\mathbf{f}_S$  can be derived from the potential functions  $F$  and  $U$ , respectively

$$\mathbf{f}_{Ti} = \frac{\partial F}{\partial q_i} \quad \text{and} \quad \mathbf{f}_S = \frac{\partial U}{\partial q_i} \quad (26)$$

In mechanical equilibrium, the first variation of the potential energies  $\pi_T$  and  $\pi_S$  must vanish

$$\delta\pi_T = \delta F = \mathbf{f}_T^T \delta \mathbf{q} - \mathbf{P}^T d\mathbf{q} \quad \text{and} \quad \delta\pi_S = \delta U = \mathbf{f}_S^T \delta \mathbf{q} - \mathbf{P}^T d\mathbf{q} \quad (27)$$

in which  $\mathbf{P}$  is the external load in work-compatibility with  $\mathbf{q}$ . For stability qualification of equilibrium, the second order increment of the potential energies  $\pi_T$  and  $\pi_S$  are needed. The equilibrium state at  $\mathbf{q}$  is stable if

$$\delta^2 \pi_T = \frac{1}{2} \delta q_i \frac{\partial^2 \pi_T}{\partial q_i \partial q_j} \delta q_j = \frac{1}{2} (\delta \mathbf{f}_T^T - \delta \mathbf{P}^T) d\mathbf{q} = \frac{1}{2} d\mathbf{q}^T \mathbf{K}_T(v) d\mathbf{q} > 0 \quad (28)$$

$$\delta^2 \pi_S = \frac{1}{2} \delta q_i \frac{\partial^2 \pi_S}{\partial q_i \partial q_j} \delta q_j = \frac{1}{2} (\delta \mathbf{f}_S^T - \delta \mathbf{P}^T) d\mathbf{q} = \frac{1}{2} d\mathbf{q}^T \mathbf{K}_S(v) d\mathbf{q} > 0$$

where the matrices

$$\mathbf{K}_T(v) = \frac{\partial^2 \pi_T}{\partial q_i \partial q_j} \quad \text{and} \quad \mathbf{K}_S(v) = \frac{\partial^2 \pi_S}{\partial q_i \partial q_j} \quad (29)$$

represent the structural tangential stiffness. These matrices are the Hessians of the potential functions  $\pi_T$  and  $\pi_S$  for isothermal and isentropic conditions, respectively. If the external load  $\mathbf{P}$  is independent of  $\mathbf{q}$ , then matrices  $\mathbf{K}_T$  and  $\mathbf{K}_S$  are equal to  $\hat{\mathbf{K}}_T$  and  $\hat{\mathbf{K}}_S$ , respectively.

### 3. The nonsmooth tangent modulus

After the short review of the tangent modulus, we focus ourselves to the nonsmooth characteristics of it. Uniaxial material behaviour will be analysed. Before detailing the consequences of the polygonal material behaviour, the short history of the nonsmooth potential theory is considered.

### 3.1. Development of the nonsmooth potential theory

For conservative mechanical systems the stability conclusions can be drawn simply from the properties of the total potential energy functional, by the Lagrange–Dirichlet theorem [42, 43]. The fundamental stability statements are based on the classical potential law

$$\sigma(\varepsilon) = \frac{\partial W(\varepsilon)}{\partial \varepsilon} \quad (30)$$

where  $\sigma = \{\sigma_{ij}\}$  and  $\varepsilon = \{\varepsilon_{ij}\}$  are the stress and strain tensors, respectively. Functional  $W(\varepsilon)$  is the smooth and convex strain energy density.

If functional  $W(\varepsilon)$  is nonsmooth but the material is reversible, the classical potential law (30) can be extended to polygonal elastic cases. Panagiotopoulos pointed out in [33, p. 85] that while the “Smooth Mechanics” is based on the notion of the classical potential, the “Nonsmooth Mechanics” is concerned with the nonsmooth and/or nonfinite convex or nonconvex superpotentials.

The generalization of the classical potential law to nonsmooth but convex potentials named superpotential was introduced by Moreau [25, 26] by using the tools of the convex analysis. The convexity of an energy function implies the monotonicity of the concerning stress-strain relation. Variational principles related to such kind of problems have the form of variational inequalities. In order to overcome the constraint of monotonicity, the notion of nonconvex superpotential was introduced by Panagiotopoulos [30] by using the generalized gradient of Clarke [7] and the results of Rockafellar [38] leading to the hemivariational inequalities in mechanical applications. In his important book Panagiotopoulos laid down the foundations of the “Nonsmooth Mechanics” and established the substationarity laws of mechanics [32]. So he obtained the generalized substationarity principles for nonconvex potentials [32, pp. 61–160; 33, pp. 107–153]. Further applications are given in [27].

The term of nonsmoothness in sense of the definitions of Panagiotopoulos [32, pp. 3–67, 33, pp. 85–95] is based on the Lipschitzian property of functions. Simply saying, for a break type discontinuity of a function  $f(x)$  at  $x$ , the Lipschitzian condition at  $x$  fulfils, while for a jump type discontinuity it does not. The existence of both the subdifferential  $\partial f(x)$  and the generalized gradient  $\partial \tilde{f}(x)$  requires the Lipschitzian property of the function at  $x$ . A point  $x_0$  is called a substationarity point of  $f(x)$  if it is a solution of the multivalued equation

$$0 \in \partial \tilde{f}(x) \quad (31)$$

where the generalized gradient  $\partial \tilde{f}(x)$  of Clarke is a set being never empty if  $f(x)$  is Lipschitzian at  $x$ . If  $f(x)$  is convex then  $\partial \tilde{f}(x)$  coincides with the subdifferential:



$\partial\bar{f}(x) = \{\text{grad}f(x)\}$ , moreover if it is also continuously differentiable at  $x$ , then  $\partial\bar{f}(x) = \text{grad}f(x)$ .

If the material is reversible but the strain energy functional  $W(\varepsilon)$  is nonsmooth, the conservative stresses can be obtained from the inclusion

$$\sigma(\varepsilon) \in \partial W(\varepsilon) \quad (32)$$

named superpotential law, introduced by Moreau in [26]. In (32)  $\partial W(\varepsilon)$  is the subdifferential of the superpotential  $W(\varepsilon)$ , a multivalued mapping as the generalization of the classical potential law (30).

Further generalization given by Panagiotopoulos in [30] aimed to extend the potential law to nonmonotone material behaviour. If the material has a nonmonotone constitutive law but the material is reversible, then the following nonconvex superpotential law is valid:

$$\sigma(\varepsilon) \in \bar{\partial} W(\varepsilon) \quad (33)$$

by using the generalized gradient of Clarke. In this way, a wide range of decreasing and even a saw-tooth form material behaviour can be dealt with.

According to Panagiotopoulos [33], the classical linear elasticity laws can be replaced by nonlinear elasticity laws  $\sigma(\varepsilon) = \partial W(\varepsilon) / \partial \varepsilon$  or more generally, by the monotone nonsmooth law  $\sigma(\varepsilon) \in \partial W(\varepsilon)$ , and even by the nonmonotone nonsmooth law  $\sigma(\varepsilon) \in \bar{\partial} W(\varepsilon)$ , if the material is reversible.

In his basic works, Panagiotopoulos deals also with the potential law of dissipative mechanical systems in [31–33]. From the view of the nonsmoothness and nonconvexity on thermodynamic bases, he obtain mainly the same stationarity conclusion as mentioned above. By introducing a general nonsmooth thermodynamic potential of dissipation and applying the superpotential law of Clarke, he gets to the generalization of the hypothesis of normal dissipation. He states that in the case of dissipation or unloading, an incremental analysis has to be applied. He deals generally with strongly nonmonotone and nonsmooth cases like saw-tooth behaviour and damage, but in aspect first of all the equilibrium and not the stability.

On the basis of the superpotential law, we can extend the concept of tangent modulus to the nonsmooth version of it, containing unloading and polygonal behaviour as well.

### 3.2. The nonsmooth tangent modulus in the case of break points in the material functions

Figure 2a shows a uniaxial stress-strain diagram  $\sigma(\varepsilon)$  representing a nonlinear inelastic material behaviour. By applying the concept of tangentially equivalent elastic structure of Bazant and Cedolin [2, p. 635], the responses of an inelastic structure can be solved in small loading steps by a series of quasielastic incremental analysis, taking the inelastic constitutive law as a thermodynamic equation of state into account. Thus, for a small step  $d\varepsilon$ , the increment of the strain energy can be considered elastic



$$dW(\varepsilon) = \sigma(\varepsilon) d\varepsilon \quad (34)$$

thus, the strain energy  $W(\varepsilon)$  at  $\varepsilon$  can as the potential function of the stresses  $\sigma(\varepsilon)$  be considered, that is, the classical potential law

$$\sigma(\varepsilon) = \frac{\partial W(\varepsilon)}{\partial \varepsilon} \quad (35)$$

can be applied. Since for small loading step, the material behaviour is supposed to be linear elastic, the increment of the stresses consists of first order term only, namely, in this case

$$d\sigma(\varepsilon, d\varepsilon) = \delta\sigma(\varepsilon, \delta\varepsilon) = \frac{d\sigma(\varepsilon)}{d\varepsilon} \delta\varepsilon = \frac{d^2W(\varepsilon)}{d\varepsilon^2} \delta\varepsilon = k_t(\varepsilon) \delta\varepsilon \quad (36)$$

in which

$$k_t(\varepsilon) = \frac{d\sigma(\varepsilon)}{d\varepsilon} = \frac{d^2W(\varepsilon)}{d\varepsilon^2} \quad (37)$$

is the tangent modulus related to the *linear comparison solid* introduced by Hill [10]. The tangent modulus  $k_t(\varepsilon)$  in Fig. 2a is the actual tangent of the stress-strain function  $\sigma(\varepsilon)$  at  $\varepsilon$ .

The tangent modulus (37) of the linear comparison solid is related to a smooth stress-strain functions  $\sigma(\varepsilon)$ , as the actual tangent of the material function. The introduction of the linear comparison solid aimed to avoid the nonsmoothness due to the switch from loading to unloading. Indeed, the classical tangent modulus aims to give a relation between the stress and strain increments, similar to the Hooke's law of the elastic materials

$$d\sigma = K_t(\varepsilon) d\varepsilon \quad \text{where} \quad K_t(\varepsilon) = \begin{cases} k_o & \text{for } d\varepsilon < 0 \text{ (unloading)} \\ k_t & \text{for } d\varepsilon > 0 \text{ (loading)} \end{cases} \quad (38)$$

where  $k_o$  is the initial elastic modulus. The point  $d\varepsilon = 0$ , at which the material changes from loading to unloading, is excluded from this expression, since at  $d\varepsilon = 0$  the function  $K_t(\varepsilon)$  has a jump. Even for avoiding this discontinuity, Hill has introduced the linear comparison solid, consequently, the unloading has been left out from the analyses.

If we want to extend the tangent modulus to the unloading too, the following stress function including the case  $d\varepsilon = 0$  is to be applied

$$\sigma(\varepsilon) = \begin{cases} \sigma_o(\varepsilon) & \text{if } d\varepsilon \leq 0 \\ \sigma_t(\varepsilon) & \text{if } d\varepsilon \geq 0 \end{cases} \quad (39)$$

in which the function  $\sigma_o(\varepsilon)$  concerns both the elastic-plastic or elastic-plastic-damage unloading, while the function  $\sigma_t(\varepsilon)$  belongs to loading only, seen in Fig. 2a. Note that function  $\sigma_t(\varepsilon)$  represents the linear comparison solid of the original nonlinear material.

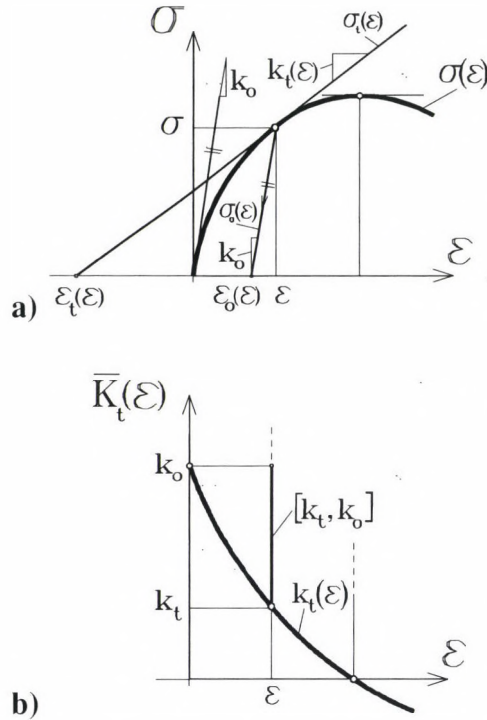


Fig. 2. The nonsmooth multivalued tangent modulus

Considering Fig. 2a, function (39) can be written in the form of

$$\sigma(\epsilon) = \begin{cases} k_o(\epsilon)[\epsilon - \epsilon_o(\epsilon)] & \text{if } d\epsilon \leq 0 \\ k_t(\epsilon)[\epsilon - \epsilon_t(\epsilon)] & \text{if } d\epsilon \geq 0 \end{cases} \quad (40)$$

in which  $k_t(\epsilon)$  and  $k_o(\epsilon)$  are the loading and unloading moduli related to the straight lines  $\sigma_t(\epsilon)$  and  $\sigma_o(\epsilon)$  at  $\epsilon$ , respectively. Strain values  $\sigma_t(\epsilon)$  and  $\sigma_o(\epsilon)$  are the intersections of the straight lines  $\sigma_t(\epsilon)$  and  $\sigma_o(\epsilon)$  with the axis  $\epsilon$ , respectively. As we can see in Fig. 2a, all these values are point by point changing, but at the strain value  $\epsilon$ , they are constant thus

$$\sigma(\epsilon) = \begin{cases} k_o(\epsilon - \epsilon_o) & \text{if } d\epsilon \leq 0 \\ k_t(\epsilon - \epsilon_t) & \text{if } d\epsilon \geq 0 \end{cases} \quad (41)$$

However, this function containing both loading and unloading is a nonsmooth function. Consequently, for obtaining the tangent modulus from it, instead of the relation in (37), subdifferentiation has to be applied

$$\bar{K}_t(\varepsilon) \equiv \bar{\sigma}(\varepsilon) = \begin{cases} \frac{d\sigma_o(\varepsilon)}{d\varepsilon} & \text{if } d\varepsilon < 0 \\ \left[ \frac{d\sigma_t(\varepsilon)}{d\varepsilon}, \frac{d\sigma_o(\varepsilon)}{d\varepsilon} \right] & \text{if } d\varepsilon = 0 \\ \frac{d\sigma_t(\varepsilon)}{d\varepsilon} & \text{if } d\varepsilon > 0 \end{cases} \quad (42)$$

resulting a multivalued function forming an interval at the condition  $d\varepsilon = 0$

$$\bar{K}_t(\varepsilon) = \begin{cases} k_o & \text{if } d\varepsilon < 0 \\ [k_t, k_o] & \text{if } d\varepsilon = 0 \\ k_t & \text{if } d\varepsilon > 0 \end{cases} \quad (43)$$

yielding the nonsmooth tangent modulus seen in Fig. 2b with an interval of  $[k_t, k_o]$  at  $d\varepsilon = 0$ . Here  $k_o$  is the initial elastic modulus and  $k_t = k_t(\varepsilon)$  is the actual tangent of function  $\sigma(\varepsilon)$ .

Thus the actual occurring stiffness  $K_t(\varepsilon)$  is the element of the set  $\bar{K}_t(\varepsilon)$  as follows

$$K_t(\varepsilon) \in \bar{K}_t(\varepsilon) \quad (44)$$

However, the set of the nonsmooth tangent modulus  $\bar{K}_t(\varepsilon)$  is changing with changing  $\varepsilon$ . If the material function  $\sigma(\varepsilon)$  is nonlinear then the solution is mathematically very difficult. That is why the idea of polygonal approximation seems to be highly reasonable. However, it leads to nonsmoothness again.

Figure 3a shows a polygonal approximation of a nonlinear material function. Naturally, the polygonal function can represent an originally polygonal material behaviour, too. For example, the composite materials or the locking behaviour or the saw-tooth type behaviour can be characterised by a polygonal constitutive law. In this case, the solution gives the correct results not an approximation, of course.

First we consider break type functions without jumps seen in Fig. 3. Later we return to the jumps, too. As we will see, to handle the polygonal behaviour seems to be the same as that of the loading-unloading, since the latter case has a simple polygonal characteristic, too.

Consider first the loading phase only. Let each segment  $i$  of the polygonal material law in Fig. 3a be specified by the relating modulus  $k_t^i$  as the constant slope of the segment  $i$  and by the strain constant  $\varepsilon_t^i$  as the intersection of the segment  $i$  and the axis  $\varepsilon$ . Thus, at the break point  $\varepsilon = \varepsilon_t$ , the material function  $\sigma(\varepsilon)$  can be written in the form of



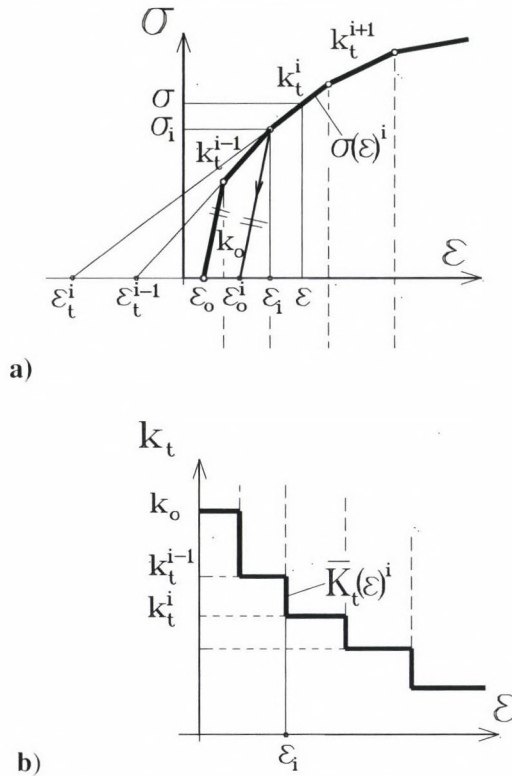


Fig. 3. The nonsmooth tangent modulus of polygonal material

$$\sigma(\epsilon)^i = \begin{cases} k_t^{i-1}(\epsilon - \epsilon_t^{i-1}) & \text{if } d\epsilon \leq 0 \\ k_t^i(\epsilon - \epsilon_i) & \text{if } d\epsilon \geq 0 \end{cases} \quad (45)$$

which is equal to the function (40) if the segment preceding the break point  $\epsilon = \epsilon_i$  can as unloading path be considered. Consequently, any path can as unloading path be handled if the sign of the strain increment  $d\epsilon$  indicates the loading or unloading characteristics of the material phases.

The nonsmooth tangent modulus  $K_t(\epsilon)^i$  belonging to the point  $\epsilon = \epsilon_i$  and seen in Fig. 3b can be obtained by subdifferentiating the function  $\sigma(\epsilon)^i$  at  $\epsilon = \epsilon_i$

$$\bar{K}_t(\epsilon)^i \equiv \bar{\partial}(\sigma(\epsilon)^i) = \begin{cases} k_t^{i-1} & \text{if } d\epsilon < 0 \\ [k_t^i, k_t^{i-1}] & \text{if } d\epsilon = 0 \\ k_t^i & \text{if } d\epsilon > 0 \end{cases} \quad (46)$$

relating to both loading and unloading when taking the preceding segment  $i-1$  as unloading path into consideration. The loading and unloading paths are indicated clearly by the sign of  $d\varepsilon > 0$  and  $d\varepsilon < 0$ , respectively.

Alternatively according to (37) the nonsmooth tangent modulus can be obtained by the second subdifferential at  $\varepsilon = \varepsilon_i$  of the nonsmooth superpotential  $W(\varepsilon)$ , too

$$K_t(\varepsilon)^i \equiv \bar{\partial}(\partial W(\varepsilon)^i) \equiv \bar{\partial}(\sigma(\varepsilon)^i) \tag{47}$$

Thus, the concept of the nonsmooth tangent modulus of polygonal material behaviour can be extended to the strain softening, namely, to damage problems. In contrast to the elastic-plastic unloading, in damaging cases, the unloading and reloading moduli are to be given individually. Consider a polygonal function of an elastic-plastic-damaging material seen in Fig. 4a. Also in the case of damaging materials, the unloading paths are linear, but in contrast to the plastic unloading, with different elastic moduli. Thus, the unloading moduli  $k_t^u$  are changing depending on the actual strains. As a typical damage property, in the case of active damage loading, the loading moduli  $k_t^i$  are negative.

