TAMPEREEN TEKNILLINEN KORKEAKOULU
Teknillinen mekaniikka

VII SUOMEN MEKANIIKKAPÄIVÄT

Tampereella 25-26. 5. 2000

NIDE 1

1. kokouspäivän esitelmät

Toimittajat
Juhani Koski ja *Simo Virtanen*

Tampere 2000

TAMPEREEN TEKNILLINEN KORKEAKOULU
Teknillinen mekaniikka
PL 589
33101 TAMPERE

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

ISBN 952-15-0414-5

Klingendahl Paino Oy
Tampere, 2000

ESIPUHE

Suomen VII Mekaniikkapäivät pidettiin Tampereella 25-26.5.2000 ja siellä pidetyt esitelmät on koottu kahteen niteeseen. Nide 1 sisältää ensimmäisen ja nide 2 toisen kokouspäivän esitelmät. Kutsuttuina ulkomaisina esitelmöitsijöinä olivat professori *Wolfram Stadler* (San Francisco State University, USA) sekä professori *Niels Olhoff* (Aalborg University, Tanska). Lisäksi tilaisuutta kunnioittivat esitelmillään professorit *Ülo Lepik* ja *Jaan Lellep* Tarton yliopistosta. Kotimaisina kutsuttuina esitelmöitsijöinä olivat professori *Martti Mikkola* (Teknillinen korkeakoulu) ja professori *Antti Pramila* (Oulun yliopisto). Niteet sisältävät yhteensä 58 artikkelia, jotka edustavat monipuolisesti teknillisen mekaniikan eri osa-alueita.

Suomen mekaniikkapäivät on järjestetty vuodesta 1982 lähtien joka kolmas vuosi (Oulu 1982, Tampere 1985, Otaniemi 1988, Lappeenranta 1991, Jyväskylä 1994, Oulu 1997). Niiden tarkoituksena on koota yhteen tutkijoita, suunnittelijoita ja opettajia, jotka työssään joutuvat tekemisiin mekaniikan ja lujuusopin ongelmien kanssa. Uutena aiheena tällä kerralla on mukaan otettu teknillisen mekaniikan opetus, jolle on varattu oma istunto.

Järjestelytoimikunta esittää lämpimät kiitoksensa kaikille esitelmijöitsijöille sekä artikkelien tekijöille, kuten myös niille lukuisille henkilöille, jotka ovat osallistuneet mekaniikkapäivien järjestelytyöhön.

Juhani Koski ja Simo Virtanen

SISÄLLYSLUETTELO

Hitsiliitokset

Elementtimenetelmä 1

Optimointi 2

Värähtelyt

Väsyminen ja lämmönsiirto

Jäännösjännitykset ja elementtimenetelmä

NIDE 2

Kutsutut esitelmät

Rakenteiden mekaniikka

Palkit

Maamekaniikka

Erityisaiheita

Elementtimenetelmä 2

A BRIEF HISTORY OF MULTICRITERIA OPTIMIZATION AND ITS APPLICATIONS IN MECHANICS

W. Stadler
School of Engineering
San Francisco State University
1600 Holloway Avenue
San Francisco, CA 94132, USA

ABSTRACT

We provide a brief history of multicriteria optimization, including applications in Mechanics. We begin with the origins of optimization, provide a basic optimization problem formulation, discuss some of the initiators of multicriteria optimization, show that a classical principle in Mechanics has a multicriteria formulation and trace the evolution of an optimization problem from its beginnings in antiquity to wide present application.

1. INTRODUCTION

Mechanics and Engineering Science may be defined as areas of instruction whose purpose it is, at least in part, to bring new methods and concepts in the sciences and mathematics to fruition by applying them in the solution of problems. Unfortunately, Mechancis has been slow to adapt and continues to maintain a traditional offering rather than innovating with courses in controls, mechatronics, multidisciplinary analysis, fuzzy systems and optimization, to name only a few. In particular, optimization as a broad design tool, has been, for the most part, ignored in the US.

In every book treating engineering design, the student is admonished to optimize his designs with vague guidelines such as, 'as strong as possible', 'as light as possible ', 'as flexible as possible', and so on, with little or no guidelines as to just how this is to be accomplished. In part, this is due to the fact that the concept of design itself is separated into two relatively disjoint approaches: One almost artful approach based on the refinement of anything that works and satisfies the constraints, making virtually no use of analysis, and a second approach which relies on detailed mathematical modeling, analysis and concepts from decision making and optimization. When one has well-accepted and tried mathematical models and computational methods for the area in which design is to be carried out, it makes no sense not to use them to obtain optimal designs. Both approaches

go back to antiquity. Not surprisingly, we know more about the origin of the analytical approach.

As far as it was possible to determine, Dido's problem is the first example of the practical use of an optimality statement. We shall first tell the story as it is told now and then indicate how it was pieced together from the historical fragments of Greek and Roman historians.

Dido's Problem. Mutto, the king of Tyre had two children, a son Pygmalion and a daughter Elissa. Upon Mutto's death, Pygmalion became king and Elissa married her uncle Sicharbas, the priest of Heracles. Pygmalion had Sicharbas assassinated to seize the priest's treasure. Elissa, appalled at her brother's action, had the treasure secretly loaded onto boats and fled with some of the nobility partial to her plight. After a stop in Cyprus, they continued westward along the northern coast of Africa, where they were welcomed by the local inhabitants. They called Elissa, Deido (wanderer), because of her many wanderings. She encountered a local ruler, King Iarbas of Numidia, whom she asked for some land where she and her retinue could settle. He allowed her to take as much as she could enclose in the hide of an ox. Dido had the hide cut into strips and then used these strips to enclose a sizable piece of land. The inhabitants kept their promise and gave her the land on which Dido eventually founded Carthage. Some time thereafter, the Lybian king wished to marry her, but she refused. On the pretext of carrying out a rite to release her from an oath not to marry, she constructed a large pyre near her chambers from which she threw herself on the pyre after it had been lighted.

This story is an amalgamation of at least three different references to Greek mythology and history.

The earliest mention of the story dealing with Elissa and her escape from Tyre is due to Timaeus (of Taromínium a Greek colony in Italy), who was exiled to Athens by Agathocles, the local tyrant. Timaeus lived from about 356 B.C. to 260 B.C. and he is considered to be the foremost third century historian whose fragments have survived in sufficient bulk to make a good assessment of his contribution. He also seems to be the originator of the myth that both Carthage and Rome were founded in the 38th year of the Olympiad or about 814-813 B.C., the generally accepted date for the founding of Carthage.

The part of the story dealing with "as much land as she might enclose with an ox hide" was penned by Publius Virgilius Marro (70-19 B.C.) in his epic tying the founding of Rome into Greek mythology at the request of Augustus, who commissioned him. In this story, she commits suicide because Aeneas leaves her. If nothing else, we now have to resolve the difficult undertaking of having committed suicide twice. One resolution is to assign the name Dido to all of the Punic queens. That is, they take on the name Dido upon becoming Queen of Carthage.

Justinius, another Roman historian, provided the additional detail of cutting the hide into strips in his book Historiae Phillippicae written in the second or third century A.D.

The attendant geography is somewhat more deterministic. Tyre is a city in today's Lebanon and Carthage (or rather what remains of it) is located on a peninsula (an easily defended location) in Libya, on the Bay of Tunis, just northeast of modern Tunis.

The earliest archeological evidence of Carthage dates from the eighth century B.C., so that the city thrived for more than 500 years until its recorded destruction by Rome at the end of the Third Punic War in 146 B.C. This is slightly at odds with Timaeus placement of the founding of Carthage in the 9th century B.C. Virgil places the mythiological founding of Rome by Aeneas, a refugee from Troy, at about 1100 B.C.

According to Ref. 2, Zenodorus provided the first proof of the isoperimetric property of the circle sometime between 340-212 B.C. The proof apparently contained a gap, which was not removed until the 19th century by Karl Weierstrass (1815-1897) in his lectures at the University of Berlin.

We may assume that Dido was aware of the isoperimetric property of the circle. Again, according to Ref. 2, cutting the hide into 1/10th inch strips would yield about 1000-2000 yards of strips with which she could have enclosed about 16-65 acres of land for a complete circle and considerably more, if taken as a semicircle along the ocean. Since Carthage was located on a peninsula, she must have done well indeed.

Figure 1. Dido purchases land for the foundation of Carthage. Engraving by Matthaus Merian the Elder, in Historische Chronika, Frankfurt a.M., Germany, 1630. (Ref. 2).

Dido's problem is and will probably remain the most celebrated optimal design problem. Thanks to its mythological setting, it can be celebrated on several levels as shown in Figures 1 and 2. Figure 3 shows that the civil engineers of the middle ages were quite aware of this result in optimal design and put it to good use.

Figure 2. J.M.W. Turner. Dido building Carthage.

Thus was born the isoperimetric problem: Among all closed plane curves of a given length to find the one that encloses the largest area.

Virtually all of the early optimization problems concerned maximization or minimization in geometric problems; there were no common methods and each problem thus was individually dealt with. However, what was found, was proven to be the optimum and thus was accepted as fact. Thus, Aristotle in the 4^{th} century B.C. already took it for granted that the circle enclosed the largest area for a given perimeter with a similar conclusion for the sphere as enclosing the largest volume for a given surface area.

The first optimization problem involving a physical phenomenon appears to have been Heron of Alexandria's probes concerning the reflection of light rays in a mirror posed in his work on *Catoptics* (reflection) in 75 A.D. (According to Ref. 3, the law of reflection had been known to Euclid, Aristotle and, probably, to Plato):

Determine the location of the point D so that the sum of the distances AD + BD is a minimum (see Figure 4).

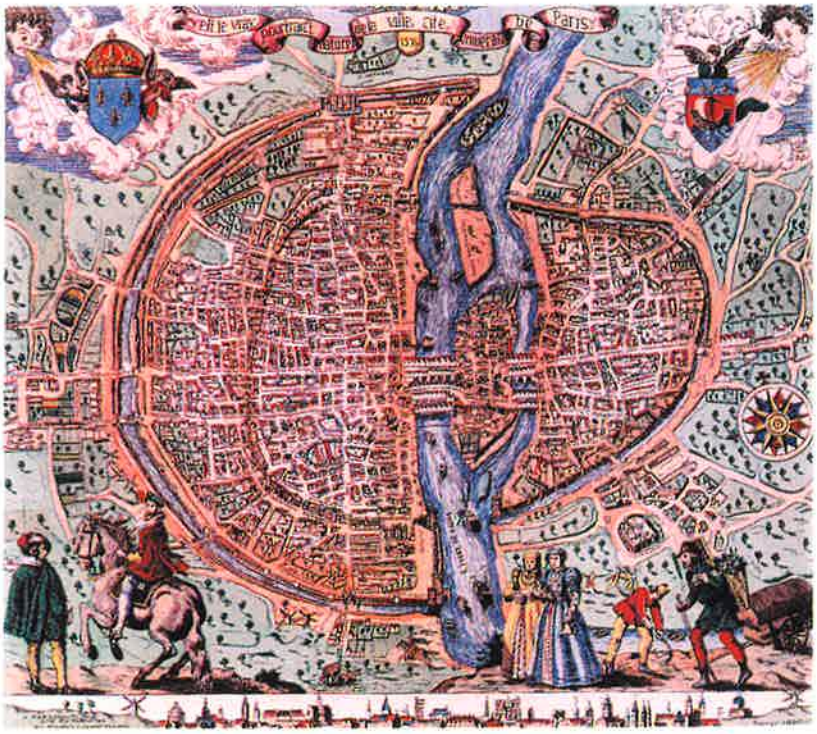


Figure 3. Medieval map of Paris.

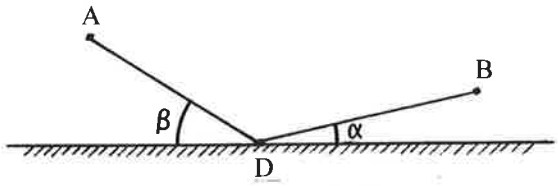


Figure 4. Heron's Problem

Eventually, this problem gave rise to the investigation of the law of refraction by Snell (1591-1626) from an experimental viewpoint and by Pierre de Fermat (1608-1665) from a theoretical point of view, concluding that light travels the minimum time path. As we shall see, this problem is the beginning of a formulation with a wide range of applications.

In reading the scientific literature from the seventeenth to the nineteenth century, we get the clear implication that nature "operates" optimally in optics, in mechanics, in thermodynamics, in fact, everywhere, much in line with the Aristotelian principle that "nature does nothing the hard way". Courant and Robbins, in their classic book *What is Mathematics?* summarize (Ref. 4): "It was observed long ago that natural phenomena often follow some pattern of maxima and minima." Leanhard Euler puts these sentiments into a metaphysical context, in line with his religious beliefs, when he writes (Ref. 5):

"...For since the fabric of the universe is most perfect and is the work of a most wise Creator, nothing whatsoever takes place in the universe in which some relation of maximum and minimum does not appear. Wherefore there is absolutely no doubt that every effect in the universe can be explained as satisfactorily from final causes, by the aid of the method of maxima and minima, as it can from the effective causes themselves."