Moreover the concept of the nonsmooth tangent modulus can be related to the so-called *locking materials*, too [41]. The initial stiffness of this materials during a loading process increases and finally, the material can become even perfectly rigid seen in Fig. 4b. In Fig. 4c the perfectly locking behaviour is illustrated. In spite of the fact that this type of materials are reversible, they can be handled similarly to the irreversible problems. The locking behaviour belongs to the family of the so-called *conditional joints* described first by Kaliszky in [15]. Further generalization of the conditional joints as subdifferential material property, and as the dual version of the reversible plastic characteristics were given by Kurutz [17, 18]. The stability conclusions due to nonsmooth behaviour is analysed also by Kurutz in [19–24].

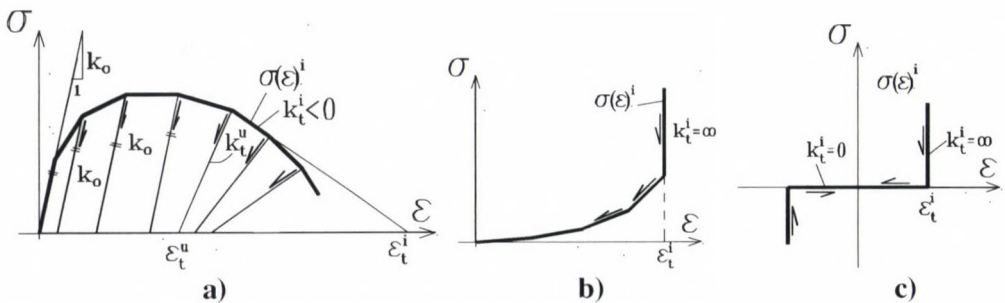


Fig. 4. Polygonal damaging and locking materials

3.3 The nonsmooth tangent modulus in the case of jumps in the material functions

There are materials applied recently, the stress-strain or force-displacement functions of which show jump type characteristics. This tendency can occur in both strain softening or strain hardening phases. The composite materials belong to the former, and the locking materials belong to the latter.

Consider the example of the perfect jump in Fig. 5a, the diagram of the perfectly rigid perfectly plastic material. In this case, the unloading and reloading take place in a perfectly rigid manner by the condition  $d\varepsilon = 0$ , manifested in a jump.

The material behaviour is characterised by the inclusion

$$\sigma(\varepsilon) \in \bar{\sigma}(\varepsilon) = \begin{cases} \sigma_2 & \text{if } d\varepsilon < 0 \\ [\sigma_2, \sigma_1] & \text{if } d\varepsilon = 0 \\ \sigma_1 & \text{if } d\varepsilon > 0 \end{cases} \quad (48)$$

according which, independently of  $\varepsilon$ , the actual stresses are the elements of the set of stresses related equally to loading, unloading and reloading.

According to (37), for the nonsmooth tangent modulus, this function needs to be sub-differentiated. However, since this function has jumps at any  $d\varepsilon = 0$ , the Lipschitz condition does not fulfil, so nor the subdifferential of Moreau, nor the generalized gradient of Clarke exists. Still, if we want to obtain the tangent modulus in such kind of Heaviside type material functions, a distributional derivative has to be applied. Thus the generalized nonsmooth tangent modulus is as follows

$$\bar{K}_t(\varepsilon) \equiv \bar{\partial}(\bar{\sigma}(\varepsilon)) = \begin{cases} 0 & \text{if } d\varepsilon \neq 0 \\ \pm (\sigma_1 - \sigma_2)\delta(d\varepsilon) & \text{if } d\varepsilon = 0 \end{cases} \quad (49)$$

where  $\delta(d\varepsilon)$  is the Dirac impulse [16]. For the condition  $d\varepsilon = 0$ , the tangent modulus forms an interval of indefinite length, namely, in the case of loading (unloading) it tends to the positive (negative) infinite.

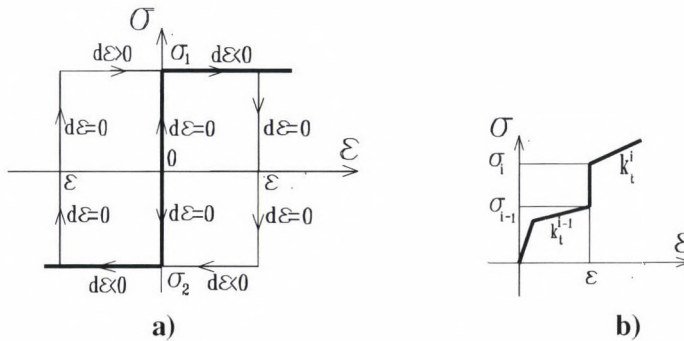


Fig. 5. Jump like materials



Naturally, an arbitrary jump  $[\sigma_{i-1}, \sigma_i]$  in the material function can equally happen for a reversible or an irreversible material at any strain value. Consider now an elastic material with a jump  $[\sigma_{i-1}, \sigma_i]$ , both preceded and followed by elastic behaviour seen in Fig. 5b. The nonsmooth tangent modulus at  $\varepsilon$  then reads

$$\bar{K}(\varepsilon) \equiv \bar{\partial}(\bar{\sigma}(\varepsilon)) = \begin{cases} k_t^{i-1} & \text{if } d\varepsilon < 0 \\ \pm (\sigma_i - \sigma_{i-1}) \delta(d\varepsilon) & \text{if } d\varepsilon = 0 \\ k_t^i & \text{if } d\varepsilon > 0 \end{cases} \quad (50)$$

since the unloading paths are equal to the loading ones.

Construct now the nonsmooth tangent modulus related to a discrete structural model.

#### 4. The nonsmooth structural tangent modulus

Let the bent structure be composed by perfectly rigid elements connected to each other (and to the support) by special springs in which the material behaviour is concentrated. The behaviour of each of the springs is characterised by a one-dimensional stress-strain function being arbitrarily nonlinear or nonconvex. We assume this functions to be approximated with polygonal form.

Let  $\sigma = \{\sigma_i\}$  and  $\varepsilon = \{\varepsilon_i\}$ ,  $i = 1, 2, \dots, a$  be the vectors of stresses and strains, respectively, where the length  $a$  of the vectors represents the number of all the stress and strain components over the structure. In our discrete model, the variables  $\sigma_i$  and  $\varepsilon_i$  are forces and deformations of the springs being work-compatible with each other. Vector  $\sigma = \sigma(\varepsilon) = \sigma_i \{\varepsilon_i\}$  contains the one-dimensional but arbitrary functions of the material behaviour.

Assume that the cinematic state of the structure can be characterised by  $n$  number of independent kinematic parameters. Then the kinematic parameters  $\mathbf{q} = \{q_i\}$ ,  $i = 1, 2, \dots, n$  are the so-called generalized coordinates of the given numerical solution.

It is well known that the potential energy functional  $\pi(\varepsilon, \mathbf{u})$  requires the compatibility conditions as subsidiary conditions. By introducing the generalized kinematic coordinates  $\mathbf{q}$ , the compatibility conditions form

$$\mathbf{u} = \mathbf{u}(\mathbf{q}), \quad \text{thus} \quad \varepsilon = \varepsilon(\mathbf{u}) = \varepsilon(\mathbf{u}(\mathbf{q})) = \varepsilon(\mathbf{q}) \quad (51)$$

consequently, the functional  $\pi(\varepsilon, \mathbf{u})$  can be reduced to a scalar function  $\pi(\mathbf{q})$  which is kinematically admissible at the same time. However, the variables  $\varepsilon$  and  $\mathbf{u}$  of the potential energy function are no more scalar, consequently, the function  $\pi(\mathbf{q}) = \pi(\varepsilon(\mathbf{u}(\mathbf{q})), \mathbf{u}(\mathbf{q}))$  is a compound function.

We assume the compatibility transformations  $\mathbf{u} = \mathbf{u}(\mathbf{q})$ ,  $\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}(\mathbf{u})$  or  $\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}(\mathbf{q})$  to be smooth functions, while the constitutive laws  $\boldsymbol{\sigma} = \boldsymbol{\sigma}(\boldsymbol{\varepsilon})$  are allowed to be nonsmooth, consequently, the functions  $\boldsymbol{\sigma} = \boldsymbol{\sigma}(\boldsymbol{\varepsilon}(\mathbf{u}(\mathbf{q}))) = \boldsymbol{\sigma}(\mathbf{q})$  can be nonsmooth. Thus, the internal potential can be nonsmooth while the external potential is always smooth.

In any case, by introducing any type of generalized coordinates, it is necessary to transform the concerning state variables to be work-compatible with the chosen generalized coordinates. So we obtain the reduced version of the stresses and external forces being work-compatible with the kinematic parameters  $\mathbf{q}$ . In order to obtain the structural tangent modulus, we need the reduction of the stresses only. Thus, here we do not deal with the external load and external potential.

Consider now the first- and second-order increments of the internal potential.

The first variation of the compound *smooth* function  $\pi_{in}(\mathbf{q}) = \pi_{in}(\boldsymbol{\varepsilon}(\mathbf{u}(\mathbf{q})))$  reads

$$\delta\pi_{in} = \frac{\partial\pi_{in}(\mathbf{q})}{\partial q_i} \delta q_i = \frac{\partial\pi_{in}(\boldsymbol{\varepsilon})}{\partial \varepsilon_l} \frac{\partial \varepsilon_l(\mathbf{u})}{\partial u_j} \frac{\partial u_j(\mathbf{q})}{\partial q_i} \delta q_i = \boldsymbol{\sigma}(\boldsymbol{\varepsilon})^T \delta \boldsymbol{\varepsilon} = \mathbf{f}(\mathbf{q})^T \delta \mathbf{q} \quad (52)$$

while for the nonsmooth internal potential, we obtain

$$\bar{\delta}\pi_{in} \equiv \bar{\partial}_i \pi_{in}(\mathbf{q}) \delta q_i \equiv \bar{\partial}_i \pi_{in}(\boldsymbol{\varepsilon}) \frac{\partial \varepsilon_l(\mathbf{u})}{\partial u_j} \frac{\partial u_j(\mathbf{q})}{\partial q_i} \delta q_i \equiv \bar{\boldsymbol{\sigma}}(\boldsymbol{\varepsilon})^T \delta \boldsymbol{\varepsilon} \equiv \bar{\mathbf{f}}(\mathbf{q})^T \delta \mathbf{q} \quad (53)$$

where the functions

$$\mathbf{f}(\mathbf{q})^T = \{f_i(\mathbf{q})\}^T = \frac{\partial\pi_{in}(\boldsymbol{\varepsilon})}{d\varepsilon_l} \frac{\partial \varepsilon_l(\mathbf{u})}{\partial u_j} \frac{\partial u_j(\mathbf{q})}{\partial q_i} = \boldsymbol{\sigma}_l(\boldsymbol{\varepsilon}) \frac{\partial \varepsilon_l(\mathbf{u})}{\partial u_j} \frac{\partial u_j(\mathbf{q})}{\partial q_i} \quad (54)$$

and

$$\bar{\mathbf{f}}(\mathbf{q})^T = \{\bar{f}_i(\mathbf{q})\}^T = \bar{\partial}_i \pi_{in}(\boldsymbol{\varepsilon}) \frac{\partial \varepsilon_l(\mathbf{u})}{\partial u_j} \frac{\partial u_j(\mathbf{q})}{\partial q_i} = \bar{\boldsymbol{\sigma}}_l(\boldsymbol{\varepsilon}) \frac{\partial \varepsilon_l(\mathbf{u})}{\partial u_j} \frac{\partial u_j(\mathbf{q})}{\partial q_i} \quad (55)$$

are the smooth and nonsmooth versions of the reduced stresses to be work-compatible with the chosen kinematic parameters  $\mathbf{q}$ . Vector  $\bar{\mathbf{f}}(\mathbf{q})$  consists of interval elements due to the set of stresses  $\bar{\boldsymbol{\sigma}}(\boldsymbol{\varepsilon})$  obtained by the subdifferentiating of the nonsmooth strain energy  $\pi_{in}(\boldsymbol{\varepsilon})$ .

Functions  $\mathbf{f}(\mathbf{q})$  and  $\bar{\mathbf{f}}(\mathbf{q})$  are the smooth and nonsmooth version of the smooth and nonsmooth stresses  $\boldsymbol{\sigma}(\boldsymbol{\varepsilon})$  and  $\bar{\boldsymbol{\sigma}}(\boldsymbol{\varepsilon})$ , respectively, being work-compatible with the chosen displacement parameters  $\mathbf{q}$ . Functions  $\mathbf{f}(\mathbf{q})$  and  $\bar{\mathbf{f}}(\mathbf{q})$  are the constitutive laws of the structure, the so-called *structural material functions*, since these functions unite the material behaviour of all the material points of the structure. The nonsmooth function  $\bar{\mathbf{f}}(\mathbf{q})$  inherits the polygonal characteristics of the material laws of each point of the structure, due to any break or jump in the material polygonal.

The structural functions  $\mathbf{f}(\mathbf{q})$  and  $\bar{\mathbf{f}}(\mathbf{q})$  can be expressed in matrix form, so, for the smooth and nonsmooth reduced stresses we have

$$\mathbf{f}(\mathbf{q})^T = \boldsymbol{\sigma}(\boldsymbol{\varepsilon})^T \mathbf{A}(\mathbf{u}) \mathbf{B}(\mathbf{q}) = \boldsymbol{\sigma}(\boldsymbol{\varepsilon})^T \mathbf{M}(\mathbf{q}) = \boldsymbol{\sigma}(\boldsymbol{\varepsilon}(\mathbf{u}(\mathbf{q})))^T \mathbf{M}(\mathbf{q}) \quad (56)$$

and

$$\bar{\mathbf{f}}(\mathbf{q})^T \equiv \bar{\boldsymbol{\sigma}}(\boldsymbol{\varepsilon})^T \mathbf{A}(\mathbf{u}) \mathbf{B}(\mathbf{q}) \equiv \bar{\boldsymbol{\sigma}}(\boldsymbol{\varepsilon})^T \mathbf{M}(\mathbf{q}) \equiv \bar{\boldsymbol{\sigma}}(\boldsymbol{\varepsilon}(\mathbf{u}(\mathbf{q})))^T \mathbf{M}(\mathbf{q}) \quad (57)$$

in which the matrices  $\mathbf{A}(\mathbf{u})$ ,  $\mathbf{B}(\mathbf{q})$  and  $\mathbf{M}(\mathbf{q})$  are the derivatives of the smooth functions  $\boldsymbol{\varepsilon}(\mathbf{u})$  and  $\mathbf{u}(\mathbf{q})$  as follows

$$\mathbf{A}(\mathbf{u}) = \{a_{ij}(\mathbf{u})\} = \left\{ \frac{\partial \varepsilon_i(\mathbf{u})}{\partial u_j} \right\} \quad (58)$$

and

$$\mathbf{B}(\mathbf{q}) = \{b_{ij}(\mathbf{q})\} = \left\{ \frac{\partial u_i(\mathbf{q})}{\partial q_j} \right\} \quad (59)$$

moreover

$$\mathbf{M}(\mathbf{q}) = \{m_{ik}(\mathbf{q})\} = \mathbf{A}(\mathbf{u}) \mathbf{B}(\mathbf{q}) = \left\{ \frac{\partial \varepsilon_i(\mathbf{u})}{\partial u_j} \frac{\partial u_j(\mathbf{q})}{\partial q_k} \right\} \quad (60)$$

For qualifying the stability, we need the second-order increment of the internal potential too, as follows for smooth cases

$$\delta^2 \pi_{in} = \frac{1}{2} \delta q_1 \frac{\partial^2 \pi_{in}(\mathbf{q})}{\partial q_1 \partial q_j} \delta q_j = \frac{1}{2} \delta q_i \left\{ \frac{\partial u_i(\mathbf{q})}{\partial q_m} \frac{\partial \varepsilon_m(\mathbf{u})}{\partial u_n} \frac{\partial^2 \pi_{in}(\boldsymbol{\varepsilon})}{\partial \varepsilon_n \partial \varepsilon_l} \frac{\partial \varepsilon_l(\mathbf{u})}{\partial u_k} \frac{\partial u_k(\mathbf{q})}{\partial q_j} + \right. \quad (61)$$

$$\left. \frac{\partial u_i(\mathbf{q})}{\partial q_m} \left( \frac{\partial \pi_{in}(\boldsymbol{\varepsilon})}{\partial \varepsilon_l} \frac{\partial^2 \varepsilon_l(\mathbf{u})}{\partial u_m \partial u_k} \right) \frac{\partial u_k(\mathbf{q})}{\partial q_j} + \left( \frac{\partial \pi_{in}(\boldsymbol{\varepsilon})}{\partial \varepsilon_l} \frac{\partial \varepsilon_l(\mathbf{u})}{\partial u_k} \right) \frac{\partial^2 u_k(\mathbf{q})}{\partial q_i \partial q_j} \right\} \delta q_j = \frac{1}{2} \delta \mathbf{f}(\mathbf{q}, \delta \mathbf{q})^T \delta \mathbf{q}$$

which can be expressed in matrix form, as well

$$\delta^2 \pi_{in} = \frac{1}{2} \delta \mathbf{q}^T \left\{ \mathbf{M}(\mathbf{q})^T \mathbf{K}_t(\boldsymbol{\varepsilon}(\mathbf{u}(\mathbf{q}))) \mathbf{M}(\mathbf{q}) + \mathbf{B}(\mathbf{q})^T \left( \boldsymbol{\sigma}(\boldsymbol{\varepsilon}(\mathbf{u}(\mathbf{q})))^T \mathbf{X}(\mathbf{u}(\mathbf{q})) \right) \mathbf{B}(\mathbf{q}) + \right. \quad (62)$$

$$\left. \left( \boldsymbol{\sigma}(\boldsymbol{\varepsilon}(\mathbf{u}(\mathbf{q})))^T \mathbf{A}(\mathbf{u}(\mathbf{q})) \right) \mathbf{Y}(\mathbf{q}) \right\} \delta \mathbf{q} = \frac{1}{2} \delta \mathbf{f}^T(\mathbf{q}, \delta \mathbf{q}) \delta \mathbf{q}$$

while for nonsmooth cases, by applying repeated subdifferentiation or distributional derivatives, we have



$$\begin{aligned} \bar{\delta}^2 \pi_{in} &\equiv \frac{1}{2} \delta q_i \bar{\partial}(\bar{\partial} \pi_{in}(\mathbf{q})) \delta q_j \equiv \frac{1}{2} \delta q_i \left\{ \frac{\partial u_i(\mathbf{q})}{\partial q_m} \frac{\partial \varepsilon_m(\mathbf{u})}{\partial u_n} \bar{\partial}_n(\bar{\partial}_i \pi_{in}(\varepsilon)) \frac{\partial \varepsilon_l(\mathbf{u})}{\partial u_k} \frac{\partial u_k(\mathbf{q})}{\partial q_j} + \right. \\ &\left. \frac{\partial u_i(\mathbf{q})}{\partial q_m} \left( \bar{\partial}_l \pi_{in}(\varepsilon) \frac{\partial^2 \varepsilon_l(\mathbf{u})}{\partial u_m \partial u_k} \right) \frac{\partial u_k(\mathbf{q})}{\partial q_j} + \left( \bar{\partial}_l \pi_{in}(\varepsilon) \frac{\partial \varepsilon_j(\mathbf{u})}{\partial u_k} \right) \frac{\partial^2 u_k(\mathbf{q})}{\partial q_i \partial q_j} \right\} \delta q_j = \frac{1}{2} \delta \bar{\mathbf{f}}(\mathbf{q}, \delta \mathbf{q})^T \delta \mathbf{q} \end{aligned} \tag{63}$$

which in matrix form shows

$$\begin{aligned} \bar{\delta}^2 \pi_{in} &\equiv \frac{1}{2} \delta \mathbf{q}^T \left\{ \mathbf{M}(\mathbf{q})^T \bar{\mathbf{K}}_t(\varepsilon(\mathbf{u}(\mathbf{q}))) \mathbf{M}(\mathbf{q}) + \mathbf{B}(\mathbf{q})^T \left( \bar{\boldsymbol{\sigma}}(\varepsilon(\mathbf{u}(\mathbf{q})))^T \mathbf{X}(\mathbf{u}(\mathbf{q})) \right) \mathbf{B}(\mathbf{q}) + \right. \\ &\left. \left( \bar{\boldsymbol{\sigma}}(\varepsilon(\mathbf{u}(\mathbf{q})))^T \mathbf{A}(\mathbf{u}(\mathbf{q})) \right) \mathbf{Y}(\mathbf{q}) \right\} \delta \mathbf{q} \equiv \frac{1}{2} \delta \bar{\mathbf{f}}^T(\mathbf{q}, \delta \mathbf{q}) \delta \mathbf{q} \end{aligned} \tag{64}$$

Here the matrices  $\mathbf{X}$  and  $\mathbf{Y}$  are three dimensional, namely, the second derivatives of the smooth functions  $\varepsilon(\mathbf{u})$  in term of  $\mathbf{u}$  and  $\mathbf{u}(\mathbf{q})$  in term of  $\mathbf{q}$ , respectively, as follows

$$\mathbf{X}(\mathbf{u}) = \{x_{ijk}(\mathbf{u})\} = \left\{ \frac{\partial^2 \varepsilon_i(\mathbf{u})}{\partial u_j \partial u_k} \right\} \tag{65}$$

and

$$\mathbf{Y}(\mathbf{q}) = \{y_{ijk}(\mathbf{q})\} = \left\{ \frac{\partial^2 u_i(\mathbf{q})}{\partial q_j \partial q_k} \right\} \tag{66}$$

Thus, finally the matrix  $\mathbf{K}_t(\varepsilon)$  in (60) is the *smooth structural tangent modulus*

$$\mathbf{K}_t(\varepsilon) = \langle K_{ij}(\varepsilon) \rangle = \left\langle \frac{\partial^2 \pi_{in}(\varepsilon)}{\partial \varepsilon_i \partial \varepsilon_j} \right\rangle = \left\langle \frac{\partial \sigma_i(\varepsilon)}{\partial \varepsilon_j} \right\rangle \tag{67}$$

while the interval matrix  $\bar{\mathbf{K}}_t(\varepsilon)$  in (64) is the *nonsmooth structural tangent modulus*

$$\bar{\mathbf{K}}_t(\varepsilon) = \langle \bar{K}_{ij}(\varepsilon) \rangle = \langle \bar{\partial}_i(\bar{\partial}_j \pi_{in}(\varepsilon)) \rangle = \langle \bar{\partial}_i(\bar{\sigma}_j(\varepsilon)) \rangle \tag{68}$$

which are both diagonal matrices since we assumed uniaxial behaviour for each material point of the structure. The nonsmooth structural tangent modulus  $\bar{\mathbf{K}}_t(\varepsilon)$  contains the nonsmooth tangent moduli of each material points related to both loading and unloading.