This postulated optimality of the universe may seem to derive from religious belief but those who held with this postulate ranged widely from the very religious Euler to those who arrived at the conclusion from their own reasoning and observation. The fact is that the extremum and minimum principles that we now so routinely employ, derive from such deeply held convictions about the efficiency of natural law. We list only a few such statements here:

Fermat's Principle of Geometric Optics. In an inhomogeneous medium a light ray traveling between two points follows a path along which the time taken is a minimum with respect to all paths joining the two points.

Principle of Minimum Potential Energy. An equilibrium of a mechanical system is stable if and only if the potential energy is a minimum.

Hamilton's Principle. The Irish mathematician W.R. Hamilton (1805-1865) proposed the following generalization of the principle of virtual work. Consider a system of n particles, assume that the external forces are conservative and the possible motions holonomic. Let $T(t)$ be the instantaneous kinetic energy of the system and let $V(t)$ be the potential energy of the external forces (we suppress the dependence on $\mathbf{x}$ and $\dot{\mathbf{x}}$). A motion $\mathbf{x}(t)$ of the system is possible if and only if

$$\delta \int_{t_0}^{t_1} (T(t) - V(t))dt = 0.$$

We have cited Fermat's Principle to show what could be deduced from the simple problem that Heron posed. We shall show subsequently that the problem has continued to evolve far beyond its initial statement to the point where it now has evident applications in electrical, civil and mechanical engineering. The principle of potential energy is one of the oldest minimum principles and we have stated Hamilton's Principle as a classical principle which can be given a multicriteria formulation.

From statements such as "nature does nothing the hard way", it is easy to infer that any possible way of doing something is either the easiest way or we may find a way which is easier. If there is only one way, it becomes easiest by default. The inference from the previously stated "natural " principles is that nature does it the easiest way; that is, nature achieves an optimum. Collectively, we thus arrive at the common problem formulation:

A quantity is to be maximized or minimized by making a suitable choice from some set of allowed choices. More concisely: Minimize $g(\mathbf{d})$ subject to $\mathbf{d} \in D$.

Note that this statement implicitly includes the possibility of comparing two different designs $\mathbf{d}_1, \mathbf{d}_2 \in D$ with

$$\mathbf{d}_1 \precsim \mathbf{d}_2 \quad \text{iff} \quad g(\mathbf{d}_1) \leq g(\mathbf{d}_2)$$

2. THE MULTICRITERIA PROBLEM

The vector optimization problem or multicriteria problem may also be succinctly stated in the form: "Optimize" $\mathbf{g}(\mathbf{d})$ subject to $\mathbf{d} \in D$, where the criterion $\mathbf{g}(\mathbf{d})$ now is given by the criterion vector

$$\mathbf{g}(\mathbf{d}) = (g_1(\mathbf{d}), g_2(\mathbf{d}), \ldots, g_N(\mathbf{d}))$$

Since we are no longer simply maximizing or minimizing a single criterion function, we must give meaning to the optimization of a vector.

In order to gain a good understanding of the multicriteria problem, it is worthwhile to dissect the whole idea of an optimization problem into its basic issues and concepts. In virtually all optimal design endeavors, it is the optimal design which is of interest and not the optimal value of some criterion function. The optimal design problem is best viewed in this light:

(1) There exists a set D of possible designs. A design $\mathbf{d} \in D$ may be characterized in a variety of ways, by a single variable or function, vectors of these or by a mix thereof. We shall present a design simply by $\mathbf{d}$.
(2) There must be a way of comparing at least two designs $\mathbf{d}_1$ and $\mathbf{d}_2$ in D. This is usually accomplished by introducing a relation $\precsim$ on D such that $\mathbf{d}_1 \precsim \mathbf{d}_2$ ($\mathbf{d}_1$ is preferred to ($\prec$) or equivalent to ($\sim$) $\mathbf{d}_2$). For obvious reasons, we call such a relation a preference relation. It may be that not all designs may be comparable to each other.
(3) There must be some notion of what constitutes a best design. Once one has achieved a comparison of different designs, it is natural to attempt to isolate a design $\mathbf{d}^* \in D$ which is better than at least a collection of other designs. Usually, this is done within the context of the preference relation $\precsim$; e.g. $\mathbf{d}^* \precsim \mathbf{d}$ for all $\mathbf{d} \in D_0 \subset D$.

Example 2.1. Suppose our design set is $D = \{x \in \mathbb{R}^1: \ x \in [0,2]\}$. The usual preference relation then is the complete or linear order $\leq$ on $\mathbb{R}^1$ restricted to the interval D. A minimum with respect to $\leq$ is given by $x^* = 0$ with $0 \leq x$ for every $x \in D$.

Rather than ordering the design set itself, suppose we have a mapping of the design set D into the reals, $\phi(\cdot)$: $D \to \mathbb{R}^1$. We may then use the usual order $\leq$ on $\mathbb{R}^1$ to induce a preference $\precsim$ on the set D with

$$\mathbf{d}_1 \precsim \mathbf{d}_2 \quad \text{iff} \quad \phi(\mathbf{d}_1) \leq \phi(\mathbf{d}_2)$$

with the obvious correlation for the optimal design. This is, of course, the concept of a utility function in economics.

Example 2.2. Let $D = \{x \in \mathbb{R}^1: x \in [0,4]\}$ and consider the mapping $g(\cdot)$: $D \to \mathbb{R}^1$ given by

$$g(x) = 1 + (x-2)^2$$

The use of the previously described approach produces

$$x_1 \precsim x_2 \text{ iff } g(x_1) \leq (x_2)$$

as illustrated in Figure 5.

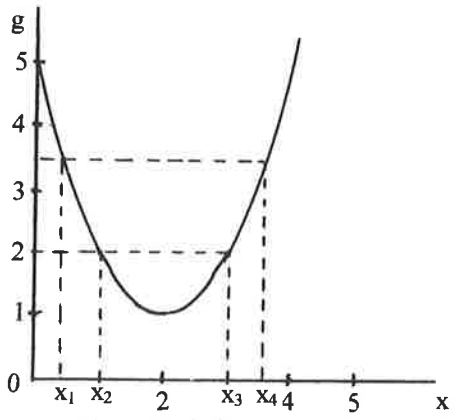


Figure 5. The imposed preference

Note that $x_1 \sim x_4$ and $x_2 \sim x_3$ and that $x_1 \prec x_2$ as well as $x_2 \prec x_4$, etc. With minimization as the basic objective, the optimal design is given by $x^* = 2$ with $x^* \precsim x$ for every $x \in D$.