Note that the structural tangent moduli  $\mathbf{K}_t(\varepsilon)$  and  $\bar{\mathbf{K}}_t(\varepsilon)$  are included in the stresses  $\boldsymbol{\sigma}(\varepsilon(\mathbf{u}(\mathbf{q})))$  and  $\bar{\boldsymbol{\sigma}}(\varepsilon(\mathbf{u}(\mathbf{q})))$ , respectively, for even strain softening or damage. However, for a global analysis related to the total domain of possible deflections, the in-

stantaneously changing tangent moduli are needed. That is why polygonal approximation is used in the material functions resulting break or jump points in the structural material function.

To extend the analysis to the instantaneously changing tangent modulus, we assumed polygonal material and consequently polygonal structural behaviour with breaks or jumps at points  $\mathbf{q}_i$  or  $\mathbf{q}_j$  of the function  $\bar{\mathbf{f}}(\mathbf{q})$  of the structural material behaviour. Thus, function (57)

$$\bar{\mathbf{f}}^T(\mathbf{q})^{ij} \equiv \bar{\sigma}(\varepsilon(\mathbf{u}(\mathbf{q})))^{ij} \mathbf{M}(\mathbf{q}) \quad (69)$$

related to the segment of  $\bar{\mathbf{f}}(\mathbf{q})$  between the points  $\mathbf{q}_i$  and  $\mathbf{q}_j$  are smooth, since the material properties are constants and so the stresses  $\bar{\sigma}(\varepsilon(\mathbf{u}(\mathbf{q})))^{ij}$  are smooth between  $\mathbf{q}_i$  and  $\mathbf{q}_j$ . But the function  $\bar{\mathbf{f}}(\mathbf{q})^{ij}$  is multivalued for a jump in the  $\bar{\sigma}(\varepsilon(\mathbf{u}(\mathbf{q})))^{ij}$  at  $\mathbf{q}_i$  or  $\mathbf{q}_j$ . So the stresses can be expressed in term of the tangent modulus, by separating them into smooth parts

$$\sigma^T(\varepsilon(\mathbf{u}(\mathbf{q})))^{ij} = (\varepsilon(\mathbf{u}(\mathbf{q})) - \varepsilon_t^{ij})^T \mathbf{K}_t^{ij} \quad \text{for } \mathbf{q} \in \text{int}\{\mathbf{q}_i, \mathbf{q}_j\} \quad (70)$$

and by using logical product "o" for the intervals, into nonsmooth parts

$$\bar{\sigma}^T(\varepsilon(\mathbf{u}(\mathbf{q})))^k \equiv (\varepsilon(\mathbf{u}(\mathbf{q})) - \bar{\varepsilon}^k)^T \circ \bar{\mathbf{K}}_t^k \quad \text{for } \mathbf{q} \in \text{front}\{\mathbf{q}_i, \mathbf{q}_j\} \quad (71)$$

where  $\mathbf{K}_t^{ij}$  is the smooth while  $\bar{\mathbf{K}}_t^k$  is the nonsmooth structural tangent modulus. Vectors  $\varepsilon_t^{ij}$  and  $\bar{\varepsilon}^k$  are the smooth and nonsmooth strain sections detailed in sections 3.2. and 3.3. The notations  $\text{int}\{\cdot\}$  and  $\text{front}\{\cdot\}$  are the internal and the frontier of the set  $\{\cdot\}$ , respectively.

In this way the reduced stresses  $\bar{\mathbf{f}}(\mathbf{q})$  can be expressed in term of the structural tangent modulus

$$\mathbf{f}^T(\mathbf{q})^{ij} = (\varepsilon(\mathbf{u}(\mathbf{q})) - \varepsilon_t^{ij})^T \mathbf{K}_t^{ij} \mathbf{M}(\mathbf{q}) \quad \text{for } \mathbf{q} \in \text{int}\{\mathbf{q}_i, \mathbf{q}_j\} \quad (72)$$

or

$$\bar{\mathbf{f}}^T(\mathbf{q})^k \equiv (\varepsilon(\mathbf{u}(\mathbf{q})) - \bar{\varepsilon}^k)^T \circ \bar{\mathbf{K}}_t^k \mathbf{M}(\mathbf{q}) \quad \text{for } \mathbf{q} \in \text{front}\{\mathbf{q}_i, \mathbf{q}_j\} \quad (73)$$

The reduced stresses represent the structural material behaviour which determine the characteristics of the equilibrium paths of the structure.

Consider now the numerical applications.

### 5. Application of the nonsmooth tangent modulus

Examples by using the nonsmooth tangent modulus are detailed in the papers of Kurutz [19–24] dealing with stability analysis of structures of nonsmooth energy functionals. Equilibrium paths and their stability qualifications are presented in the mentioned papers. Global stability analysis related to the total domain of the possible deflections is investigated, the nonsmooth functions of the structural tangent stiffness is analysed. Reversible and irreversible problems are equally detailed, loading-unloading and polygonal approximation of highly nonlinear problems are also dealt with. Strain softening, damage and localization, or certain combinations of material behaviour are investigated, too.

Thus, a short illustration of the advantage of the nonsmooth tangent modulus is presented here.

It is frequently applied in stability analyses that the structure is divided into perfectly rigid parts, and the material properties are concentrated to certain springs at the connecting points, the joints of the structure [1]. By using this kind of illustrative examples the advantage of the nonsmooth tangent modulus can be made evident.

Figure 6a shows the structure of total length  $l$  consisting of two rigid elements of length  $l/2$ , by assuming the middle joint to be in the half of the structural height. The system has one degree of kinematic freedom, since the geometrical state of the structure can be characterized by a single parameter  $q$ . Let this parameter be the angle of rotation  $\vartheta_A$  at the support joint  $A$ , so  $q = \vartheta_A$ .

Consider now the vector  $\mathbf{u}(\mathbf{q})$  of the kinematically admissible displacement functions

$$\mathbf{u} = \begin{bmatrix} u_{Bx} \\ u_{By} \\ u_{Dy} \end{bmatrix} = \begin{bmatrix} l \sin q / 2 \\ l(1 - \cos q) / 2 \\ l(1 - \cos q) \end{bmatrix} = \begin{bmatrix} u_1(q) \\ u_2(q) \\ u_3(q) \end{bmatrix} = \mathbf{u}(q) \quad (74)$$

and the vector of the strains  $\boldsymbol{\varepsilon}(\mathbf{u}(\mathbf{q}))$

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \vartheta_D \\ \vartheta_A \\ \vartheta_B \end{bmatrix} = \begin{bmatrix} q \\ q \\ 2q \end{bmatrix} = \boldsymbol{\varepsilon}(q) \quad (75)$$

As we can see the displacement functions  $\mathbf{u}(q)$  are nonlinear, while the strain functions  $\boldsymbol{\varepsilon}(q)$  are linear in the chosen parameter  $q$ . We will see that due to the strain linearity, the function of the reduced stresses  $\bar{\mathbf{f}}(q)$  is segment by segment linear.



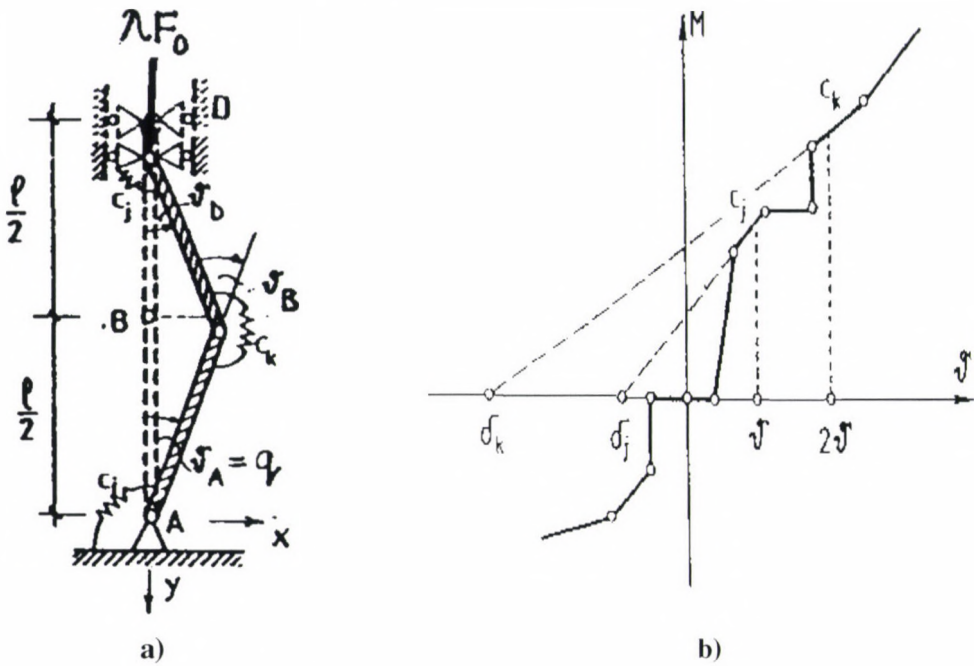


Fig. 6. The example for illustrating the nonsmooth tangent stiffness

The external potential of the dead load  $\lambda \mathbf{F}_o^T = \lambda [0 \ 0 \ F_o]$  is  $\pi_{ex}(\mathbf{q}) = -\lambda \mathbf{F}_o^T \mathbf{u}(\mathbf{q})$ , in the case of a single vertical load  $\lambda F_o$  on the top of the structure. Thus, we have

$$\pi_{ex}(q) = -\lambda F_o l (1 - \cos q) \tag{76}$$

The internal potential depends on the material behaviour which is concentrated to the springs applied at the joints A, B and D of the structure. The springs are characterized equally by the same polygonal moment-rotation function  $M(\vartheta)$  seen in Fig. 6b. A general polygonal stress-strain diagram is considered. Each linear part  $i$  of it, included the horizontal and vertical parts as well, can be characterized by two data: the modulus of elasticity  $c_i$  and the related rotation  $\delta_i$  as the rotation at the intersection of the segments of the polygon with the coordinate axis  $\vartheta$ . General elastic-plastic-damaging behaviour is modelled. The spring constant  $c$  is variable, it changes stepwise linearly within the interval of  $0 \leq c_i \leq \infty$  in any segment  $i$  of the material polygon representing the actual elastic, rigid, or even locking or damaging character of the material phase.

Consider first an elastic nonsmooth problem.

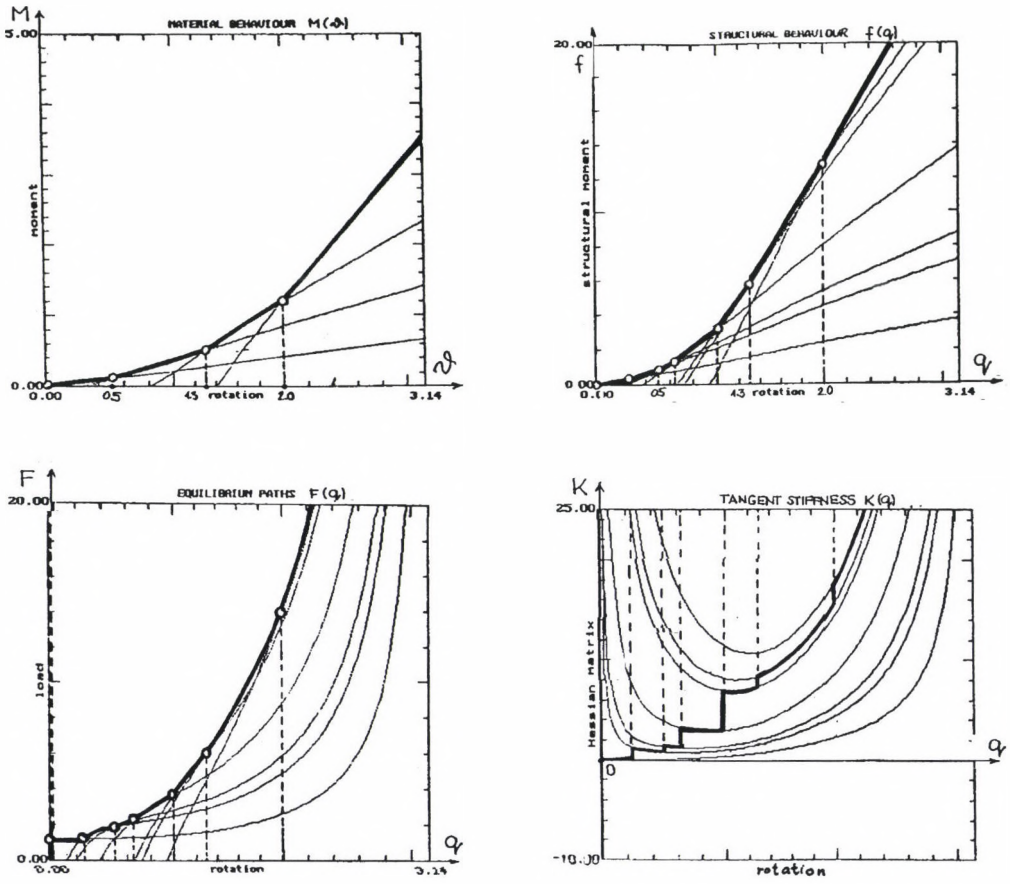


Fig. 7. Stability analysis of structures with nonlinear material by applying polygonal approximation

5.1. Nonsmooth reversible problems

Figure 7a shows the diagram  $M(\vartheta)$  of a linearly hardening elastic material applied uniformly to each spring of the structure. The concerning reduced stresses  $f(q)$  are seen in Fig. 7b. The concerning equilibrium paths  $\lambda(q)$  of the structure are seen in Fi. 7c. For the sake of simplicity this time we follow the behaviour of the structure in the right-hand side interval  $0 \leq q \leq \pi$  only.

We have shown in papers Kurutz [19–24] that the equilibrium path of the structure having nonsmooth material can be obtained as the envelope of the component equilibrium paths related to each segments of the reduced stresses, namely, the structural material behaviour  $f(q)$ .

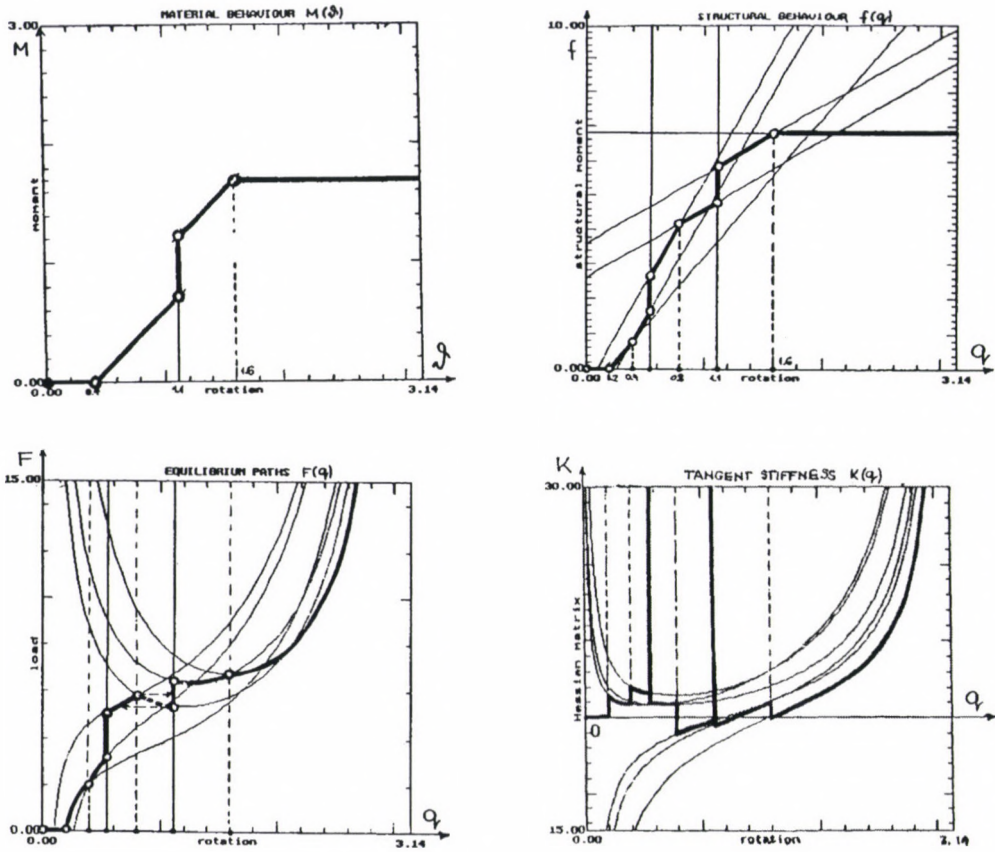


Fig. 8. Stability analysis of structures with originally polygonal material

In the one dimensional case, the stability of equilibrium at the points of the equilibrium paths can be qualified by a simple sign control of the related functions of the tangent stiffness. For qualifying the equilibrium paths, we need the second subdifferential of the total nonsmooth superpotential. The nonsmooth functions of the structural tangent stiffness  $K(q)$  are seen in Fig. 8d. For regular points the tangent stiffness is a single value, while for singular points it is an interval. In order to find the critical load, we consider the inclusion

$$0 \in \det K(q) \tag{77}$$

knowing that the determinant of an interval matrix forms an interval, too. In this one-dimensional case, this matrix has a single element, and, even if it is an interval, it is equal to its determinant in itself. We can conclude that for material hardening (softening) the upper (lower) envelope of the component paths results in the nonsmooth equilibrium path



of the structure. This statement coincides with the fact that in the case of the material softening (hardening), the load bearing capacity of the structure decreases (increases).

Figure 8 show the example of a general reversible polygonal behaviour: a gradually hardening material law with vertical and horizontal jumps. Here, to obtain the nonsmooth tangend modulus and the functions of the structural tangend stiffness, distributional derivatives were applied. In Fig. 8d the intervals of indefinite length are seen related to the jumps in the material behaviour  $M(\vartheta)$  of each joint, and, consequently, the jumps in the structural material behaviour  $f(q)$ .

We can conclude that by applying polygonal approximation for nonlinear constitutive laws and by using the concept of nonsmooth tangend modulus, considerable good solutions can be obtained for nonlinear problems.

## 5.2. Nonsmooth irreversible problems

In contrast to the plastic limit analysis where the plastic hinges exhibit no softening, in many important cases, however, the absence of softening cannot be guaranteed. For example, in pre-stressed reinforcement concrete beams, the softening of the hinge is caused by strain softening of concrete due to micro-fracturing.

Let the elastic-plastic-damaging behaviour of the joints be represented by the material function  $M(\vartheta)$  seen in Fig. 9a. In contrast to the plastic unloading, in the damage zone the unloading moduli are specified individually. Here we do not detail unloading and localization, these questions are analysed in Kurutz [23, 24].