We emphasize again that the criterion values themselves are generally of lesser importance than their use in comparing different designs in design space.

The same overall approach is used in the multicriteria problem. The only difference is that the mapping now is a vector-valued map

$$\mathbf{g}(\cdot): D \to \mathbb{R}^N$$

with $\mathbf{g}(\mathbf{d}) = (g_1(\mathbf{d}), g_2(\mathbf{d}), \ldots, g_N(\mathbf{d}))$ as the criterion vector. A preference $\precsim_N$ is introduced on $\mathbb{R}^N$ and is then used to induce a preference $\precsim_D$ on D with

$$\mathbf{d}_1 \precsim_D \mathbf{d}_2 \text{ iff } \mathbf{g}(\mathbf{d}_1) \precsim_N \mathbf{g}(\mathbf{d}_2)$$

We shall trace the historical development of this problem along with its most prevalent solution concept in the next section.

3. INITIATORS OF MULTICRITERIA OPTIMIZATION

The relatively late appearance of the consideration of several criteria stems from the fact that it is much harder to imagine an optimal choice, given the engrained view of the single criterion problem where the ideas of maximum and minimum evolved together with the usual number system and its ordering. Having grasped this concept, basic computation generally is no more difficult than that for the single criterion case.

The development of economic theory provided the impetus for the multicriteria problem. Such problems were first considered in economic utility theory and welfare theory. The subsequent inception of game theory provided further insight and the more or less simultaneous evolution of the mathematics of orderings on sets eventually put all of the concepts on solid mathematical footing.

Formally, we take the beginning of economic theory to be the *Wealth of Nations* written by Adam Smith, and his contemporary view of society expressed in: "A vile maxim of the masters of mankind ... All for ourselves and nothing for other people." One of the early ideas was that of a utility function $g(\cdot)$: $D \to \mathbb{R}^1$ for *each* consumer, indicating which goods or choices he preferred and it was assumed that individual choices were made in harmony with his utility function; i.e.,

$$\mathbf{d}_1 \precsim \mathbf{d}_2 \text{ iff } g(\mathbf{d}_1) \le g(\mathbf{d}_2)$$

The collective good or simultaneous consideration of the previously individually active consumers is the basic idea of welfare theory and the first realization of the vector maximum problem. That is, we wish to obtain an optimal decision $\mathbf{d}^* \in D$ which optimizes all of the criteria simultaneously or, equivalently, provides for the optimum of the vector $\mathbf{g}(\mathbf{d})$.

Francis Ysidro Edgeworth (1845 – 1926). He was the first to consider individuals with competing preferences, more precisely, competing utilities. He generally visualized a situation between two individuals and then extrapolated the results to groups of individuals in his book *Mathematical Psychics* (Ref. 6). The two person situation is often illustrated with the "Edgeworth Box", an illustration first used by the statistician Arthur Lyon Bowley (1869 – 1957).

Suppose we consider two economic agents A_1 and A_2 with convex utilities $g_1(x,y)$ and $g_2(x,y)$ which they seek to maximize by trading the commodities (x,y) with a total endowment (a,b); that is, $0 \le x \le a$, $0 \le y \le b$. Their consumption set thus is the quadrant bounded by (a,b). The quadrant is arranged in such a way that A_1 reaches his total endowment at (a,b) and A_2 reaches it at (0,0). In terms of the distribution of goods at a point (x,y), A_1 receives the amount (x,y) while A_2 gets (a – x, b – y). The situation is illustrated in Figure 6.

In the sketch, the level curves of A_1 are the solid arcs emanating from $g_1(0,0)$ and those of agent A_2 are dashed and emanate at $g_2(0,0)$ taken to coincide with (a,b). The arcs are the

lines of indifference for each agent, indicating that the agent is indifferent for all allocations of goods associated with a given indifference set. Edgeworth then defines an equilibrium point for the economy (defines optimality) in the following manner:

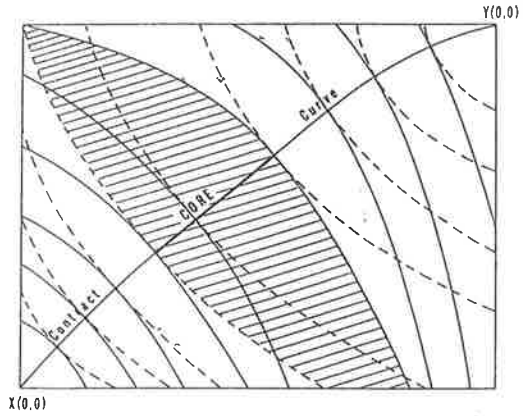


Figure 6. The Edgeworth Box and the Core (of a cooperative game)

"It is required to find a point (x*,y*) such that in *whatever direction* we take an infinitely small step, $g_1(\cdot)$ and $g_2(\cdot)$ do not increase together but that while one increases the other decreases."

With this in mind, he arrives at what he terms the contract curve traced out by points at which the gradients of $g_1(x,y)$ and $g_2(x,y)$ have the same direction. Furthermore, if we take a small step away from a point on the contract curve, e.g., we move into the upper cusped region, then $\partial g_1/\partial x < 0$ and $\partial g_2/\partial x > 0$ as required by the definition. He discusses many relevant implications and extensions of the basic definition, most of which are mentioned in Ref. 6.

Vilfredo Federico Damaso Pareto (1848 – 1923). His parents were from Genoa and had moved into exile in France for political reasons. When Vilfredo was born in 1848, they first named him Fritz Wilfried, perhaps because of some admiration for the German democrats, and only later did they turn this into Vilfredo.

Vilfredo was trained as and initially worked as a civil engineer. He held strong political views and his socialist leanings eventually made him one of the founders of fascism. Anecdotal lore has it that Mussolini attended some of his lectures in Political Economy at Lausanne, where Pareto had accepted Leon Walrus' position.

Pareto was a Greek and Latin scholar who developed an abiding interest in economics through his close friendship with the economist Maffeo Pantaleoni. Vilfredo's early training in mathematics quickly placed him at the forefront of mathematical economics where he was the first to use variational methods in deducing necessary conditions for economic equilibrium. He was well aware of Edgeworth's work and he particularly

prided himself in making use only of the level curves in his analysis rather than assuming the existence of utility functions. His somewhat rambling definition of the "maximum ophelimity" of a collectivity (Ref. 7) and his associated variational necessary conditions have evolved somewhat in what we now take to be the definition of a Pareto optimum. Indeed this optimum should more appropriately be termed an Edgeworth-Pareto optimum or EP-optimum in a more abbreviated form. We write the Edgeworth - Pareto definition in the now accepted mathematical form:

Definition. *Edgeworth-Pareto Optimality.* Let $\mathbf{d}^* \in D$. Then $\mathbf{d}^*$ is an EP-optimal decision iff

$$\mathbf{g}(\mathbf{d}) \leq \mathbf{g}(\mathbf{d}^*) \Rightarrow \mathbf{g}(\mathbf{d}) = \mathbf{g}(\mathbf{d}^*)$$

for all $\mathbf{d}^*$-comparable $\mathbf{d} \in D$.

Here $\leq$ is the so-called *natural* order on $\mathbb{R}^N$; the $\mathbf{d}^*$-comparable requirement is added to the definition because $\leq$ is a partial order, implying that not *all* of the criteria values may be comparable to each other. Minimization of the individual criteria was taken to be the basic desideratum in this definition.

Even in this succinct mathematical form, it isn't all that clear what is meant by an EP-optimum. Some more or less converse statements do provide some insight.

Perhaps the most obvious and useful from a computational point of view is that a criterion value $\mathbf{g}(\mathbf{d})$ cannot be EP-optimal if there exists a $\mathbf{g}_0 \in A = \mathbf{g}(D)$ such that $g_{0i} < g_i(\mathbf{d})$ for i = 1,2,...,N, when minimization of each $g_i(\cdot)$ is taken as the basic objective. Similarly, we are at an EP-optimal point if we cannot deviate from it, even minutely, without increasing at least one of the criteria. The EP-optimal set of designs is that set where we truly need to compromise in moving from one design to another in that a further decrease in one component would require an increase in at least one of the other components. To some extent the value of multicriteria optimization thus lies in all of the bad designs that are eliminated rather than in the particular attributes of the EP-optimal designs themselves.

Georg Cantor (1845 – 1918). The concepts of Edgeworth and Pareto grew out of the socioeconomic viewpoint of what constituted an economic equilibrium; in essence, the search for a point at which the consumers and the producers would be satisfied with what they got. Indeed, it was shown later that subject to some convexity assumptions on preferences, and on production costs and profits, that there exists an EP-optimal point at which the producers maximize their profit – all in all not a bad economic situation if realizable in practice.

The mathematical background for such statements was provided by Cantor in his work on ordered sets (Ref. 8). (Recall that ordering the design set D to be able to compare designs is our basic objective). In this context, EP-optimality is only one optimality concept among an infinity of possible ones. As we shall see, however, there does seem to be something natural about the concept of EP-optimality in that it is possible to cast some

classical extremum principles in mechanics within this multicriteria framework. Furthermore, natural structures appear to be optima for suitably chosen criteria.

Orderings come in a variety of colors and guises ranging from partial to complete preorders to partial and linear orders; they can be smooth and connected or not, be imposed on arbitrary spaces or on $\mathbb{R}^N$, thus serving as a device for comparing virtually anything. Within the economic context, the so-called natural order on $\mathbb{R}^N$ is of greatest interest since it provides the *partial* order on the criteria values on which we have based the mathematical concept of EP-optimality. (The word "partial" refers to the fact that not *all* points are comparable to one another). The natural order on $\mathbb{R}^N$ is defined in terms of the usual order on the reals with $\mathbf{x},\mathbf{y} \in \mathbb{R}^N$ and

$$\begin{aligned} \mathbf{x} \leq \mathbf{y} &\quad \text{iff } x_i \leq y_i \text{ for every } i \in I = \{1,2,\ldots,N\} \\ \mathbf{x} \leq \mathbf{y} &\quad \text{iff } x_i \leq y_i,\ \mathbf{x} \neq \mathbf{y}, \text{ for } i \in I \\ \mathbf{x} \ll \mathbf{y} &\quad \text{iff } x_i < y_i, \text{ for every } i \in I \end{aligned}$$

We can easily get a geometric view of this situation. As before, we take minimization as the basic objective. At any given point $\mathbf{x}$, we imagine a cone with vertex at $\mathbf{x}$ and with the cone being a translate of the positive orthant of $\mathbb{R}^N$. The vertex $\mathbf{x}$ then is better than all of the points $\mathbf{y}$ in the cone; that is, $\mathbf{x} \leq \mathbf{y}$. Points outside the cone are not comparable to $\mathbf{x}$; e.g., the point $\mathbf{z}$ Points $\mathbf{x}^*$ on the dark border are the EP-optimal criteria values for the set A. We have depicted the situation in the criteria space $\mathbb{R}^N$. The ordering on $\mathbb{R}^N$ is thus used to induce an ordering in the design space D with $\mathbf{d}^*$ given by $\mathbf{g}(\mathbf{d}^*) = \mathbf{x}^*$, for example. The situation is depicted in Figure 7.

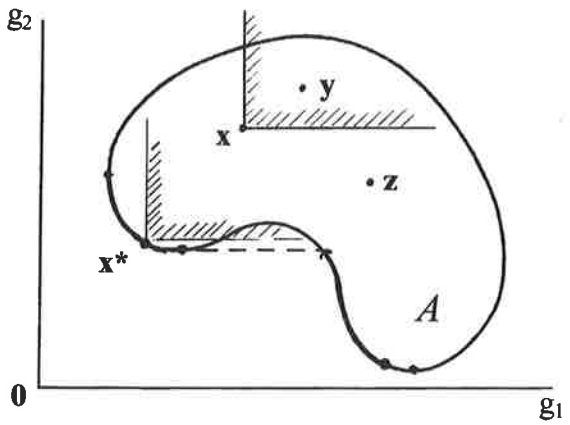


Figure 7. EP-optimality on $\mathbb{R}^N$

Here, the economic definition of the optimum was probably given without an awareness of

the ordering. Otherwise, it is common to define optimality relative to the ordering that has been imposed.

Classical mathematics has led us to view $\leq$ on $\mathbb{R}^1$ as the only plausible or natural ordering with the consideration of a single criterion and its maximum or minimum as a consequence. It is gradually becoming apparent that the use of a single criterion is a straightjacket which tends to distort actual decision making where compromises between conflicting choices are the norm. As a consequence, there also has been no use of multicriteria models to describe natural phenomena.

Emile Borel (1871 – 1956). A third and final area contributing to the solution of the vector maximum problem is provided by game theory. Not games of chance which have an ancient history eminently traced by Florence David in her book *Gods, Games and Gambling*, but games where the psychology of the player and his choices may affect the outcome. Borel writes (Ref. 9): "We will consider a game where the winning depends simultaneously on chance and on the skill of the player."

The minimax theorem incorporates the first definition of an optimal strategy for a game. Von Neumann laid claim to the founding of game theory by virtue of being the first to provide a general proof of the theorem in 1929 (Ref. 10), even though Borel had demonstrated the validity of the theorem for n = 3 and n = 5 as early as 1923. Von Neumann was aware of the work of Borel, but he never cited it; indeed, he only cites Borel's text on probability theory.