Figure 9b shows the nonsmooth function  $f(q)$  of the structural material behaviour resulted by the simultaneously different material phases of each joints  $A$ ,  $B$  and  $D$ . The simultaneity of the different material phases depends on the actual strains at the joints controlled by the actual rotations  $\varepsilon(q)$ , namely, the compatibility transformations (75). Thus while the joints  $A$  and  $D$  have a rotation  $q$  with the modulus  $c_j$ , the middle joint  $B$  suffers a rotation twice larger using the modulus  $c_k$ . Thus by applying overall unloading, the unloading moduli can be different if the joint  $B$  reaches the damage phase. In Fig. 9c and d the nonsmooth functions of the equilibrium paths  $\lambda(q)$  and the concerning structural tangential stiffness functions  $K(q)$  are illustrated. Due to the gradual strain softening, the tangential stiffness tends to be indicating instabilities.

As a conclusion of the presented one-dimensional examples, we can state that by using polygonal approximation and the nonsmooth tangend modulus, any uniaxial nonlinear and irreversible material behaviour can be handled. However, in the case of higher dimension, the mathematical difficulties increase.

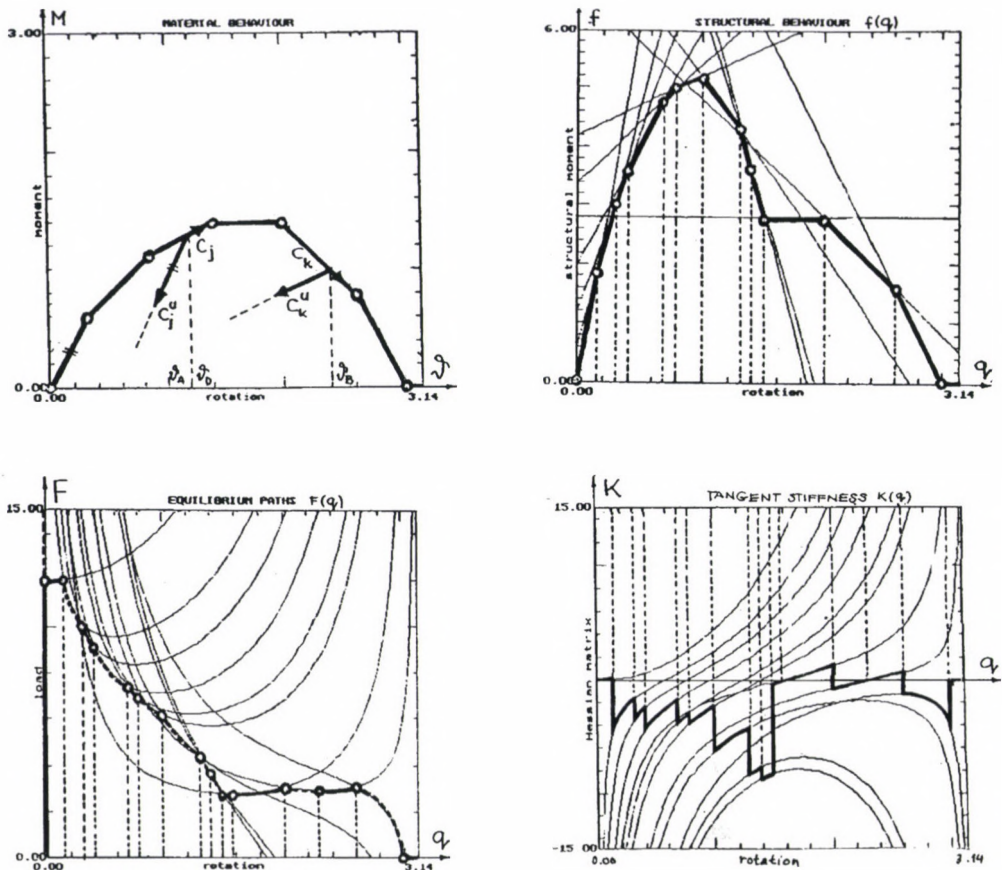


Fig. 9. Stability analysis of structures with damaging material by applying polygonal approximation

## 6. Conclusions

In this paper the material and the structural tangent moduli are in focus. After a short historical review, where the century long development of the tangent modulus was detailed, we introduced the nonsmooth version of it.

For global stability analysis of structures with nonlinear materials, the instantaneously changing tangent modulus is required. This leads to mathematical difficulties. By applying polygonal approximation, nonsmooth problems are obtained. The nonsmooth tangent modulus of the polygonal behaviour can be considered as the nonsmooth tangent modulus of loading-unloading of irreversible materials.

As a conclusion, the tangent modulus containing both material loading and unloading is always multivalued. The nonsmooth tangent modulus related to the break points of a



material polygon forms intervals of finite length, while related to the jump points of a material polygon forms intervals of infinite length concerning to the Dirac-impulse. The nonsmooth tangent modulus of polygonal material behaviour can be applied to the cases of strain softening and damage, too.

The concept of nonsmooth tangent modulus of the material point containing equally loading and unloading or polygonal approximation of nonlinear materials, can be extended to the whole structure, yielding to the nonsmooth structural tangent modulus. The generalized nonsmooth structural tangent modulus is multivalued. In the case of uniaxial material behaviour, it forms a diagonal interval matrix. For a break (jump) type material discontinuity, the intervals are finite (infinite).

For stability analyses of inelastic systems, the concept of tangentially equivalent elastic structure is needed and a quasi-elastic analysis is investigated.

One-dimensional illustrations for simple discrete structures with uniaxial material laws helped to prove the advantage of the nonsmooth material and structural tangent modulus.

### Acknowledgement

The present study was supported by the National Development and Research Foundation (OTKA T015851, OTKA T022622, OTKA T023929, OTKA T025256, MKM FKFP 0397/97).

### References

1. Augusti, G.: Stabilita' di strutture elastiche elementari in presenza di grandi spostamenti, *Atti Accad. Sci. Fis. Mat. Napoli Serie 3.a.* 4. No.5. 1964.
2. Bazant, Z. P., Cedolin, L.: *Stability of Structures. Elastic, Inelastic, Fracture and Damage Theories*, Oxford University Press, New York Oxford, 1991.
3. Benallal, A., Billardon, R., Geymonat, G.: Some mathematical aspects of the damage softening rate problem, In: *Cracking and Damage – Strain Localization and Size Effect*, ed. by J. Mazars, Z.P. Bazant, pp. 247–257, Elsevier, London, New York, 1989.
4. Benallal, A., Billardon, R., Geymonat, G.: Bifurcation and localization in rate-independent materials, In: *Bifurcation and Stability of Dissipative Systems*, Ed. by Q.S. Nguyen, CISM Lecture Notes 327, pp. 1–44, Springer, 1993.
5. Billardon, R., Doghri, I.: Localization bifurcation analysis for damage softening elastoplastic materials, In: *Cracking and Damage - Strain Localization and Size Effect*, ed. by J. Mazars, Z.P. Bazant, pp. 295–303, Elsevier, London, New York 1989.
6. Bruhns, O. T.: Bifurcation problems in plasticity, In: *The Constitutive Law in Thermoplasticity*, CISM Lecture Notes 281, pp. 461–540, ed. by Th. Lehmann, Springer, Wien, New York, 1984.
7. Clarke, F. H.: Generalized gradients and applications, *Trans. Am. Math. Soc.* 205. 247–262, 1975.
8. Del Piero, G., Sampaio, R.: Unified treatment of damage and plasticity based on a new definition of microfracture, *Univ. Udine*, 062, 1989.
9. Halphen, B., Nguyen, Q. S.: *Sur les materiaux standards generalizes*, *J. de Mecanique* Vol. 1. 1975.
10. Hill, R.: A general theory of uniqueness and stability in elastic-plastic solids, *J. Mech. and Phys. of Solids* Vol.6. 236–249. 1958.



11. Hill, R.: Uniqueness criteria and extremum principles in self adjoint problems of continuum mechanics, *J. Mech. Phys. Solids*, 10, 185. 1962.
12. Hill, R.: On the classical constitutive relations for elastic-plastic solids. In: *Recent Progress in Applied Mechanics* 241. Wiley, New York, 1967.
13. Hill, R.: Aspects of invariance in solid mechanics, In: *Advances in Applied Mechanics* Vol. 8. ed. by Yih C. S., Academic Press, 1978.
14. Janson, J., Hult, J.: Fracture mechanics and damage mechanics a combined approach, *Journ. Mec. Appl.* Vol. 1. 69-84. 1977.
15. Kaliszky, S.: The analysis of structures with conditional joints, *Journal of Structural Mechanics*, Vol. 6. 195-210. 1975.
16. Keener, J.P.: *Principles of Applied Mathematics. Transformation and Approximation*, Addison Wesley Publ. Comp. California, 1988.
17. Kurutz, M.: Generalized conditional joints as subdifferential constitutive models, *Zeitschrift für Angewandte Mathematik und Mechanik, ZAMM*, Vol. 65. No. 5. T347-T348. 1985.
18. Kurutz, M.: Analysis of generalized conditional joints as subdifferential constitutive models, *Journal of Mechanics of Structures and Machines*, Vol. 15. No. 2. 123-151. 1987.
19. Kurutz, M.: Stability of simple structures with nonsmooth energy functionals in the one-dimensional case, *Acta Technica Sci. Hung.* Vol. 102. No. 3-4. 291-314. 1989.
20. Kurutz, M.: On the nonsmooth stability analysis, *Zeitschrift für Angewandte Mathematik und Mechanik, ZAMM*, Vol. 72. No. 4. T114-T117. 1991.
21. Kurutz, M.: Stability of structures with nonsmooth nonconvex energy functionals, *European Journal of Mechanics, A/Solids*, Vol. 12. No. 3. 347-385. 1993.
22. Kurutz, M.: Equilibrium paths of polygonally elastic structures, *Journal of Mechanics of Structures and Machines*, Vol. 22. No. 2., 181-210. 1994.
23. Kurutz, M.: Equilibrium paths of polygonally damaging structures, Part one: The nonsmooth nonconvex stability problem, *International Journal of Damage Mechanics*, Vol. 5. No. 1. 16-41. 1996.
24. Kurutz, M.: Equilibrium paths of polygonally damaging structures. Part two: One dimensional example for nonsmooth damage and localization, *International Journal of Damage Mechanics*, Vol. 5. No. 1. 42-67. 1996.
25. Moreau, J. J.: *Fonctionelles soudenferetiables* C. R. Acad. Sc. Paris 257A. 4117-4119. 1963.
26. Moreau, J. J.: *La notion de surpotential et les liaisons unilaterales en elastostatique*, C. R. Acad. Sc. Paris, 267A, 954-957. 1968.
27. Moreau, J. J., Panagiotopoulos, P.D. (eds): *Nonsmooth Mechanics and Applications*, CISM Lecture Notes, No. 302. Springer, NewYork, Wien, 1988.
28. Nguyen, Q. S.: *Stabilite et bifurcation des systemes dissipatifs standards a comportement independant du temps physique*, R. C. Acad. Sc. Paris Vol. 310. 1990.
29. Nguyen, Q. S.: *Bifurcation and stability of time-independent standard dissipative systems*. In: *Bifurcation and Stability of Dissipative Systems*, ed. by Q. S. Nguyen, CISM Lecture Notes 327, pp. 45-94, Springer, 1993.
30. Panagiotopoulos, P. D.: *Nonconvex superpotentials in the sense of F. H. Clarke and applications*, *Mech. Res. Comm.* 8. 335-340. 1981.
31. Panagiotopoulos, P. D.: *Nonconvex energy functions. Hemivariational inequalities and substationary principles*, *Acta Mechanica*, 42. 160-183. 1983.
32. Panagiotopoulos, P. D.: *Inequality Problems in Mechanics and Applications. Convex and Nonconvex Energy Functions*, Birkhauser, Basel, 1985.
33. Panagiotopoulos, P. D.: *Nonconvex superpotentials and hemivariational inequalities*. In: *Quasi-differentiability in mechanics*, ed. by J. J. Moreau and P. D. Panagiotopoulos, CISM Lecture Notes 302, pp. 83-176. Springer, NewYork, Wien, 1988.
34. Raniecki, B., Bruhns, O. T.: *Bounds to bifurcation stresses in solids with nonassociated plastic flow law at finite strain*, *Int. J. Mech. Phys. Solids* Vol. 29. 153-172. 1981.
35. Rice, J. R.: *Inelastic constitutive relations for solids: an internal variable theory and its application in metal plasticity*, *J. Mech. Phys. Solids*, Vol. 19. 1971.

36. Rice, J. R.: The localization of plastic deformations, In: *Theoretical and Applied Mechanics*, ed. by W.T. Koiter, North Holland P.C. pp. 207–220. 1976.
37. Rice, J. R., Rudnicki, J. W.: A note on some features of the theory of localization of deformation, *Int. J. Solids and Structures*, Vol. 16. 596–605. 1980.
38. Rockafellar, R. T.: *Convex Analysis*, Princeton Univ. Press, Princeton, 1970.
39. Shanley, F. R.: The column paradox, *J. Aero. Sci.* 13. 678. 1946.
40. Shanley, F. R.: Inelastic column theory, *J. Aero. Sci.* 14. 261. 1947.
41. Suquet, P. M.: Locking materials and hysteresis phenomena, In: *Unilateral Problems in Structural Analysis*. Eds G. Del Piero and F. Maceri, CISM Lecture Notes 288. pp. 339–373, Springer, Wien, New York, 1985.
42. Thompson, J. M. T., Hunt, G. W.: *A General Theory of Elastic Stability*, Wiley, London, 1973.
43. Thompson, J. M. T., Hunt, G. W.: *Elastic Instability Phenomena*, Wiley, Chichester, 1984.

## POSSIBILISTIC ESTIMATIONS OF IMPERFECTIONS IN STRUCTURAL ANALYSIS\*

Nándori, E.<sup>1</sup> and Kovács, M.<sup>2</sup>

<sup>1</sup>*Technical University of Budapest, Department of Lightweight Structures and Chassis,  
Bertalan L. u. 2, H-1111 Budapest, Hungary*

<sup>2</sup>*Eötvös University, Department of Operations Research, Rákóczi út 5, H-1088 Budapest, Hungary*

(Received: 30 January, 1997)

This work is designed to find a method permitting microgeometrical imperfections occurring in the course of construction and operation of structures to be taken into consideration in structural analysis. We will show that possibilistic solutions can be found by solving linear programming problems. Since good program packages are available for solving this problem by computer, the estimates are obtained quite simply.

### 1. Introduction

In the course of preparation for construction, manufacture or operation of any structure, made of steel or being of mixed construction, imperfections occur which can be included in five groups arbitrarily:

- macrogeometrical imperfections (deviation from the straight of a straight rod, length of a rod element other than its nominal value, etc.);
- microgeometrical imperfections (random changes in wall thickness, changes in cross-section, etc.);
- imperfections resulting from deviation of the magnitude, position and direction of load from the prescribed values;
- errors in boundary conditions;
- material defects.

Distinction shall be made between two cases according to the occurrence and detection of imperfections.

- In the first case, the imperfections are detected by high-accuracy measurements in the course of construction of the structure, and then the structure so defined is analyzed. The structure defined anew by improving the deviations – except material defects –

\*Supported by Hungarian National Science Foundation (OTKA) under the grant # T100 and T019455.



can be recalculated, it does not mean a new problem. Such deterministic procedure can take place in the case of high-value individual structures.

- In the second case, the imperfections cannot be detected accurately because either the costs are too high or the imperfections result from a material defect. An accurate definition of any imperfection will be especially difficult if products of series production are investigated, where imperfections occur randomly.

Discussed in this work are random imperfections occurring in products produced in series because both the reliability and the service life of these products depend on these random phenomena.

Different imperfections in the structure may have different consequences.

The effects of macrogeometrical imperfections occur in statistically indetermined structures only (no kinematic load results from macrogeometrical imperfections in statically determined structures). This random kinematic load resulting from imperfections of this kind has been studied by more authors assuming either the distribution function of the imperfections or certain statistical informations (moments) [1–3].

For certain problems, e.g. for determination of the stresses of ring-shaped beams under internal or external pressure or any other load, there are well-known analytical methods available. However, the geometry of the cross-section may change as a result of some macrogeometrical error and at this instant, also the solution to the problem will be completely different, the rotational-symmetric solution being no longer valid in the case in question.

In [4] the changes in the critical load of tubes of elliptic cross-section under external pressure were investigated. It was shown here that the value of the critical load decreases extremely rapidly if the deviation of the small-axis-to-large-axis ratio increases.

Microgeometrical imperfections affect the stresses of the structure in case of both statically determined and indetermined structures. Obviously, we will obtain stresses other than the calculated ones if the cross-sectional data differ from the calculated values. The case of indetermined structures is more complicated because here the microgeometrical imperfection may be combined with the kinematic load.

The effects of random microgeometrical imperfections on rod structures are discussed in [5–7].

The effects of random microgeometrical imperfections on the stability and supercritical behaviour (buckling, plastic hinge, etc.) have not been investigated yet. This question is especially significant in the description of the behaviour of structures in accidents (overturning, collision of vehicles). Local deviations may considerably affect the behaviour of structures in accidents.

The effects of random deviations of the position, magnitude and direction of load acting upon the structure have not been investigated yet at all.

Interestingly, as a result of random load imperfections that is, deviation of the position and direction of load from the prescribed value, strength problems may face us in certain cases instead of bifurcation stability problem. For example the description of the behaviour of a straight rod under central compression is a stability problem while the investigation of a rod under eccentric pressure is a strength problem. Hence, a signifi-

cant random deviation of the position of load from centric position changes the nature of the problem as well. A similar problem may result from random deviations of the line of action of load; e.g. if the direction of the axial compressive load of a clamped rod is other than axial.

Load imperfections may closely bear upon part of the macrogeometrical imperfections: e.g. if the axis of a straight road differs from the straightness, then the problem facing us is rather a strength problem than a stability problem.

The solution to the problem is fundamentally affected by the boundary conditions. "Pure" boundary conditions are usually defined for the mathematical solutions: no angular displacement is permitted by the bracing or no sinking is permitted by the support, etc. In practice, completely different boundary conditions are usually present in combination, the bracing being, in fact, a partial bracing (elastic bracing, elastic support or elastically embedded hinge, etc.) only. In case of the actual structures made on the basis of a prescribed documentation, variations may occur and the different realized structures operate under different boundary conditions with a different solution series associated with each.

As a rule, the imperfections listed contribute to the damage of the structures. The realized structures in which the imperfections mentioned above may occur incidentally become more sensitive with their load capacity and reliability decreasing in general and the extent of reduction in load capacity and reliability can, and shall, be determined by means of some mathematical method (e.g. probability theory).

The material characteristics used in the calculations are used to take the material defects into consideration. The tensile strength or the yield point and/or the modulus of elasticity are practically characteristic of the material possessing some material defect and thus the real structural characteristics are taken into account in the calculations. Hence, material defects are not dealt with specifically in this work and they are not included among the imperfections investigated in this work although their influence is considerable.

Affected by the imperfections described is not only the static but also the dynamic behaviour of the structure, moreover, certain imperfections may result in nonlinearities in the system.

The properties of the structure also change as a function of time. What we are talking about is not the stabilization of the state adjusted initially (with the analysis performed for these conditions) but the occurrence of any imperfection (wear, loosening, local cracks, etc.) in the course of operation.

The literature is rich in exact or approximate solutions to technical problems of quite a number. Also methods to investigate the effects of different specific imperfections or dimensional imperfections falling within the scope of this study are known [8]. However, the stochastic implications of this that is, the fact that the different imperfections occur randomly and each product shows different properties, have been investigated occasionally and not in full.



## 2. Methods to treat microgeometrical imperfections

This work is designed to find a method permitting microgeometrical imperfections occurring in the course of construction and operation of structures to be taken into consideration in structural analysis.

Methods offering themselves for use in structural analysis are

- probability calculation,
- interval arithmetic,
- fuzzy methods and among these, so-called possibilistic methods.

The conditions and/or the advantages of the use of any of these methods can be specified.

To use methods based on probability theory, the distribution or density functions or at least some moments of the probability variables must be known. Data of a large number about the phenomena associated with these imperfections should be available to obtain these informations, however, such data are seldom available. Thus what we encounter in case of the probability models are rather expert opinions than reliable statistical characteristics. Another considerable disadvantage of the probability calculation methods is the extraordinary complexity of the calculations (calculation of multiple integrals, solution of nonlinear systems even in case of linear problems, etc.). In practical problems, only finite perturbations in the data are possible and thus so-called limited density functions must be used for the actual calculations, this is also a problem contributing to the complexity of the calculations. Another problem is that the coefficients of the stochastic equations to be solved are not independent random variables even in case the parameters of the different imperfections can be considered to be independent.