The minmax theorem deals with the existence of a saddle point for the payoff matrix of a rectangular game. A little insight into matrix games helps to appreciate the power of the theorem. Consider the payoff matrix

$$\begin{bmatrix} 6 & -2 & -1 & 0 \\ 4 & 3 & -1 & 3 \\ 3 & 4 & 1 & 2 \\ 2 & -3 & -2 & -1 \end{bmatrix}$$

and suppose a play of the game consists of P_1 choosing a row and P_2 choosing a column with neither having prior knowledge of the others choice. After they choose, player P_2 pays player P_1 the amount listed in the matrix. For example, if P_1 plays r2 and P_2 plays c2, then P_2 pays P_1 3 units of whatever the payoff consists of. A minus sign means that P_1 has to pay P_2.

With this in mind, it is clear that P_1 would never play r4, since he can do better by playing r1, r2 or r3 no matter what P_2 plays. We say that r4 is *dominated* by the other rows; that is, $(2,-3,-2,-1) < (4,3,-1,3)$.

Suppose now that we think of $a_{ij} = f(i,j)$, a real-valued function defined on the matrix element locations. A saddle point of the matrix game is a location (i_0,j_0) such that

$$a_{ijo} \leq a_{iojo} \leq a_{ioj}$$

or, equivalently,

$$a_{iojo} = \max_i \min_j a_{ij} = \min_j \max_i a_{ij}$$

Put another way, there is a selection (i_0,j_0) and an a_{iojo} such that a_{iojo} is at once the minimum of its row and the maximum of its column.

The desirability of playing this saddle point lies in the fact that by choosing r3, P_1 will receive at least 1 and by choosing c3, P_2 can keep P_1 from getting more than 1. This reasoning forms the basis for calling this choice the *optimal* choice for the game and the corresponding payoff the *value* of the game.

If every game had such a saddle point, then we would be done. Unfortunately, this is not the case.

As usual, we may expand our horizons and raise our expectations when we settle for probably rather than definitely. Suppose we are given an m x n game matrix A and we assume that P_1 plays row i with probability x_i and P_2 plays column j with probability y_j. This gives rise to the ordered tuples $\mathbf{x} = (x_1,x_2,\ldots,x_m)$ and $\mathbf{y} = (y_1,y_2,\ldots,y_n)$ which we constrain to belong to the simplexes $S_m = \{\mathbf{x}: x_1 + x_2 + \ldots + x_m = 1\}$ and $S_n = \{\mathbf{y}: y_1 + y_2 + \ldots + y_n = 1\}$. Clearly, x_i and y_j may also be thought of as the frequencies with which P_1 chooses row i and P_2 chooses column j, respectively.

We term $\mathbf{x} \in S_m$ and $\mathbf{y} \in S_n$ *mixed strategies* for P_1 and P_2, respectively. If both players play these mixed strategies, then the expectation of player P_1 is given by

$$E(\mathbf{x},\mathbf{y}) = \sum_{j=1}^{n} \sum_{i=1}^{m} a_{ij}x_iy_j$$

Again, we look for the saddle point as an ideal solution to such a game. That is, if there exists a pair $(\mathbf{x}^*,\mathbf{y}^*)$ such that

$$E(\mathbf{x},\mathbf{y}^*) \le E(\mathbf{x}^*,\mathbf{y}^*) \le E(\mathbf{x}^*,\mathbf{y})$$

then we term $(\mathbf{x}^*,\mathbf{y}^*)$ an optimal mixed strategy for P_1 and P_2 and $E(\mathbf{x}^*,\mathbf{y}^*)$ the value of the game to P_1.

Minmax Theorem. Let A be a given game matrix and let $E(\mathbf{x},\mathbf{y})$ be the expectation with $\mathbf{x} \in S_m$ and $\mathbf{y} \in S_n$. Then

$$\max_{\mathbf{x}\in S_m} \min_{\mathbf{y}\in S_n} E(\mathbf{x},\mathbf{y})$$

and

$$\min_{\mathbf{y}\in S_n} \max_{\mathbf{x}\in S_m} E(\mathbf{x},\mathbf{y})$$

both exist and are equal.

This is essentially the theorem proven by von Neumann in 1929 and it formed the nucleus for the unsurpassed study of game theory by von Neumann and Morgenstern entitled *The Theory of Games and Economic Behavior* (Ref. 11) first published in 1944. A very readable reference to game theory is McKinsey (Ref. 12) and we have used this text as our desk reference for this discussion.

Game theory evolved to include numerous other reasonable outcomes of games of conflict such as Nash equilibria, Stackelberg solutions and a plethora of others. This simply expresses the fact that there is an infinity of possible game equilibria which may be reasonably defined just as there is an infinity of possible optimality concepts between competing orderings which may be defined on the decision space. In due course, there also appeared the cooperative game where the players got more by cooperating with each other rather than competing against each other. It is not surprising that the optimality or equilibrium concepts of these cooperative games have much in common with those of the vector maximum problem.

The remaining contributions to the early development of the vector maximum problem are computational rather than conceptual. A first step in that direction was a scalarization of the problem as part of the development of welfare theory. This consists of the introduction of a welfare function

$$W(\cdot): \quad \mathbb{R}^N \rightarrow \mathbb{R}^1$$

defined over a subset of the criteria space $\mathbb{R}^N$. Of course, the most obvious such function is a linear combination of the criteria

$$W(\mathbf{g}) = c_1 g_1 + c_2 g_2 + \ldots + c_N g_N$$

and the concept actually evolved from this view to the acceptance of a general function of the criteria as conceived by Bergson in 1948 (Ref. 13). In this context, it is easy to show the following result.

Lemma (Ref. 14). Suppose $W(\cdot)$ is a compromise (welfare) function which is monotonically increasing in each of its arguments. Then $\mathbf{d}^* \in D$ such that

$$W(\mathbf{g}(\mathbf{d}^*)) = \max\{W \circ \mathbf{g}(\mathbf{d}): \mathbf{d} \in D\}$$

is an EP-optimal decision.