The advantage of the interval arithmetic is that it is enough to know expert opinion only on the extreme bounds of the occurring imperfections. The calculation method is relatively simple although separation of cases of quite a number is required depending on the sign of the variables. A disadvantage is, however, that no information can be obtained about the rate of the error-propagation.

In fuzzy methods (e.g. [9–10]), the expert's estimate identifies not only the interval where the imperfections can be move but also the expectable reliability of this interval (the higher the value of the membership function, the more the given point can be expected to fall within the perturbation interval). Similarly, changes in the set of possible solutions can be investigated in accordance with the level of acceptance (value of the membership function). In this paper we will show that possibilistic solutions can be found by solving linear programming problems. Since good program packages are available for solving this problem by computer, the estimates are obtained quite simply. The problems solved under conditions acceptable on different acceptability levels supply information on the propagation rate of the imperfections. Moreover, these methods do not require the independence of the variables.



### 3. Mathematical background of possibilistic estimates

#### 3.1. The concept of fuzzy numbers

Let  $x$  be a real variable. If  $\xi$  is a real realization of  $x$ , the indicator function of variable  $x$  will be

$$\chi_{\{x\}} = \begin{cases} 1, & \text{if } x = \xi \\ 0, & \text{if } x \neq \xi \end{cases}$$

If the realization of  $x$  is uncertain, then a fuzzy number  $m : \mathbb{S} \rightarrow [0, 1]$  can be used to give the possibility of any point of the real line

We say that  $\mu : \mathbb{S} \rightarrow [0, 1]$  is a *fuzzy number of finite support* if

- 1)  $\mu(x)$  is continuous on  $\mathbb{S}$ ,
- 2) there exists a triad  $-\infty < b \leq a \leq c < \infty$  such that
  - $\sim \mu(a) = 1$ ;
  - $\sim \mu(x) = 0 \quad \forall x \in (-\infty, b) \cup (c, \infty)$ ;
  - $\sim \mu(x)$  is strictly increasing on  $[b, a]$  and strictly decreasing on  $[a, c]$ .

Due to the above conditions, the subset  $C_r(m) = \{x \in \mathbb{S} : \mu(x) \geq r\}$  of the real line containing the points for which the realization level is at least  $r$  forms a finite interval for any given  $0 < r \leq 1$ .  $[b, c]$  is the support of the fuzzy number for which  $[b, c] = \overline{\{x \in \mathbb{S} : \mu(x) \geq r\}}$  (here the overline indicates the topological closure of the set).

Hereafter we use only fuzzy numbers of finite support, therefore the attribute "of finite support" will be omitted for the most part.

Intervals  $C_r(\mu)$  uniquely determine the side functions

$$\ell(r) = \inf\{x \in \mathbb{S} : x \in C_r(\mu)\}$$

$$u(r) = \sup\{x \in \mathbb{S} : x \in C_r(\mu)\}$$

of the fuzzy number  $\mu$  which have the following properties:

1.  $\ell(r)$  is continuous and strictly increasing on  $[0, 1]$ ;
2.  $\ell(0) = b, \ell(1) = a$ ;
3.  $u(r)$  continuous and strictly decreasing on  $[0, 1]$ ;
4.  $u(0) = c, u(1) = \alpha$ .

Conversely, any pair of functions  $\ell(r), u(r)$ , having the four properties given above, uniquely determines a fuzzy number:

$$\mu(x) = \begin{cases} 1, & \text{if } x = a, \\ 0, & \text{if } x \leq b \text{ or } x \geq c, \\ \sup\{r \in [0, 1] : \ell(r) \leq x\}, & \text{if } x \in [b, c], \\ \sup\{r \in [0, 1] : u(r) \geq x\}, & \text{if } x \in [b, a] \end{cases}$$

Let  $FN$  denote the set of fuzzy numbers.

Let a special subset of fuzzy numbers be defined in the following way:

Let  $g(r)$  be a monoton decreasing continuous function on  $[0, 1]$  with the boundary properties  $g(0) = g_0 < \infty$ ,  $g(1) = 0$ . The fuzzy numbers defined by side functions  $\ell(r) = -\delta g(r) + \alpha$ ,  $u(r) = \delta g(r) + \alpha$  will be called *g-induced symmetric fuzzy numbers*. (In this case  $a = \alpha$ ,  $b = \alpha - \delta$ ,  $c = \alpha + \delta$ . The set of the  $g$ -induced symmetrical fuzzy number will be denoted by  $FN_{g,\alpha}$ .)

With a fixed  $g$  the  $g$ -induced symmetrical fuzzy number is uniquely determined by  $(\alpha, \delta)$ , where  $\alpha$  is the center and  $\delta$  is the spread of the fuzzy number. It is quite easy to see that for the fuzzy number  $\mu = (\alpha, \delta) \in FN_g$  the membership function is given by

$$u(x) = \begin{cases} 1, & \text{if } x = \alpha, \\ g^{-1}\left(\frac{|x - \alpha|}{\delta}\right), & \text{if } |x - \alpha| \leq \delta \text{ and } \delta \neq 0. \\ 0, & \text{otherwise} \end{cases}$$

If  $\delta = 0$ , then  $\mu(x) = \chi_{\{\alpha\}}(x)$ , where  $\chi_A$  is the characteristic function of the set  $A$  (Fig. 1a).

Most frequently, generator function  $g$  is chosen as follows:

a)  $g(r) = 1 - r$ .

In this case  $g_0 = 1$  and

$\sim \mu = (\alpha, 0)$  is the characteristic function of the single point set  $\{\alpha\}$  i.e. it is the description of the exact of  $\alpha$  (see Fig. 1a).

$\sim \mu = (\alpha, \delta)$  is the description of a triangular fuzzy number with the center  $\alpha$  and support set  $[\alpha - \delta, \alpha + \delta]$  (see Fig. 1b).

$$\mu(x) = \begin{cases} 1 - \frac{|x - \alpha|}{\delta}, & \text{if } \alpha - \delta \leq x \leq \alpha + \delta. \\ 0, & \text{otherwise} \end{cases}$$

b)  $g(r) = \sqrt{1 - r}$  (Fig. 1c).

In this case  $g(0) = 1$  and

$$\mu(x) = \begin{cases} 1 - \frac{|x - \alpha|^2}{\delta^2}, & \text{if } \alpha - \delta \leq x \leq \alpha + \delta. \\ 0, & \text{otherwise} \end{cases}$$

c)  $g(r) = (1 - r)^2$  (Fig. 1d).  
 Here  $g(0) = 1$  and

$$\mu(x) = \begin{cases} 1 - \sqrt{\frac{|x - \alpha|}{\delta}}, & \text{if } \alpha - \delta \leq x \leq \alpha + \delta. \\ 0, & \text{otherwise} \end{cases}$$

d)  $g(r) = \sqrt{-\ln((1 - \lambda)r + \lambda)}$  (Fig. 1e).  
 In this case  $g_0 = \sqrt{-\ln \lambda}$  and

$$\mu(x) = \frac{\exp\left(-\frac{|x - \alpha|^2}{\delta^2}\right) - \lambda}{1 - \lambda}.$$

This function is obtained from the normal density function

$$\mu(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{|x - \alpha|^2}{2\delta^2}\right)$$

shifting it by  $-\lambda$  and normalizing the obtained function such that its maximum to be 1. Therefore this fuzzy number can be considered as an analogy of the normal density function cut on the significance level  $\lambda$ .

A fuzzy set  $\mu \in FN^n$  on  $\mathbb{S}^n$  will be considered an a fuzzy vector if it is obtained as a Cartesian product of fuzzy numbers (Fig. 1e), i.e.

$$\mu(\mathbf{x}) = \mu(x_1, \dots, x_n) = \min(\mu(x_1), \dots, \mu(x_n)).$$

### 3.2. Fuzzified functions and relations

Let  $f: \mathbb{S}^m \rightarrow \mathbb{S}$ .  $\hat{f}: F^m \rightarrow F$  is a possibilistic fuzzy extension of  $f$  for the fuzzy vectors  $\mu \in FN^m$  if the possibilistic function value  $\hat{f}(\mu)(y) = \hat{f}(\mu_1, \dots, \mu_m)$  is calculated by the Zadeh's extension principle [9]:



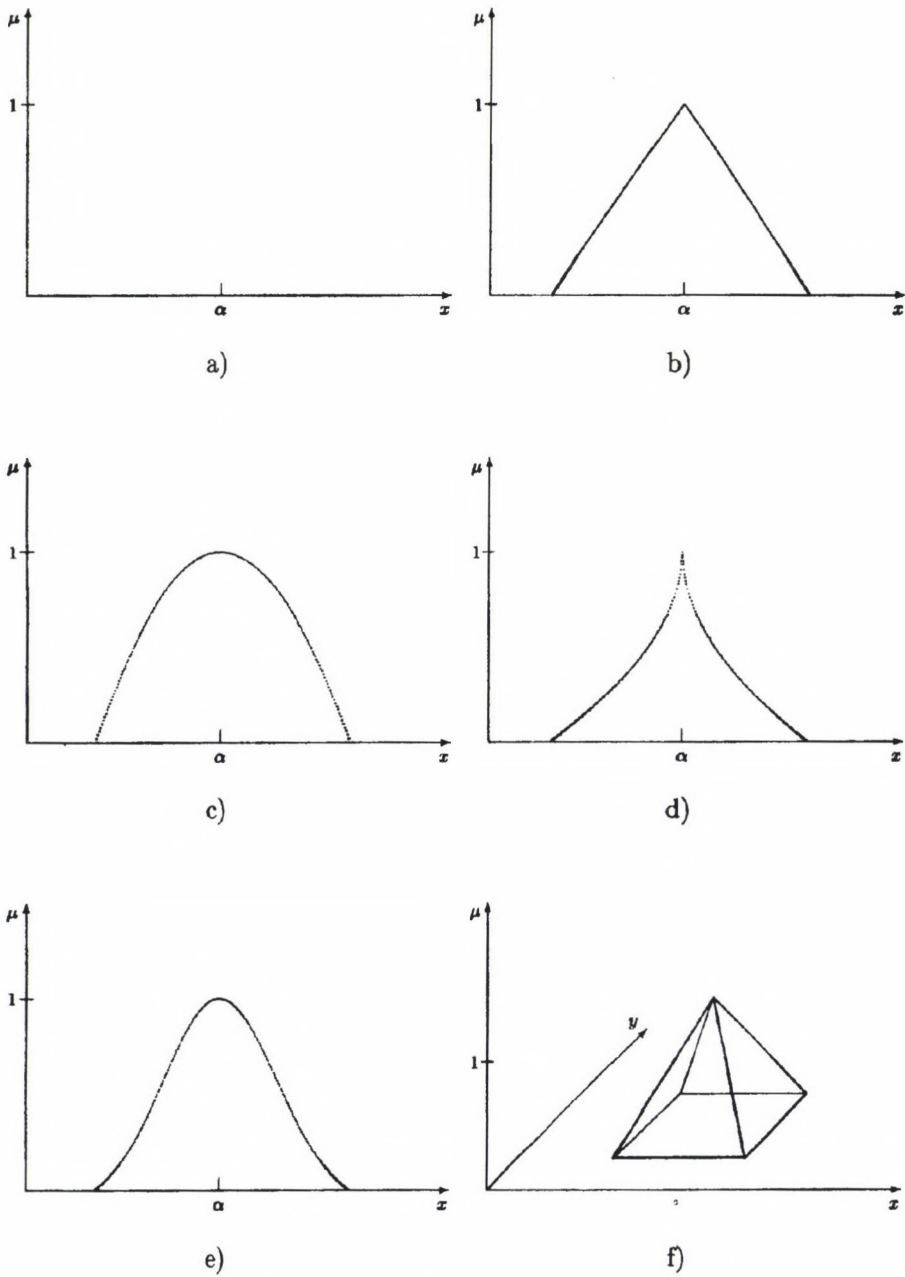


Fig. 1

$$\begin{aligned}
 \text{Poss}(y=f(\mu)) &= \hat{f}(\mu)(y) = \hat{f}(\mu_1, \dots, \mu_m)(y) \\
 &= \begin{cases} \sup_{\mathbf{a} \in A(y)} \mu(\mathbf{a}) & \text{if } A(y) \neq 0 \\ 0, & \text{otherwise} \end{cases} \\
 &= \begin{cases} \sup_{(a_1, \dots, a_m) \in A(y)} \mu(a_1, \dots, a_m), & \text{if } A(y) \neq 0 \\ 0, & \text{otherwise} \end{cases} \\
 &= \begin{cases} \sup_{(a_1, \dots, a_m) \in A(y)} \min(\mu_1(a_1), \dots, \mu_m(a_m)), & \text{if } A(y) \neq 0, \\ 0, & \text{otherwise} \end{cases} \tag{1}
 \end{aligned}$$

where  $A(y) = \{\mathbf{a} = (a_1, \dots, a_m) \in \mathbb{S}^m : y = f(\mathbf{a})\}$ .

Particularly, if  $f$  is a linear function, i.e.  $f(\mathbf{a}, \mathbf{x}) = \sum_{i=1}^n a_i x_i$ , and assume that the possibilistic values of the parameters  $a_i$  are given by  $g$ -induced symmetrical fuzzy number, i.e.  $\mu_i = (\alpha_i, d_i)$ , then it was proved in [11] that the possibilistic function value  $\tilde{f}(\mu, \mathbf{x})(y)$  at the point  $\mathbf{x} \in \mathbb{S}$  belongs also to  $FN_g$  (Fig. 2), namely:

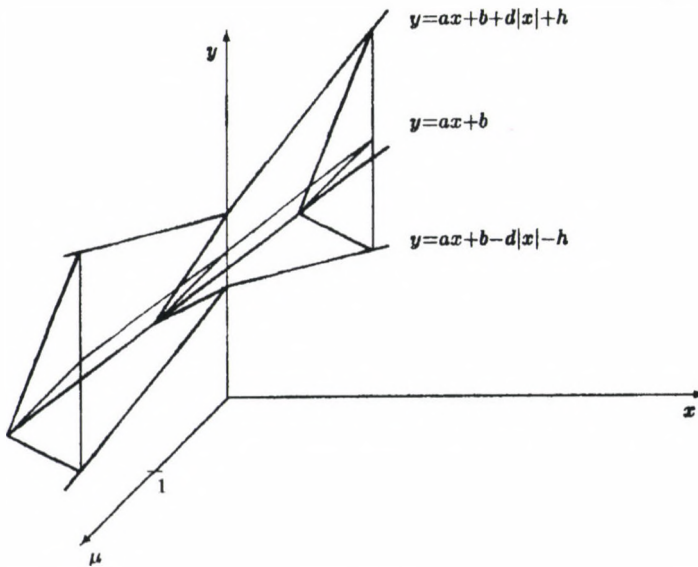


Fig. 2

$$\hat{f}(\boldsymbol{\mu}, \mathbf{x}) = \left( \sum_{i=1}^n \alpha_i x_i, \sum_{i=1}^n d_i |x_i| \right),$$

otherwise

$$\begin{aligned} \text{Poss} \left( y = \sum_{i=1}^n a_i x_i \right) &= \hat{f}(\boldsymbol{\mu}, \mathbf{x})(y) = \\ &= \begin{cases} 1, & \text{if } y = \sum_{i=1}^n \alpha_i x_i \\ g^{-1} \left( \frac{\left| y - \sum_{i=1}^n \alpha_i x_i \right|}{\sum_{i=1}^n d_i |x_i|} \right), & \text{if } \sum_{i=1}^n d_i |x_i| \neq 0 \text{ and} \\ & \left| y - \sum_{i=1}^n \alpha_i x_i \right| \leq g(0) \sum_{i=1}^n d_i |x_i| \\ 0, & \text{otherwise} \end{cases} \end{aligned}$$

Let the possibilistic value of the equality of two fuzzy number be defined by the following formula

$$\text{Poss}(\mu = \nu) = \sup_{x=y} \min(\mu(x), \nu(x)).$$

It is easy to see ([11]) that

$$\text{Poss}(\mu = \nu) = \text{Poss}(\mu - \nu, \chi_{\{0\}}),$$

where the membership function of  $\mu = \nu$  is computed by (1). Moreover, it is also obvious that

$$\text{Poss}(\mu = \chi_{\{0\}}) = \mu(0).$$

If  $\mu = (a, \delta) \in FN_g$  and  $\chi_{\{0\}} = (0, 0)$ , then

$$\text{Poss}(\mu = \chi_{\{0\}}) = \begin{cases} 1, & \text{if } \alpha = 0 \\ g^{-1} \left( \frac{|\alpha|}{\delta} \right) & \text{if } |\alpha| \leq \delta \text{ and } \delta \neq 0 \\ 0, & \text{otherwise} \end{cases}$$



3.3. Error estimation for linear equation systems using possibilistic methods

Let the linear system

$$\sum_{j=1}^n a_{ij} x_j - a_{i0} = 0, \quad i = 1, \dots, m \tag{2}$$

be given, where the error tolerance of the coefficients is modelled by possibilistic methods, i.e. fuzzy numbers  $\tilde{a}_{ij} \in FN_g, i = 1, \dots, m, j = 0, \dots, n, \tilde{a}_{ij} = (\alpha_{ij}, d_{ij})$  are considered given instead of  $a_{ij} \in \mathbb{S}$ . In that case the possibilistic value of the solution of the system is a fuzzy set on  $\mathbb{S}^n$  with the membership function

$$\sigma(\mathbf{x}) = \min_{i=1, \dots, m} \sigma_i(\mathbf{x})$$

where

$$\sigma_i(\mathbf{x}) = \text{Poss} \left( \sum_{j=1}^n \tilde{a}_{ij} x_j - \tilde{a}_{i0} = \chi_{\{0\}} \right) = \begin{cases} 1, & \text{if } \sum_{j=1}^n \alpha_{ij} x_j - \alpha_{i0} = 0 \\ g^{-1} \left( \frac{\left| \sum_{j=1}^n \alpha_{ij} x_j - \alpha_{i0} \right|}{\sum_{j=1}^n d_{ij} |x_j| + d_{i0}} \right), & \text{if } \sum_{j=1}^n d_{ij} |x_j| + d_{i0} \neq 0 \text{ and} \\ & \left| \sum_{j=1}^n \alpha_{ij} x_j - \alpha_{i0} \right| \leq g(0) \left( \sum_{j=1}^n d_{ij} |x_j| + d_{i0} \right) \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

is the possibilistic value of the satisfaction of the  $i$ -th equation.

If  $\mathbf{x}^*$  is the solution of the exact system (2), then  $\sigma_i(\mathbf{x}^*) = 1$  for every  $i \in \{1, \dots, m\}$ .

The conditions of the first and second rows in  $\sigma_i(\mathbf{x})$  determine a neighbourhood of the solution of the exact system such that all interior points from this neighbourhood have positive possibility. For this reason, the solution set of the inequality system

$$\left| \sum_{j=1}^n \alpha_{ij} x_j - \alpha_{i0} \right| \leq g(\lambda) \left( \sum_{j=1}^n d_{ij} |x_j| + d_{i0} \right), \quad i = 1, \dots, n \quad (4)$$

describes the points which can be considered a solution with a possibility at least  $0 \leq \lambda \leq 1$ , i.e. for which  $\sigma_i(\mathbf{x}) \geq \lambda \forall i \in \{1, \dots, m\}$  (see Fig. 5 in the Example 2).

To find the possible perturbation interval for every coordinate of the solution let us consider the optimization problems

$$\min x_i \quad \text{and} \quad \max x_i, \quad i = 1, \dots, m \quad (5)$$

subject to (4). With  $\lambda = 0$  the solutions of these problems define the lower and upper bounds of the perturbed solution of (2), that is these values define an  $n$ -dimensional interval which comprises all the solutions of the perturbed system (Fig. 5).

The constraints of the optimization problems (5)–(4) are piecewise linear and there is no nonnegativity condition for the variables.

However, problems (5)–(4) can be linearized introducing new variables. Let

$$u_j = \max(0, x_j) \quad \text{and} \quad v_j = \max(0, -x_j), \quad j = 1, \dots, n.$$

Then

$$x_j = u_j - v_j, \quad |v_j| = u_j + v_j, \quad u_j, v_j \geq 0 \quad \text{and} \quad u_j v_j = 0.$$

Carrying out the substitutions and omitting the conditions  $u_j v_j = 0$  one obtains  $2n$  linear programming problems, which differ only in their objective function. Therefore the complexity of the calculation can be reduced solving the duals of the problems (5)–(4), because in this case the  $2n$  problems can be formulated by one common simplex tableau. In the course of the solution, condition  $u_j v_j = 0$  can be taken into consideration when choosing the pivot element under condition that variables  $u_j$  and  $v_j$  cannot get into the basis simultaneously. However, the disadvantage of this method is that the exclusion one of the pair  $u_j, v_j$  cannot be ensured automatically by every one of the LP packages and thus this must be ensured individually.