The inception of linear and nonlinear programming that followed then made possible the routine calculation of maxima and minima as well as EP-optima when the problem included inequality constraints. George Dantzig introduced linear programming and the simplex method in 1948 (Ref. 15) and the process was quite likely used by Koopmans (Ref. 16) in his linear production theory model where he introduced the term *efficiency* to characterize the EP-optimal solutions in 1951. Indeed, both worked together at the Cowles Commission. The general theory of nonlinear programming with inequality constraints was first dealt with by Karush in 1939 (Ref. 17), followed by a paper by F. John in 1948,

appropriately titled "Extremum Problems with Inequalities as Subsidiary Conditions" (Ref. 18) and eventually culminated in the well-known paper by Kuhn and Tucker in 1951 (Ref. 19). The last paper also included a section on the vector maximum problem, along with a statement of necessary conditions for EP-optimality. For further reading on the history of multicriteria optimization, see also Refs. 20,21, and 22.

4. EVOLUTION OF AN OPTIMIZATION PROBLEM

Practical optimization in an economic context has been around as long as mankind ranging from cooperative behavior where all goods and property were communal, to gaining an advantage in trade, to the amassing of great wealth by stealth or force. A knowledge of geometry was probably first put to practical use in the building of structures in Egypt many centuries before the rise of mathematics in Greece. According to Ref. 2: "It is likely that the first major advance that the Greeks made was to consider mathematical concepts (such as numbers and geometric figures) to be abstractions, creations of the human mind, and not part of the real world". It is this abstraction that allowed them to extend their work beyond practical use and to ask questions whose answers might not be immediately applicable to some physical problem or process. We now briefly trace the evolution of one such optimization problem that began with an application, then went to abstraction, and whose generalization now is a widely applicable research topic.

The story begins with the previously mentioned work of Heron concerning the reflection of light-rays in one or more mirrors. Dealing with problems on triangles was a natural extension of these investigations.

(1) *Given two sides a and b of a triangle to determine the maximum area.* The answer is a right triangle whose two legs are a and b.
(2) *Given the area a and one side c of a triangle to determine the triangle for which the sum of the remaining sides is the smallest.* Since A = ½ch, h is given. For a given h, a + b is smallest when the triangle is isosceles.
(3) *Given one side c and the sum a + b of the other two sides to find among all such triangles the one with the largest area.* This is the converse statement of the previous problem. The isosceles triangle has the largest area.

These problems and their solutions seem to have been known to the Greeks.

(4) **Fermat's Triangle Problem**. *Three cities A,B,C are to be connected by a system of roads. To find the system of least possible length* . The problem was posed by Fermat in 1646 and was solved by Evangelista Toricelli, Galileo's student, in 1646. The problem was resurrected by Jacob Steiner (1796 – 1863) a professor at the University of Berlin. This is no longer a simple triangle problem. We now are to find a point P (Fermat point) such that the total length of the paths joining P with the vertices is a minimum. For an acute triangle, the point is located in the interior of the triangle in such a way that the lines connecting P with the vertices radiate at 120° intervals (Figure 8(a)). For an obtuse triangle, the shortest network is one that connects the

vertex with the obtuse angle to the other two points (Figure 8(b)). According to Kuhn Ref. 23), the proof of this exceptional case was given by F. Heinen in 1834.

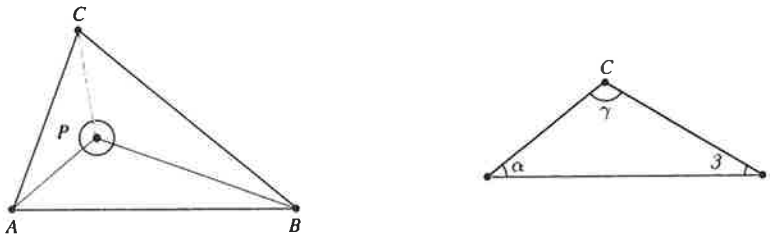


Figure 8. Fermat's Problem

(5) **Schwarz's Triangle Problem.** Hermann Amadeus Schwarz (1843 – 1921) was a Professor of Mathematics first at Goettingen and subsequently in Berlin. *Given an acute triangle, find an inscribed triangle with smallest possible perimeter.* The answer is the altitude triangle. That is, a triangle whose vertices are at the base points of the altitudes (Figure 9).

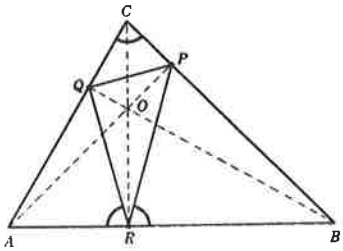


Figure 9. Schwarz's altitude triangle

We note that at each of these vertices, the adjacent angles are equal and conclude that a light ray would follow the perimeter of this triangle. For this reason, the altitude triangle is also called a light triangle. It is the only light triangle with three sides. There do exist other inscribed closed light paths that are not minimal.

(6) **Generalized Fermat Problem.** *Find a point P in the plane the sum of whose distances from n given points is a minimum.* This generalization was proposed by Simpson in his book *Doctrine and Application of Fluxions* (London, 1750) who also contributed to the solution of Fermat's Problem.

(7) **Jarnik and Kössler's Problem** (Ref. 24). They proposed the following extension of the road network problem: *Find the shortest network which interconnects n points in the plane.* This problem is currently a very active research area (see Ref. 25, for example) with applications in electrical, civil and mechanical engineering (Figure 10).

At this point, one might ask what the previous string of problems has to do with multicriteria optimization. Generalizations of any given problem are not always as obvious as they seem after the generalization has been voiced. It took more than 2000 years from Heron"s problem to the generalization of Jarnik and Kössler and it took almost the same amount of time to go from the contemplation of a single criterion to the simultaneous consideration of several criteria.