Besides the definition of the limits of error for each coordinate, also the limits of error for any of the linear function of  $\mathbf{x}$  can be determined without increasing the computational complexity of the problem, since also here, the constraints conditions are the same as discussed above.

### 4. Calculation of rod systems

The fundamental equation of the rod system by matrix force method is

$$\mathbf{B}^T \mathbf{R} \mathbf{B} \mathbf{x} + \mathbf{B}^T \mathbf{R} \mathbf{a} = 0, \tag{6}$$

where

- B** – matrix of  $\mathbb{S}^{m \times n}$ , it describes the displacements of the statically determined basic system under unit loads on the points of the virtual intersection;
- R** – blockdiagonal flexibility matrix of  $\mathbb{S}^{m \times m}$ ;
- a** – vector of  $\mathbb{S}^m$ , it describes the displacements of the basic system under external loads;
- x** – vector of  $\mathbb{S}^n$  for the statically indetermined internal forces (moments) acting on the points of virtual intersections;
- L = Bx + a** – internal forces of the indetermined structure;
- m** – number of rod (beam sections  $\times k$ , where
 
$$k = \begin{cases} 1 & \text{for constant} \\ 2 & \text{for linear (bending moments)} \\ 3 & \text{for piecewise second degree} \\ & \text{internal force (moment) function} \end{cases};$$
- n** – degree of redundancies.

The uncertain informations are given in the flexibility matrix **R**. The structure of this matrix is

$$\mathbf{R} = \text{diag} \langle r_1 \mathbf{M}_1, \dots, r_k \mathbf{M}_k \rangle,$$

where

- $r_i = \frac{\ell_i}{6IE}$  for the bending moments under concentrated load, while  $r_i = \frac{\ell_i}{30IE}$  for bending moments under distributed load of constant intensity per section, where

- ~  $\ell_i$  is the *i*-th section length;
- ~  $IE$  is the stiffness for bending;

- $\mathbf{M}_i, i = 1, \dots, k$  are symmetric positive definite matrices and  $\sum_{i=1}^k \dim \mathbf{M}_i = m$ . For

example,

- ~  $\mathbf{M}_i = \mathbf{I}$  (unit matrix) for constant internal forces,
- ~  $\mathbf{M}_i = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$  for linearly varying internal forces,
- ~  $\mathbf{M}_i = \begin{pmatrix} 4 & 2 & -1 \\ 2 & 16 & 2 \\ -1 & 2 & 4 \end{pmatrix}$  for piecewise second order internal forces.



Every  $\mathbf{M}_i$  can be decomposed in the form  $\mathbf{M}_i = \mathbf{C}_i^T \mathbf{C}_i$ . For example, if  $\mathbf{M}_i = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ ,

$$\text{then } \mathbf{C}_i = \begin{pmatrix} \frac{\sqrt{3+1}}{2} & \frac{\sqrt{3-1}}{2} \\ \frac{\sqrt{3-1}}{2} & \frac{\sqrt{3+1}}{2} \end{pmatrix}.$$

Consequently

$$\mathbf{R} = \text{diag} \left\langle \mathbf{C}_1^T \text{diag} \langle r_1 \rangle \mathbf{C}_1, \dots, \mathbf{C}_k^T \text{diag} \langle r_k \rangle \mathbf{C}_k \right\rangle$$

Let the matrix  $\mathbf{B}$  and the vector  $\mathbf{a}$  be partitioned according to the block structure of  $\mathbf{R}$ :

$$\mathbf{B}^T = [\mathbf{B}_1^T, \dots, \mathbf{B}_k^T], \quad \mathbf{a}^T = [\mathbf{a}_1^T, \dots, \mathbf{a}_k^T].$$

With this partition we obtain

$$\mathbf{B}^T \mathbf{R} \mathbf{B} = [\mathbf{B}_1^T \mathbf{C}_1^T, \dots, \mathbf{B}_k^T \mathbf{C}_k^T] \text{diag} \langle \text{diag} \langle r_1 \rangle, \dots, \text{diag} \langle r_k \rangle \rangle \begin{bmatrix} \mathbf{C}_1 \mathbf{B}_1 \\ \vdots \\ \mathbf{C}_k \mathbf{B}_k \end{bmatrix},$$

and

$$\mathbf{B}^T \mathbf{R} \mathbf{a} = [\mathbf{B}_1^T \mathbf{C}_1^T, \dots, \mathbf{B}_k^T \mathbf{C}_k^T] \text{diag} \langle \text{diag} \langle r_1 \rangle, \dots, \text{diag} \langle r_k \rangle \rangle \begin{bmatrix} \mathbf{C}_1 \mathbf{a}_1 \\ \vdots \\ \mathbf{C}_k \mathbf{a}_k \end{bmatrix}.$$

Introducing the notations

$$\bar{\mathbf{B}} = \begin{bmatrix} \mathbf{C}_1 \mathbf{B}_1 \\ \vdots \\ \mathbf{C}_k \mathbf{B}_k \end{bmatrix}, \quad \bar{\mathbf{a}} = \begin{bmatrix} \mathbf{C}_1 \mathbf{a}_1 \\ \vdots \\ \mathbf{C}_k \mathbf{a}_k \end{bmatrix}, \quad \bar{\mathbf{R}} = \text{diag} \langle \text{diag} \langle r_1 \rangle, \dots, \text{diag} \langle r_k \rangle \rangle$$

the fundamental equation will be obtained in the form

$$\bar{\mathbf{B}}^T \bar{\mathbf{R}} \bar{\mathbf{B}} \mathbf{x} + \bar{\mathbf{B}}^T \bar{\mathbf{R}} \bar{\mathbf{a}} = 0, \tag{7}$$

where  $\bar{\mathbf{R}}$  is now a diagonal matrix consisting of the values  $r_i$  of uncertain information about the flexibility matrix only. Thus upon investigation of the error propagation, no distinctions need be made between the cases with diagonal and blockdiagonal flexibility matrix as the propagation is structurally similar. (Hereinafter we use no overlining for notation by which we understand either the original or the transformed  $\mathbf{B}$  and  $\mathbf{a}$  depending on the flexibility matrix.)

Let

$$\mathbf{b}_i^T = [\mathbf{b}_{i1}^T, \dots, \mathbf{b}_{ik}^T] \quad \text{and} \quad \mathbf{a}^T = [\mathbf{a}_1^T, \dots, \mathbf{a}_k^T]$$

be denote the partitioned form of the  $i$ -th row vector of matrix  $\mathbf{B}^T$  and the vector  $\mathbf{a}^T$ , respectively. With this notation the system of equations (6) has the following form:

$$\sum_{j=1}^n \left( \sum_{s=1}^k \langle \mathbf{b}_{is}^T, \mathbf{b}_{js}^T \rangle r_s \right) x_j + \sum_{s=1}^k \left| \langle \mathbf{b}_{is}^T, \mathbf{a}_s \rangle \right| r_s = 0, \quad i = 1, \dots, n \tag{8}$$

where  $\langle \cdot, \cdot \rangle$  denotes the scalar product of two vectors.

Assume that the uncertain values  $r_j$  are given by the fuzzy numbers  $\tilde{r}_i = (r_i, \varepsilon_i r_i) \in FN_g$ , i.e. it is assumed that the possible perturbation of the values  $r_i$  is direct proportion to the nominal values. Since the coefficients of (6) are linearly dependent on  $r_i$ , these coefficients are also fuzzy numbers in  $FN_g$  given by

$$\tilde{f}_{ij} = \left( \sum_{s=1}^k \langle \mathbf{b}_{is}^T, \mathbf{b}_{js}^T \rangle r_s, \sum_{s=1}^k \left| \langle \mathbf{b}_{is}^T, \mathbf{b}_{js}^T \rangle \right| r_s \varepsilon_s \right), \quad i, j = 1, \dots, n,$$

and

$$\tilde{f}_{i0} = \left( \sum_{s=1}^k \langle \mathbf{b}_{is}^T, \mathbf{a}_s \rangle r_s, \sum_{s=1}^k \left| \langle \mathbf{b}_{is}^T, \mathbf{a}_s \rangle \right| r_s \varepsilon_s \right), \quad j = 1, \dots, n,$$

Using the formula (3) one can calculate the possibilistic value of that the  $i$ -th equation is being satisfied.

The conditions according to (4) can be obtained by substitutions

$$\alpha_{ij} = \sum_{s=1}^k \langle \mathbf{b}_{is}^T, \mathbf{b}_{js}^T \rangle r_s, \quad i, j = 1, \dots, n,$$

$$\alpha_{i0} = \sum_{s=1}^k \langle \mathbf{b}_{is}^T, \mathbf{a}_s \rangle r_s, \quad i = 1, \dots, n,$$

$$d_{ij} = \sum_{s=1}^k \left| \langle \mathbf{b}_{is}^T, \mathbf{b}_{js}^T \rangle \right| r_s \varepsilon_s, \quad i, j = 1, \dots, n,$$

$$d_{i0} = \sum_{s=1}^k \left| \langle \mathbf{b}_{is}^T, \mathbf{a}_s \rangle \right| r_s \varepsilon_s, \quad i = 1, \dots, n.$$

Essentially, the perturbation interval of the solution can be defined without introducing new variables if the solution associated with the nominal value of  $r_i$  or at least the sign of the coordinates of the solution is known. This latter is often immediately obtained on the basis of technical considerations with regard to the forces acting on the points of virtual intersection. In such cases, the coordinates need to be optimized only in that ortant in which the nominal solution is present. This permits the symbols for absolute value to be removed from (4). Should 0 be obtained for one of the perturbation limits of any coordinate, it must be examined separately whether there is a solution in the ortant providing an opposite sign for this coordinate. If so, the system cannot necessarily be considered stable.

The final stress of the redundant structure can also be estimated by means of the above method.

Let  $\beta_i$  denote the  $i$ -th row vector of the matrix  $\mathbf{B}$ . Then the  $i$ -th coordinate of  $\mathbf{L} = \mathbf{B}\mathbf{x} + \mathbf{a}$  is  $l_i(\mathbf{x}) = \langle \beta_i, \mathbf{x} \rangle + a_i$ . The interval of the possible perturbation of the different force coordinates is given by  $[\min l_i(\mathbf{x}), \max l_i(\mathbf{x})]$ , where the minimum and maximum should be taken subject to the constraints given by (4). Upon stability considerations, here we investigate the domain without sign reversal only.

Note that the perturbation interval obtained by the above method for the final internal force vector is in general narrower than it would be obtained if the extreme values of the forces associated with the points of the virtual intersections were used to determine the extreme values of the coordinates of  $\mathbf{L}$  because the feasible domain of the latter problem is an  $n$ -dimensional interval containing the actual possible range.

## 5. Examples

### Example 1

Let us consider the structure illustrated in Fig. 3.

The parameters of the structure are given by the matrices  $\mathbf{B}$ ,  $\mathbf{R}$  and the vector  $\mathbf{a}$  as follows

$$\mathbf{B} = \begin{bmatrix} 0 \\ 2 \\ 1 \\ 2 \\ -2 \\ 4 \\ 2 \\ 2 \end{bmatrix}, \quad \mathbf{R} = \begin{bmatrix} 2 & 1 & & & & & & & \\ 1 & 2 & & & & & & & \\ & & 2 & 1 & & & & & \\ & & 1 & 2 & & & & & \\ & & & & 4 & 2 & & & \\ & & & & 2 & 4 & & & \\ & & & & & & 7.5 & & \\ & & & & & & & 15 & \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 8 \\ 0 \\ 16 \\ 0 \\ 8 \end{bmatrix}.$$



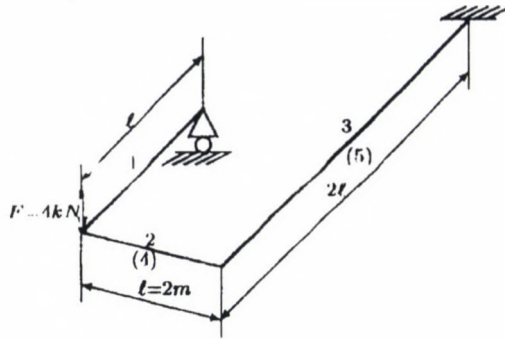


Fig. 3

Here the nominal value of the uncertain flexibility matrix parameters

$$r_1 = 1, r_2 = 1, r_3 = 2, r_4 = 7.5, r_5 = 15.$$

The system to be solved, depending on parameters of the flexibility matrix, is

$$(8r_1 + 8r_2 + 24r_3 + 4r_4 + 4r_5)x + (32r_2 + 96r_3 + 16r_5) = 0,$$

that is the nominal solution is

$$x^* = -3.013,$$

and the final internal force vector belonging to the nominal solution is

$$\mathbf{L}^T = [ \underbrace{0, -6.026}_1, \underbrace{0, 1.1974}_2, \underbrace{6.026, 3.948}_3, \underbrace{-6.026}_4, \underbrace{1.974}_5 ].$$

Here the numbers under the braces indicate the section numbers.

Let the fuzzy elements of the flexibility matrix be triangular fuzzy numbers (i.e. generated by the side function  $g(t) = 1 - t$  given by

$$\tilde{r}_1 = (1, \varepsilon), \quad \tilde{r}_2 = (1, \varepsilon), \quad \tilde{r}_3 = (2, 2\varepsilon), \quad \tilde{r}_4 = (7.5, 7.5\varepsilon), \quad \tilde{r}_5 = (15, 15\varepsilon),$$

where  $\varepsilon < 1$ . Then the fuzzy equation to be solved is

$$(334, 334\varepsilon)x + (704, 704\varepsilon) = 0.$$

From this we obtain the membership function of the fuzzy set of possible solutions

$$\sigma(x) = \begin{cases} 1 - \frac{|154x + 464|}{(154|x| + 464)\varepsilon}, & \text{if } |154x + 464| \leq (154|x| + 464)\varepsilon. \\ 0, & \text{otherwise} \end{cases}$$

The perturbation interval obtained by optimization is

$$\frac{1+\varepsilon}{1-\varepsilon}x^* \leq x \leq \frac{1-\varepsilon}{1+\varepsilon}x^*.$$

For the vector of final internal forces we obtain that

$$\mathbf{B}\mathbf{x} + \mathbf{a} \ni \left[ \begin{array}{c} [0, 0] \\ \left[ 2\frac{1+\varepsilon}{1-\varepsilon}x^*, 2\frac{1-\varepsilon}{1+\varepsilon}x^* \right] \\ [0, 0] \\ \left[ 2\frac{1+\varepsilon}{1-\varepsilon}x^*+8, 2\frac{1-\varepsilon}{1+\varepsilon}x^*+8 \right] \\ \left[ -2\frac{1-\varepsilon}{1+\varepsilon}x^*, -2\frac{1+\varepsilon}{1-\varepsilon}x^* \right] \\ \left[ 4\frac{1+\varepsilon}{1-\varepsilon}x^*+16, 4\frac{1-\varepsilon}{1+\varepsilon}x^*+16 \right] \\ \left[ 2\frac{1+\varepsilon}{1-\varepsilon}x^*, 2\frac{1-\varepsilon}{1+\varepsilon}x^* \right] \\ \left[ 4\frac{1+\varepsilon}{1-\varepsilon}x^*+8, 4\frac{1-\varepsilon}{1+\varepsilon}x^*+8 \right] \end{array} \right] \left. \begin{array}{l} \vphantom{\left[ \right.} \right\} 1 \\ \vphantom{\left[ \right.} \right\} 2 \\ \vphantom{\left[ \right.} \right\} 3 \\ \vphantom{\left[ \right.} \right\} 4 \\ \vphantom{\left[ \right.} \right\} 5 \end{array} \right.$$

Here the numbers after the brackets refer to the corresponding section.

Markedly, the domain of imperfection obtained for the internal forces (bending moments or torsion) will not be symmetric for the nominal values even in this relatively simple case.

### Example 2

Let the examined structure be given in Fig. 4.

Here

$$\mathbf{B} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{R} = \begin{bmatrix} 2 & 1 & & & & \\ 1 & 2 & & & & \\ & & 2 & 1 & & \\ & & 1 & 2 & & \\ & & & & 4 & 2 \\ & & & & 2 & 4 \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 2000 \\ 0 \\ 0 \end{bmatrix}.$$

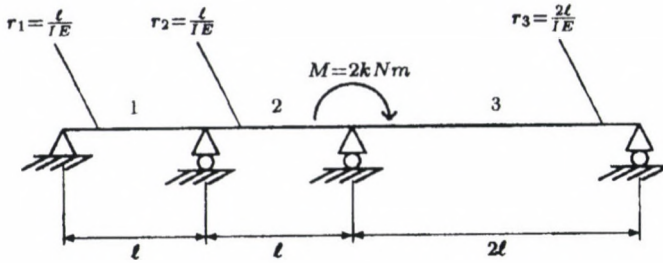


Fig. 4

The equation system depending on the uncertain parameter of the flexibility matrix is

$$\begin{bmatrix} 2(r_1 + r_2) & r_2 \\ r_2 & 2(r_2 + r_3) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 2000r_2 \\ 4000r_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

For the nominal values of the parameters we have

$$r_1 = 1, r_2 = 1, r_3 = 2.$$

Then the nominal solution will be

$$x_1^* = -347.826, x_2^* = -608.697,$$

furthermore, for the stress vector belonging to the nominal forces we obtain

$$\mathbf{L}^T = \left[ \underbrace{0, -347.826}_1, \underbrace{-347.826, 1391.303}_2, \underbrace{-608.697, 0}_3 \right] \quad \text{number of sections}$$

If the triangular fuzzy numbers

$$\tilde{r}_1 = (r_1, \varepsilon r_1), \quad \tilde{r}_2 = (r_2, \varepsilon r_2), \quad \tilde{r}_3 = \left( r_3, \frac{3}{2} \varepsilon r_3 \right)$$

are used to model the imperfect flexibility matrix elements, the membership function of the fuzzy solution set will be determined by the minima of the fuzzy sets

$$\sigma_1(x_1, x_2) = 1 - \frac{|4x_1 + x_2 + 2000|}{\varepsilon(4|x_1| + |x_2| + 2000)}$$

and

$$\sigma_2(x_1, x_2) = 1 - \frac{|x_1 + 10x_2 + 4000|}{\varepsilon(|x_1| + 4|x_2| + 4000)}.$$





Example 3

The structure investigated in this example is given in Fig. 6. Data matrices associated with this structure are

$$\mathbf{B} = \begin{bmatrix} 0 & 0 & 0 \\ -1.414 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -1.414 \\ 0 & 0 & 0 \\ 1.414 & -2 & 0 \\ 0 & 0 & 0 \\ 0 & 2 & -1.414 \\ 0 & 0 & 0 \\ -0.707 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -0.707 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ -0.707 & 0 & 0 \\ 0 & 0 & -0.707 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} 0 \\ 20 \\ 20 \\ 60 \\ 60 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

} bending moments  
  
  
  
  
  
  
  
  
  
  
  
  
 } normal forces

$$\mathbf{R} = \text{diag} \left\langle \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, 60, 120, 60, 120, 60, 60, 60, 60, 60 \right\rangle.$$

Let the nominal values of the flexibility matrix elements be

$$r_1 = 1, r_2 = 2, r_3 = 1, r_4 = 1, r_5 = 1, r_6 = 60, r_7 = 120,$$

$$r_8 = 60, r_9 = 120, r_{10} = r_{11} = r_{12} = r_{13} = 60.$$

Then the system corresponding to (6) is formulated as follows:

$$\begin{aligned}
 (4r_1 + 4r_4 + 0.5r_6 + r_{10} + 0.5r_{11})x_1 - 5.656r_4x_2 - 56.568r_1 &= 0 \\
 -5.656r_4x_1 + (24r_2 + 8r_4 + 8r_5 + r_7 + r_9)x_2 - 2.828r_5x_3 - 480r_2 &= 0 \\
 -2.828r_5x_2 + (4r_3 + 4r_5 + 0.5r_8 + 0.5r_{12} + r_{13})x_3 - 169.704r_3 &= 0.
 \end{aligned}$$

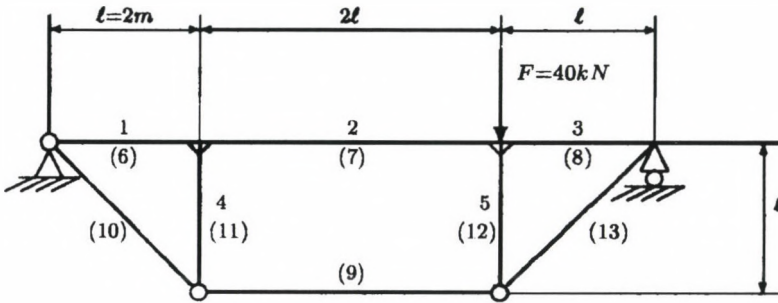


Fig. 6

From this system we obtain the nominal solution

$$x_1 = 0.054, \quad x_2 = 3.157, \quad x_3 = -1.256,$$

and the corresponding final internal force vector

$$\mathbf{L}^T = [0, -0.713, -0.6.314, -6.314, 1.776, 0, -5.601, 0, 8.090, 0, -0.356, -3.157, 0.888, 3.157, 0.581, -0.356, 0.888, -1.256].$$

If the imperfect flexibility matrix element are modeled by triangular fuzzy number with the center of their nominal values and spreads equal to 10% of the nominal values, i.e.

$$\begin{aligned} \tilde{r}_1 = \tilde{r}_3 = \tilde{r}_4 = \tilde{r}_5 &= (1, 0.1), \quad \tilde{r}_2 = (2, 0.2), \\ \tilde{r}_6 = \tilde{r}_8 = \tilde{r}_{10} = \tilde{r}_{11} = \tilde{r}_{12} = \tilde{r}_{13} &= (60, 6), \quad \tilde{r}_7 = \tilde{r}_9 = (120, 12), \end{aligned}$$

then we have to solve the possibilistic system of equation

$$\begin{aligned} (128, 12.8)x_1 + (-5.657, 0.566)x_2 + (0, 0)x_3 + (-56.568, 5.657) &= (0, 0) \\ (-5.657, 0.566)x_1 + (304, 30.4)x_2 + (-2.828, 0.283)x_3 + (-960, 96) &= (0, 0) \\ (0, 0)x_1 + (-2.828, 0.283)x_2 + (128, 12.8)x_3 + (169.704, 16.97) &= (0, 0) \end{aligned}$$

Each equation determines a fuzzy set of the own possible solutions with the membership functions

$$\begin{aligned} \sigma_1(x_1, x_2, x_3) &= 1 - \frac{|128x_1 - 5.656x_2 - 56.56|}{12.8|x_1| + 0.566|x_2| + 5.656} \\ \sigma_2(x_1, x_2, x_3) &= 1 - \frac{|-5.657x_1 + 304x_2 - 2.828x_3 - 960|}{0.566|x_1| + 30.4|x_2| + 0.283|x_3| + 96} \\ \sigma_3(x_1, x_2, x_3) &= 1 - \frac{|-2.828x_2 + 128x_3 + 169.704|}{0.283|x_2| + 12.8|x_3| + 16.97} \end{aligned}$$



From here, taking into consideration the signs of the nominal solution, the inequality system

$$\begin{aligned}
 128(1 - \varepsilon)x_1 - 5.656(1 + \varepsilon)x_2 &\leq 56.57(1 + \varepsilon) \\
 -128(1 + \varepsilon)x_1 + 5.656(1 - \varepsilon)x_2 &\leq 56.57(1 - \varepsilon) \\
 -5.657(1 + \varepsilon)x_1 + 304(1 - \varepsilon)x_2 + 2.828(1 - \varepsilon)y_3 &\leq 960(1 + \varepsilon) \\
 5.567(1 - \varepsilon)x_1 - 304(1 - \varepsilon)x_2 - 2.828(1 - \varepsilon)y_3 &\leq -960(1 - \varepsilon) \\
 -2.828(1 + \varepsilon)x_2 - 128(1 + \varepsilon)y_3 &\leq -169.704(1 - \varepsilon) \\
 2.828(1 + \varepsilon)x_2 + 128(1 - \varepsilon)y_3 &\leq 169.704(1 + \varepsilon)
 \end{aligned}$$

determine the perturbation range, where  $y_3 = -x_3$  and  $\varepsilon = 0.1$ . Performing the optimization for the coordinates of the internal forces and final internal forces, respectively, we obtain

$$x_1 \in [0.456, 0.749], \quad x_2 \in [2.581, 3.867], \quad x_3 \in [-1.564, -1.0278],$$

and

$$\mathbf{Bx} + \mathbf{a} \in \left[ \begin{array}{c} [0, 0] \\ [18.94, 19.36] \\ [12.27, 14.85] \\ [-52.27, -54.85] \\ [61.45, 62.21] \\ [0, 0] \\ [-7.01, -4.20] \\ [0, 0] \\ [6.62, 9.90] \\ [0, 0] \\ [-0.48, -0.32] \\ [-3.87, -2.58] \\ [0.73, 1.11] \\ [2.58, 3.87] \\ [0.45, 0.75] \\ [-0.48, -0.32] \\ [0.73, 1.11] \\ [-1.56, -1.03] \end{array} \right] \left. \begin{array}{l} \vphantom{\left[ \begin{array}{c} \end{array} \right]} \\ \vphantom{\left[ \begin{array}{c} \end{array} \right]} \\ \vphantom{\left[ \begin{array}{c} \end{array} \right]} \\ \vphantom{\left[ \begin{array}{c} \end{array} \right]} \\ \vphantom{\left[ \begin{array}{c} \end{array} \right]} \\ \vphantom{\left[ \begin{array}{c} \end{array} \right]} \\ \vphantom{\left[ \begin{array}{c} \end{array} \right]} \\ \vphantom{\left[ \begin{array}{c} \end{array} \right]} \\ \vphantom{\left[ \begin{array}{c} \end{array} \right]} \\ \vphantom{\left[ \begin{array}{c} \end{array} \right]} \\ \vphantom{\left[ \begin{array}{c} \end{array} \right]} \\ \vphantom{\left[ \begin{array}{c} \end{array} \right]} \\ \vphantom{\left[ \begin{array}{c} \end{array} \right]} \\ \vphantom{\left[ \begin{array}{c} \end{array} \right]} \\ \vphantom{\left[ \begin{array}{c} \end{array} \right]} \\ \vphantom{\left[ \begin{array}{c} \end{array} \right]} \\ \vphantom{\left[ \begin{array}{c} \end{array} \right]} \\ \vphantom{\left[ \begin{array}{c} \end{array} \right]} \end{array} \right\} \begin{array}{l} \text{bending moments} \\ \\ \\ \\ \text{normal forces} \end{array}$$

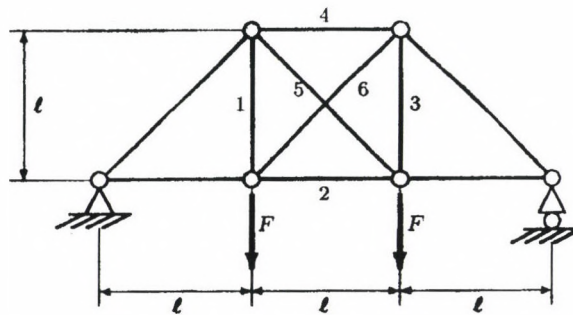


Fig. 7

#### Example 4

Also the stability problem can be studied on the structure illustrated in Fig. 7. Here the structural data are, as follows:

$$\mathbf{B} = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 1 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{R} = \begin{bmatrix} \frac{1}{30} & & & & & & \\ & 1 & & & & & \\ & & \frac{1}{30} & & & & \\ & & & 1 & & & \\ & & & & \sqrt{2} & & \\ & & & & & \sqrt{2} & \\ & & & & & & 0 \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}.$$

The elements in the diagonal of the flexibility matrix mean at the same time the nominal values of the flexibility matrix parameters. Thus the nominal solution of the equation (6) is

$$x^* = \frac{\sqrt{2}}{31 + 30\sqrt{2}},$$

and the vector of the corresponding final internal forces is

$$\mathbf{L} = \left[ \frac{30 + 60\sqrt{2}}{31 + 60\sqrt{2}}, \frac{30 + 60\sqrt{2}}{31 + 60\sqrt{2}}, \frac{30 + 60\sqrt{2}}{31 + 60\sqrt{2}}, -\frac{32 + 60\sqrt{2}}{31 + 60\sqrt{2}}, \frac{\sqrt{2}}{31 + 60\sqrt{2}}, \frac{\sqrt{2}}{31 + 60\sqrt{2}} \right].$$

Should the flexibility matrix be treated parametrically, the parametric equation to be solved will be

$$\left(\frac{1}{2}r_1 + \frac{1}{2}r_2 + \frac{1}{2}r_3 + \frac{1}{2}r_4 + r_5 + r_6\right)x = \left(-\frac{1}{\sqrt{2}}r_1 - \frac{1}{\sqrt{2}}r_2 - \frac{1}{\sqrt{2}}r_3 - \frac{1}{\sqrt{2}}r_4\right).$$

Let the fuzzy parameters of the flexibility matrix be

$$(r_i, \varepsilon_i), \quad i = 1, \dots, 5.$$

Then the fuzzy equation corresponding to (6) will be

$$\left(\frac{31}{30} + 2\sqrt{2}, \sqrt{2}\left(1 + \frac{1}{30}\right)\varepsilon\right)x + \left(-\frac{\sqrt{2}}{30}, \frac{31\sqrt{2}}{30}\varepsilon\right) = (0, 0).$$

The possibilistic fuzzy solution set of this equation is defined by the membership function

$$\sigma(x) = \begin{cases} 1 - \frac{\left| (31 + 60\sqrt{2})x - \sqrt{2} \right|}{\varepsilon \left( (31 + 60\sqrt{2})|x| + 31\sqrt{2} \right)}, & \text{if } x \in C \\ 0 & \text{otherwise} \end{cases}$$

where

$$C = \left\{ x \in \mathbf{R} : \frac{1 - 31\varepsilon}{1\varepsilon} x^* \leq x \leq \frac{1 + 31\varepsilon}{1 + \varepsilon} x^* \right\}$$

is the perturbation interval of the solution. Obvious, if

$$\frac{1}{31} \leq \varepsilon \leq 1,$$

then the point 0 will be contained in the perturbation interval, which means that as a result of the imperfections, the force may act in the opposite direction and this may upset the stability of the system. Similarly, a reversal may take place in the last two coordinates of the final stress for which the following interval vector is obtained:



$$\mathbf{Bx} + \mathbf{a} \in \left[ \begin{array}{l} \left[ \frac{31 + 60\sqrt{2}(1 - \varepsilon)}{(1 - \varepsilon)(31 + 60\sqrt{2})}, \frac{31 + 60\sqrt{2}(1 + \varepsilon)}{(1 + \varepsilon)(31 + 60\sqrt{2})} \right] \\ \left[ \frac{31 + 60\sqrt{2}(1 - \varepsilon)}{(1 - \varepsilon)(31 + 60\sqrt{2})}, \frac{31 + 60\sqrt{2}(1 + \varepsilon)}{(1 + \varepsilon)(31 + 60\sqrt{2})} \right] \\ \left[ \frac{31 + 60\sqrt{2}(1 - \varepsilon)}{(1 - \varepsilon)(31 + 60\sqrt{2})}, \frac{31 + 60\sqrt{2}(1 + \varepsilon)}{(1 + \varepsilon)(31 + 60\sqrt{2})} \right] \\ \left[ \frac{31 + 60\sqrt{2}(1 - \varepsilon)}{(1 - \varepsilon)(31 + 60\sqrt{2})}, \frac{31 + 60\sqrt{2}(1 + \varepsilon)}{(1 + \varepsilon)(31 + 60\sqrt{2})} \right] \\ \left[ \frac{\sqrt{2}(1 - 31\varepsilon)}{(1 - \varepsilon)(31 + 60\sqrt{2})}, \frac{\sqrt{2}(1 + 31\varepsilon)}{(1 + \varepsilon)(31 + 60\sqrt{2})} \right] \\ \left[ \frac{\sqrt{2}(1 - 31\varepsilon)}{(1 - \varepsilon)(31 + 60\sqrt{2})}, \frac{\sqrt{2}(1 + 31\varepsilon)}{(1 + \varepsilon)(31 + 60\sqrt{2})} \right] \end{array} \right]$$

### References

1. Michelberger, P.: Berechnung der durch die Fertigungsungenauigkeiten des Fahrgestells hervorgerufenen Montagespannungen mit Hilfe des Matrizen-Kraftgrößen-Verfahrens. *Acta Technica Hung.*, 73 (3–4): 335–347, 1972
2. Michelberger, P.: Auswirkung der Fertigungs- und Montierungsungenauigkeiten auf die Beanspruchung der Stabkonstruktionen. *Acta Technica Hung.*, 75 (1–4): 301–308, 1973
3. Michelberger, P.–Futó, P.–Keresztes, A.: Analysis for stresses caused in vehicles with the aid of statistical methods. *Acta Technica Hung.*, 83 (1–2): 93–101, 1976
4. Michelberger, P.–Szilágyi, L.: Ein Näherungsverfahren zur Festigkeitsberechnung von dünnwandigen, elliptischen Rohren. *VDI-Zeitschrift*, 102 (26): 1235–1238, 1960
5. Nándori, E.: Deduction de la modification à partir du problème général de liaison des structures statiquement indéterminées. *Periodica Politechnica*, 3 (1): 52–56, 1975
6. Michelberger, P.–Nándori, E.–Kovács, M.: The influence of some nonlinearities for design of vehicle structures. *Periodica Politechnica*, 12 (1–2): 51–57, 1984
7. Michelberger, P.–Szabó, Z.: Load examination of vehicle-body of reinforced cylindrical shell in case of kinematic load. In: *Proc. 15th Congress of International Council of Aeronautical Sciences (ICAS)*, London, Vol. II, 1467–1474, 1986
8. Michelberger, P.–Nándori, E.–Kovács, M.: Effect of the change of cross-sectional characteristics on the force distribution of vehicle frames. *Acta Technica Ac. Sci. Hung.*, 98 (3–4): 345–366, 1985
9. Zadeh, L. A.: Fuzzy sets. *Inf. Control*, 8: 338–353, 1965
10. Dubois, D.–Prade, H.: *Fuzzy Sets and Systems. Theory and Applications*. Academic Press, New York, 1980
11. Kovács, M.: A stable embedding of ill-posed linear systems into fuzzy systems. *Fuzzy Sets and Systems*, 45 (3) 305–312, 1992

# THE UNDERGROUND WATERPROOFING OF THE UNIVERSITY SPORT CENTRE OF BUDAPEST

Széll, M.

*Technical University of Budapest, Department of Building Construction,  
Műegyetem rkp. 3, H-1521 Budapest, Hungary*

(Received: 23 July 1997)

The article focuses on the waterproofing of the underground structural elements of the University Sports Centre which was designed to serve the purposes of the EXPO as well. The construction, which was stopped in 1997, will be completed in time to serve the millennium celebrations.

Due to the high groundwater level of the plot (Lágymányos) and the architectural concept that solved the contradiction between the strict building regulations and the design program by underground buildings, the engineers designing the foundations and the waterproofing faced a difficult task. The ground water level (+101.60 maB) equals the parapet height of the windows on the ground floor. If the water table risen above this level the building would have to be flooded, which would result in extraordinary stresses on the waterproofing. The curving contour of the sports hall and the effects of precipitation on the roof surfaces covered with soil, at the ramps and slot presented further difficulties. On the slurry walls of the dressing and training facilities block active waterproofing was constructed.

The design of the waterproofing, due to the difficulty of the task and the high risk factor, was a result of the work of a group of design specialists. This indicates that in the future, according to the increasing complexity of design tasks, the architect's team will have a new member: the building construction design specialist.

## 1. The design phases

The University Sport Centre – which was supposed to be used for the purposes of the EXPO first – is being constructed in the 11th district of Budapest, in Lágymányos. The project was developed in the A & D Studio that is led by the two winner architects of the competition held in 1993, Profs Antal Lázár and Péter Magyar.

The Architect Studio asked for the cooperation of subdesigners Sándor Horváth, Dr László Nagy, Rita Pataky and the author for the elaboration of the design of the waterproofing.

The article would like to describe the underground waterproofing.

The three main components of the centre are:

- the sports hall,
- the dressing and trainings rooms,
- the swimming pool,

which are connected directly and by a passage called “the slot” that is sunk into the ground but open on the top. These buildings all have their own transformer buildings that supply the electricity.

We prepared the plans of the waterproofing between May 1994 and September 1996 in several stages, continuously. The construction of the waterproofing began in August 1995.

In the summer of 1997, the task was the temporary sealing of the sports hall and the dressing and training buildings already constructed, until further development.

## 2. The factors influencing the waterproofing: the building and hydrological conditions

The group of buildings contains both hall- and cell-like spaces. The masses – due to the site conception plan – are all situated mainly underground. Only the coverings of the swimming pool and the sports hall will rise above the ground level.

Every building is different from the points of view of architectural design, structure and foundation. The sports hall of elliptical floor plan, which is covered by a reticulated dome, has a reinforced concrete raft foundation, while the spaces around it and the grandstand rest on pile foundations. The elliptic side wings have mat foundations. The engineers designed piles for the swimming pool (which is shaped similarly to the sports hall) and the adjoining spaces and grandstands, and reinforced concrete raft foundations for the side wings (Fig. 1).

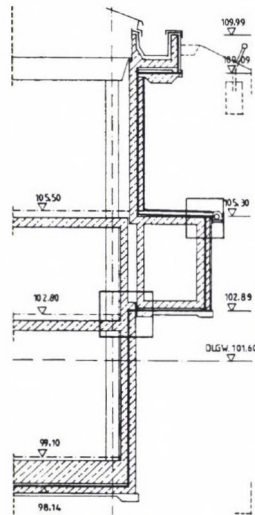


Fig. 1. Connection of air duct to the hall (vertical section)



Under the western wing of the dressing and training rooms, as a result of the modification carried out in the spring of 1995, a ca. 5500 m<sup>3</sup> three storey underground garage was constructed with a shifted schedule. In the first phase only one level was executed, the remaining two will be done by means of mining. This block is supported by slurry walls that extend into the watertight clay base material. The transformer building has a raft foundation.

Constructing the ramp and the "slot" was the most important aspect of the static design, while the waterproofing (against ground and rain water) was the most crucial constructional problem to solve.

The blocks with different foundation types are separated by expansion joints, for making the movement differences possible. According to calculations, the subsidence difference between the units with pile and raft foundations, and slurry wall and raft foundations does not exceed 20 mm, even in case of saturated soil. However, if the object has to be flooded and the water reaches the +102.00 maB level, the subsidence difference may be greater than the indicated value. The predicted movements of the building are in close relation with the changes of the hydrological conditions.

In the winter of 1993/94, when the architectural design defining the conception of the building begun, the maximum water level was +101.00 maB, equal to the existing sports ground level. The geotechnical expert's opinion prepared in the spring of 1994, however, predicted a higher possible water level. The maximum water level predicted for 100 years frequency is 103.00 maB, which, on the other hand, means a design value of +103.50 maB together with the 50 cm safety margin. The client – after considering the lifetime of the building, the cycles of renewal and the financial aspects of the different solutions – decided to have the level of protection against ground water on +101.60 maB. This corresponds to the parapet heights of the windows of the training rooms, so in case of a higher water level the building has to be flooded. The ground water is moderately aggressive.

### 3. General design principles

In the tender documentation we defined the requirements the waterproofings have to fulfill, and made a suggestion to apply a certain type of waterproofing. Protection against water pressure had to be used, in the case of the building concerned, up to the level of +101.60 maB, above which DPC was also sufficient. On the joining roof surfaces covered with soil, on the ramps and in the "slot", the effect of rain water had also to be considered.

The waterproofing has to fulfill the following requirements:

- it is waterproof up to the given pressure,
- it resists the chemical impact of the ground water,
- it is root-proof,
- can be executed also during winter,
- quick construction, easy assembly into sheets,

- minimal need for auxiliary equipments, e.g. in case of high walls,
- the buildings that are constructed one after the other are able to join each other in a waterproof way by penetrating shield or joints,
- the joints that are subject to severe loading may be solved in a safe way,
- the fine details resulting from the geometry of the waterproofing may be executed easily,
- the material is certified by “ÉMP” as usable for the given purpose,
- the suggested material was applied against pressure water on a more than 3000 m<sup>2</sup> reference surface in Hungary.

#### 4. Materials of the waterproofing

Considering the requirements listed above we suggested the use of softened, weldable plastic membrane (PVC) waterproofing, together with the necessary protective and deviding layers. The purpose is well served by the Hungarian-made HUNGISOL-B membrane. We prescribed the following material thicknesses:

- |   |             |
|---|-------------|
| – in case of pressure water             | min. 1.5 mm |
| – in case of water vapour               | min. 1.0 mm |
| – at the connecting horizontal sections | min. 1.5 mm |

“HUNGISOL FLEECE” (min 320 g/m<sup>2</sup>) was applied as a levelling layer below the isolation. The ready-made waterproofing is protected by HUNGISOL MV sheets on the horizontal and vertical surfaces.

Theoretically, other materials would have also been suitable for these conditions and place. However, we had to take into account other factors besides their individual performance data: the materials of the joined building structures, the possible reactions that may occur at the touching surfaces, and the entire system of layers. It is a well-known fact that at nonchangeable places, e.g. under raft foundations, only such a material is suitable which has a lifetime as long as the protected structure, or even longer. On the slurry wall foundations of the underground garages of the block that contains the dressing and training rooms, due to the requirements different from those of other places, we suggested active waterproofing. On these sections a plastic drain layer was applied (DÖRKEN MS). In between the head-beam and the slurry wall foundation, in order to achieve a load-transmitting connection, we designed a sheet steel waterproofing (Fig. 2).

The standard referring to the dimensioning of thermal technology – MSZ 04-140-2/1991 – deals with the energetics of a building as a unit, so the exact thicknesses of the thermal insulations are determined by the subdesigner of energetics. At special details, besides the general energetical calculations, the individual requirements also have to be satisfied in order to minimize the cold-bridge effect.



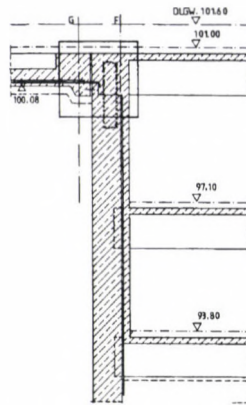


Fig. 2. The expansion joint connection of the one storey high and deep basement blocks of the dressing and training unit

## 5. Requirements concerning dryness and protection against rainfall

Inside the sports hall and the swimming pool, total dryness is required, which called for a waterproofing. In the block of the dressing and training rooms there are different prescriptions for the dryness of the different rooms:

- in case of the gyms and the connected storage rooms, total dryness is reasonable, while
- in case of the tunnel of the public works, at the ramp and in the garage, relative dryness is sufficient.

Both the ramps and the slot are open structures, subject to the effects of the rainwater and the rising groundwater, but going around with the waterproofing from above and below is unreasonable. The protection of the structures against floatation is solved according to the statical plans, the moisture protection is provided by the watertight concrete. The problems here were the making of a sufficient space for work and the expansion joints.

The draining of the rainwater is done by liftover, according to the sanitary-engineering plans.

In the transformer houses the requirement of total dryness has to be fulfilled because of the cells' sensitivity to water. The rainwater entering the circular inner courtyard might cause the corrosion of the raft foundation and the floor duct inside it. Eventually it might fill the structure, waterproofed tank-like from below, with water, so prevention necessary. We suggested that the structural concrete should slope toward the gully in the courtyard. The collected water is carried away by lifting over into a connecting pit. We applied cement mortar upgraded with synthetic resin against the rainwater. The covering consists of compressed concrete blocks laid in sand.



## 6. Waterproofing geometry and solutions

The sheet waterproofing of the buildings is often on different levels, even within one unit. The floor levels connected to the expansion joints are often different or changing. The waterproofing of the structures against pressure was made on the supporting walls, the DPC was provided with a protective structure subsequently. Between the vertical wall surfaces with DPC there are horizontal sections that may be viewed as intensive green roofs, so they have to be equipped with filtering and drain layers, and their drainage has to be solved (Fig. 3). Due to the geometry, 15–20 cm thick supporting walls had to be built up to the +102.89 maB level at some places, so the thicker, 1.5 mm PVC sheets were applied there.

Joining to the waterproof plastic sheet of the one-storey high block of dressing and training rooms, at the section surrounded by the slurry wall foundation, we applied an active waterproofing method. There is a shift in the horizontal floor waterproofing of the +100.08 maB level, after the expansion joint. As a force transmitting connection is required between the headbeam and the slurry wall foundation, which necessitates the reinforcing bars to be lead through, the waterproofing above the slurry wall foundation is made of sheet steel. The protection of the wall and the floor against moisting is executed by high capacity surface drainages, which is supplemented on the walls of the rooms on the training rooms, requiring total dryness, by further wall DPC and a lining wall built in front of it. In case of the public works' tunnel at the level of the dressing rooms and the ramps there is no need for complete dryness, therefore at these places the surface drainage and the 15 cm thick reinforced concrete lining wall in front of it pro-

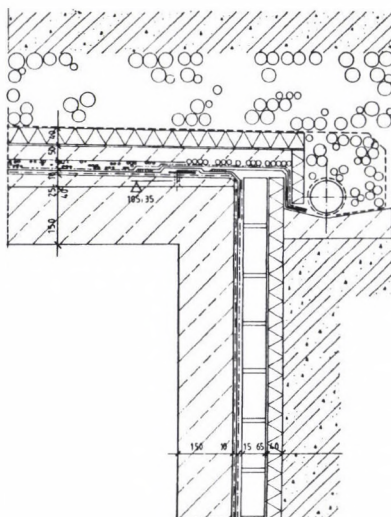


Fig. 3. Joint of wall and roof waterproofing covered with earth

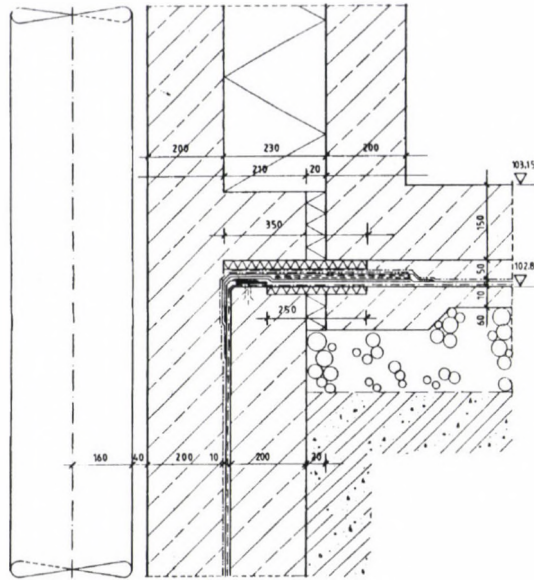


Fig. 4. Expansion joint of hall and side wing

vides sufficient protection. The plastic surface drainage runs down on the entire wall surface except for the beam pockets that serve as supports for the slabs, guiding the occasionally penetrating water below the floor. From the drain layer under the floor, the water is driven away by a total of 6 wells (2 in every section).

## 7. Expansion joints

At the expansion joints (dilatation gaps), connections of waterproofings different in material and in solution had to be provided. The predicted movement of the connected blocks relative to each other caused further difficulties. At sections with moisture proofing a thicker sheet of the DPC material was used to develop a so-called shielded dilatation (Fig. 4). At the section subjected to water pressure, the solution is a dilatation support strip (TROCAL-AF) clamped into the reinforced concrete structure, made of high capacity 5 mm thick PVC, and a pressure reduction profile (Fig. 5).

The dilatation profiles of the sheet waterproofing is driven into the mass concrete. The watertightness is solved according to the labyrinth principle: the sheets are edged by steel profiles.

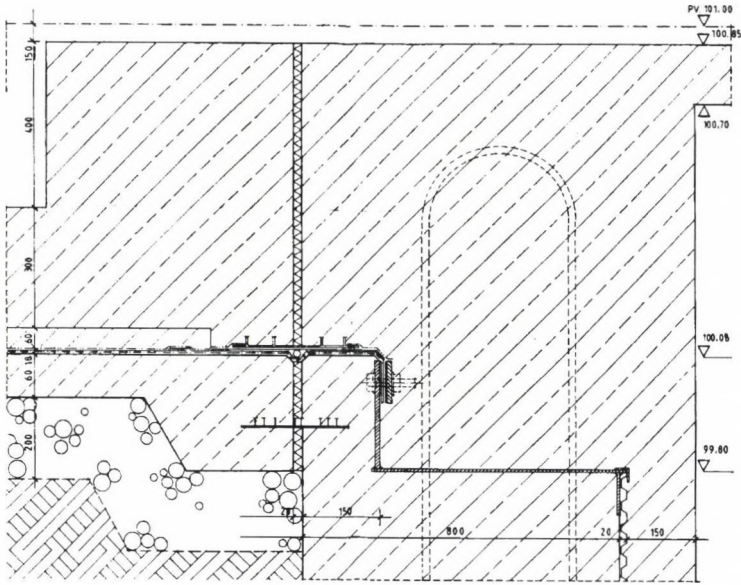


Fig. 5. Expansion joint subjected to the pressure of groundwater

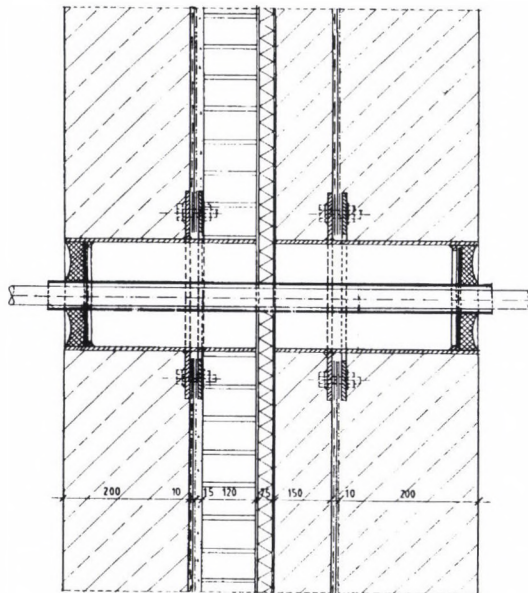


Fig. 6. Layout of electric cables between the dressing and training unit and the transformer block



### **8. Penetration through the waterproofing**

The pile foundations and the piers are connected by reinforcement in many places. The waterproofing there is sheet steel compressing edges. At the connection of PVC sheets, the compression force determined by calculation was arrived at by the prescribed tightening of the fixing nuts. The waterproofing is penetrated by sanitary-engineering ducts at different levels. They are installed with compressing edges and packing boxes at sections of waterpressure, and by collars and shields made of their own material above the +101.60 maB level, according to the requirement of moisture proofing.

In between the transformer building and the block of dressing rooms, in the zone of ground moisture, a stock of cables had to be driven through the DPC. As the penetration crosses the expansion joint, two separate sheet metal protective boxes had to be built in. Their internal sheeting and protective piping make deformations and the subsequent replacement of cables possible (Fig. 6).

### **9. The protection of the waterproofing in case of a staged construction process or flooding**

The waterproofing of the underground building parts, according to the original building concept (sports hall used for EXPO purposes first) was based on the plan that every building – the sports hall, the training rooms and the swimming pool – would be built in more or less one, continuous process. After the calling off of the EXPO project it became obvious that staging may be expected. If a longer pause elapses between the construction of units, the danger of the flooding of any existing building from the direction of a not yet built one (in case of a possible rise of the level of the ground water) must not be neglected. This must be prevented by sealing the waterproofing of the already built units, providing the possibility of adequate subsequent connections to it at the same time.

The waterproofing may be subjected to an extremely high pressure in case it has to be flooded, which becomes necessary if the water level is high. This must be avoided by all means, as long as possible. The protection of the building from the ground water, above the level of the door openings (+101.00 maB) up to the level of floatation, must be provided by temporary solutions (sandbags, etc.). After flooding, the dewatering of the building units must be carried out in such a way that the subsidence difference of the separate units should be as little as possible.

### **10. Summary**

Besides the three volumes of the tender documentation, the four volumes of working drawings, including 70 pages of text, 75 M=1:25 sections and about 110 M =1:5 details, presented the following information about the underground waterproofing of the University Sports Centre:

- system of requirements,
- isolation geometry,
- layers,
- junction details and
- material quantities.

The dimensions and complexity of the buildings, the unfavourable soil and ground water conditions and the accelerated design process ment a great challenge for every designer involved in the project, including us.

The creators of the waterproofing all teach building construction, and their design practice is of the same type, as well. Their everyday practice justifies the recognition that, as the design tasks become more and more complex, the team of designers must include a new member: the subdesigner of building construction. The architect is busy with the problems of form and function, the structural engineer is responsible for the loadbearing and stability, while the sanitary engineer is responsible for the services of the building. The new member of the team may contribute successfully in two ways:

- designing the individual constructions and substructures of the building as in the case presented in this article,
- achieving the suitable unity of solutions, materials and technologies of building construction, with adjusting the performances to the requirements. In this case the building construction specialist cooperates with the architect from the concept phase until the detail drawings continuously, helping the birth of suitable, healthy or even intelligent buildings with his/her special knowledge.



## BOOK REVIEW

Amália Iványi: **Hysteresis models in electromagnetic computation**

Akadémiai Kiadó, Budapest, 1997, ISBN 963 7416 0

In English, 229 pages, 17 × 25 cm, Hardbound

Hysteresis in ferromagnetic materials is a very interesting from both practical and theoretical point of view. Its practical importance in electrical engineering (especially in power engineering is well known in electrical machines, iron cores, shielding layers, etc. From a theoretical point of view hysteresis is a nonlinear process with memory which, as a consequence, cannot be characterised with a function but more sophisticated approaches are required. In most cases hysteresis is accompanied by anisotropy of the material, too.

The physical backgrounds of hysteresis are dealt in the literature from a physicist point of view. The computational methods to be used in electrical engineering calculations appear scattered in various books and papers. This book aims to give a survey of the different models giving a very short introduction on the physical backgrounds and a very detailed treatise of the applied calculation method. We get not only a description of the known models and their applications but a critical description of their advantages and possible shortages, too. These are based not only on the different publications but on the author's board experiences at the calculation of electromagnetic problems involving ferromagnetic materials.

The book is divided in two parts, about 180 pages net. It is accomplished with a very detailed list of references, authors and subject index.

The first part gives the critical description of the different models of hysteresis, in particular the "analytical", the "dynamic", the Langevin type, the Preisach type (a very detail description), the Stoner–Wohlfahrt type and the Chua type model. As already mentioned we get not only the descriptions of the different models but their properties with emphasis on their applications at numerical electromagnetic field calculation.

The second part covers the author's own scientific researches. They provide the applications of the most important hysteresis models dealt previously. The discussed topics are energy in oriented lamination, a magnetic shielding, anisotropic material in rotational field. We get here an introduction in different methods for the calculation of special electromagnetic problems (field quantities, scalar potential and vector potential ap-



proach). In this part we find also some comparison of the results calculated by the author and the results calculated or measured by other experts. More such comparison would be useful because it is not easy to estimate the errors caused by the necessary simplifications and neglects applied during the various calculations.

Summarising: the book is an excellent reference of its topic. It can be recommended for electrical and software engineers interested in electromagnetic field calculations involving ferromagnetic materials with hysteresis and anisotropy.

*G. Fodor*

PRINTED IN HUNGARY

Akadémiai Nyomda, Martonvásár

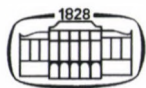
MAGYAR  
TUDOMÁNYOS AKADÉMIA  
KÖNYVTÁRA

# ACTA TECHNICA

ACADEMIAE SCIENTIARUM HUNGARICAE

EDITOR-IN-CHIEF: P. MICHELBERGER

VOLUME 107



AKADÉMIAI KIADÓ, BUDAPEST 1995-96







ACTA TECHNICA

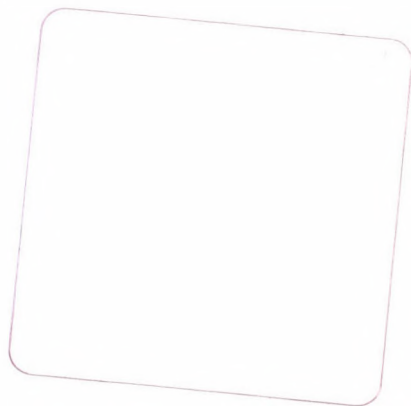
Volume 107 Nos 1-4

<i>Bagi, K.</i> : Geometrical modelling of granular assemblies .....	1
<i>Bhargava, R. R. – Tyagi, N. K.</i> : Unbounded plate with two circular-arc cracks with coalesced plastic zones closed by variable pressure .....	149
<i>Bódi, I. – Fűzy, J. – Klopka, Z.</i> : Deformation and stress analysis of hyperbolic paraboloid shells taking into account realistic boundary conditions .....	175
<i>Bojtár, I.</i> : Numerical estimation of the elastic properties of cemented granular materials with microstructure .....	17
<i>Bojtár, I. – Gálos, M. – Rechterisz, Á.</i> : Determination of fracture mechanical properties on rock's boring core material .....	163
<i>Ding, Dajun</i> : Semiempirical theory of masonry strength .....	181
<i>Dombay, G. – Piriou, Ph. – Dukan, S. – Kiene, L.</i> : Modeling the influence of environmental parameters and residence time on microbial drinking water quality in distribution systems .....	195
<i>Dulácska, E.</i> : The effect of ductility in combination with the P- $\Delta$ effects, to be taken into consideration in the seismic design of compressed structures .....	219
<i>Galaskó, Gy. – Gáspár, Zs. – Nouri-Baranger, T. – Leon, J. C. – Trompette, P. – Veron, P.</i> : Comparison of tent structures calculation in Hungary and France .....	27
<i>Gáspár, Zs. – Károlyi, Gy. – Popper, Gy.</i> : Application of Gröbner basis theory to find global equilibrium paths of the simple arc .....	37
<i>Györgyi, J.</i> : Use of finite element method for testing structures for dynamical effects caused by moving vehicles .....	49
<i>Hankó, Z.</i> : On the division of the kinetic energy (power) content in turbulent flow. An approximate analytical review .....	405
<i>Keshava, K. M. – Shesha, P. M. N.</i> : Application of a new optimization procedure for the design of a proportional weir .....	265
<i>Klopka, Z. – Ládi, N.</i> : Numerical examination of the stiffening system of tall buildings by using object-oriented programming .....	277
<i>Kozák, I.</i> : Continuum mechanics. Summary of comparison of solid and fluid mechanics .....	303
<i>Kurutz, M.</i> : Imperfection sensitivity of the classical stable-symmetric bifurcation problem modified by bilinear material .....	63
<i>Kurutz, M.</i> : Tangent modulus for nonsmooth materials .....	339
<i>Lovas, A.</i> : Mechanical models of the human skull-brain system .....	83
<i>Nándori, E. – Kovács, M.</i> : Possibilistic estimations of imperfections in structural analysis .....	369
<i>Nédli, P.</i> : Elastic-plastic analysis of frames in case of geometrical nonlinearity .....	99
<i>Popper, Gy. – Károlyi, Gy.</i> : Gröbner basis calculations with applications to mechanics .....	109
<i>Szabó, J. – Tarnai, T.</i> : On the quasi-inverse of the matrix of a system of linear equations .....	115
<i>Széll, M.</i> : The underground waterproofing of the University Sport Centre of Budapest .....	395

<i>Tarnai, T. – Gáspár, Zs.:</i> Packing of equal circles in a square .....	123
<i>Vásárhelyi, A. – Lógó, J.:</i> Dynamic elasto-plastic analysis of structures by mathematical programming .....	137

BOOK REVIEW

<i>Iványi, A.:</i> Hysteresis Models in Electromagnetic Computation ( <i>G. Fodor</i> ) .....	405
---	-----



## NOTICE TO CONTRIBUTORS

Papers in English are accepted on condition that they have not been previously published or accepted for publication.

Only original papers will be published and a copy of the Publishing Agreement will be sent to the authors of papers accepted for publication. Manuscripts will be processed only after receiving the signed copy of the agreement.

Manuscripts in two copies (the original type-written copy plus a clear duplicate one) complete with figures, tables, and references should be sent to

*Acta Technica*  
Nádor u. 7. I. 118  
Budapest, Hungary  
H-1051

Although every effort will be made to guard against loss, it is advised that authors retain copies of all material which they submit. The editorial board reserves the right to make editorial changes.

*Manuscripts* should be typed double-spaced on one side of good quality paper with proper margins and bear the title of the paper and the name(s) of the author(s). The full postal address(es) of the author(s) should be given in a footnote on the first page. An abstract of 50 to 100 words should precede the text of the paper. The approximate locations of the tables and figures should be indicated in the margin. An additional copy of the abstract is needed. Russian words and names should be transliterated into English.

*References.* Only papers closely related to the author's work should be referred to. The citations should include the name of the author and/or the reference number in brackets. A list of numbered references should follow the end of the manuscript.

References to periodicals should mention: (1) name(s) and initial(s) of the author(s); (2) title of the paper; (3) name of the periodical; (4) volume; (5) year of publication in parentheses; (6) numbers of the first and last pages. Thus: 5. Winokur, A.—Gluck, J.: Ultimate strength analysis of coupled shear walls. American Concrete Institute Journal 65 (1968) 1029-1035

References to books should include: (1) author(s)' name; (2) title; (3) publisher; (4) place and year of publication. Thus: Timoshenko, S.—Gere, J.: Theory of Elastic Stability. McGraw-Hill Company. New York, London 1961

*Illustrations* should be selected carefully and only up to the necessary quantity. Black-and-white photographs should be in the form of glossy prints. The author's name and the title of the paper together with the serial number of the figure should be written on the back of each print. Legends should be brief and attached on a separate sheet. Tables, each bearing a title, should be self-explanatory and numbered consecutively.

Authors will receive proofs which must be sent back by return mail.

Authors will receive 50 reprints free of charge.