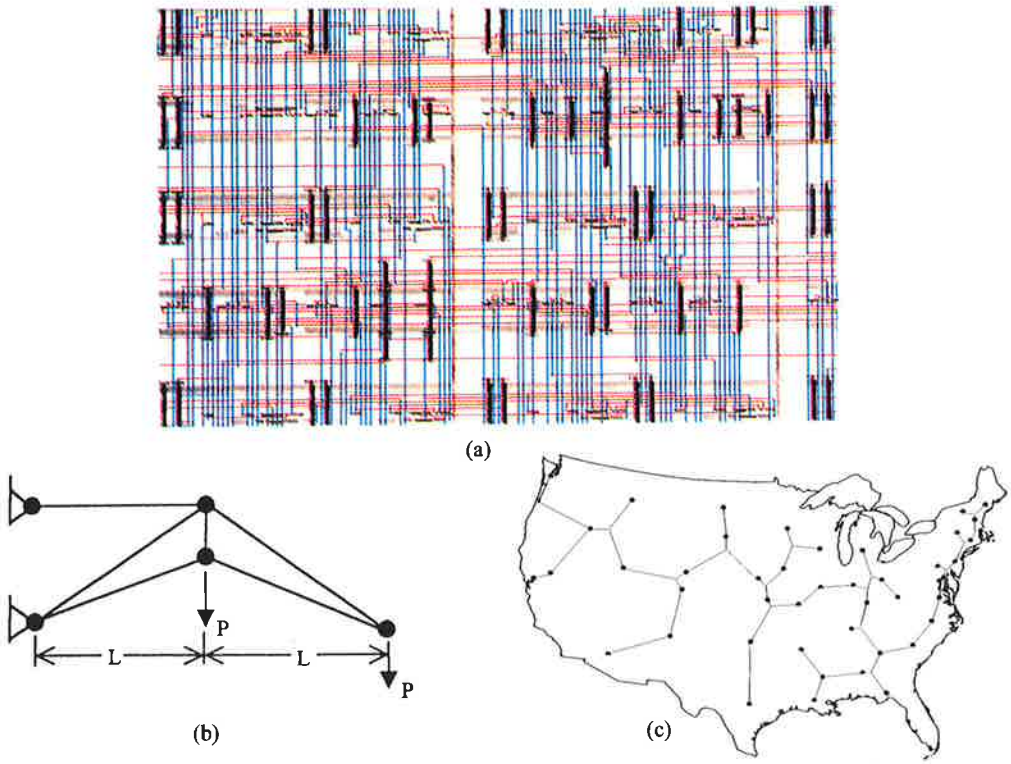


Figure 10. Current applications of the network problem. (a) Routing plan for the ZORA chip. (b) Fermat-Steiner truss. (c) Fermat tree connecting 49 cities in the US.

5. THE MULTICRITERIA VEIN IN MECHANICS

The first engineering application of the subject was published by Stadler (Ref. 26) and the method has since been applied to a wide variety of engineering problems ranging from axially symmetric extrusion, to the design of automatic braking systems to the design of the dish for a radio telescope. These and more may be found in a number of monographs (Refs. 14,27,and 28). The possible applications are limited only by the imagination, particularly so, since most problems in engineering require a trade-off between conflicting criteria.

Rather than dealing with a specific application, we shall take a multicriteria view of two classical principles, whose original formulation already involves two criteria, the potential energy and the kinetic energy. The investigation is an outgrowth of the author's concept of Natural Structural Shapes (Ref. 29), an application of multicriteria optimization to optimal structural design based on the simultaneous minimization of the mass and the strain energy of the loaded structure. The result was the following connection between Hamilton's principle and Rayleigh's principle which be now briefly describe (see also Ref. 30).

Suppose we consider a conservative elastic system with instantaneous kinetic energy $T(t;\mathbf{d})$ and potential energy $V(t;\mathbf{d})$, $\mathbf{d} \in D$, being a design variable. Introduce the criteria

$$g_1(\mathbf{d}) = \int_{to}^{t_1} V(t;\mathbf{d})\, dt \quad \text{and} \quad g_2(\mathbf{d}) = \int_{to}^{t_1} T(t;\mathbf{d})\, dt$$

with the basic intent of minimizing $g_1(\mathbf{d})$ and maximizing $g_2(\mathbf{d})$ subject to $\mathbf{d} \in D$ and with $V(t;\mathbf{d}) + T(t;\mathbf{d}) = \text{constant}$.

It can be shown that the ratio $R(\mathbf{d}) = g_1(\mathbf{d})/g_2(\mathbf{d})$ is a minimum if and only if $g_1(\mathbf{d})$ is a minimum and $g_2(\mathbf{d})$ is a maximum or equivalently, if and only if

$$G_0(\mathbf{d}) = g_1(\mathbf{d}) - g_2(\mathbf{d})$$

is a minimum. Thus, the modified Rayleigh Quotient is a minimum if and only if the usual Hamiltonian is a minimum.

With these results in mind, we now formulate the extremal conditions for the corresponding multicriteria principle. The corresponding multicriteria extremum principle is based on the necessary conditions for EP-optimality. These may be obtained from those for single criterion problems by replacing the single criterion with the linear combination of criteria,

$$G(\mathbf{d}) = c_1 g_1(\mathbf{d}) - c_2 g_2(\mathbf{d})$$

with the ultimate result that the ratio c_2/c_1 multiplies the frequency or, equivalently, that the frequency parametrizes the set of EP-optima. Thus, the multicriteria principle is a generalization of the original principle admitting a *family* of possible motions.

6. CONCLUSION

We close our presentation with some comments on the teaching of optimization. From the preceding, it is evident that optimization is a vast and vibrant field whose results pervade all of engineering and the sciences ranging from minimum principles in mechanics to the optimal packing of cells in a honeycomb. Thus, optimization courses should be core courses in a general curriculum and they should be central to curricula in Engineering Science and Mechanics.

Optimization has occupied some of the greatest minds from Galileo to Huygens to Euler to Lagrange to Rayleigh and Hamilton. Yet the power of optimization methods is barely realized, is rarely taught at an undergraduate level and is considered to be a difficult topic at the graduate level. This is a fallacy. Design optimization should be taught at every level of instruction. If the subject is to be widely used, it is imperative that it be taught at the undergraduate level. Indeed, the subject lends itself nicely to a structured approach since its methods range from finite mathematics to the calculus of variations and optimal control.

The level of the presentation thus could be adjusted to the students' concurrent learning experience.

The author has taught a senior course to students with the usual background in physics and calculus and with courses in dynamics, circuits, strength of materials and systems analysis. As their final examination, they were asked to criticize and review randomly selected articles on optimization in their respective areas of interest or specialization. Two of the students found errors in the papers they reviewed, one of them so much so that the results of the paper were all but meaningless. All of them found the subject to be stimulating and challenging, but manageable.

7. REFERENCES

(1) Grunal, P.: *The Dictionary of Classical Mythology*, translated by A.R. Maxwell-Hyslop. English translation published by Basil Blackwell Publishers, New York, 1986. Originally published in French as *Dictionnaire de la Mythologie Greque et Romaine*, Presses-Universitaire de France, Paris, 1951.

(2) Hildebrandt, S. and A. Tromba: *The Parsimonious Universe (Shape and Form in the Natural World)*, Copernicus, Springer Verlag, New York, New York,1995

(3) Boyer, Carl, B.: *A History of Mathematics*, John Wiley and Sons, Inc., New York, New York,1968.

(4) Courant, R. and H. Robbins: *What is Mathematics?*, Oxford University Press, New York, New York, 1941.

(5) Euler, L.: *Methodus Inveniendi Lineas Curvas Maximi Minimive Proprietate Gaudentes.* (Lausanne and Geneva, 1744). English translation by W.A. Oldfather, C.A. Ellis, and D.M. Brown, ISIS 20 (1933)68.

(6) Edgeworth, F.Y.: *Mathematical Psychics*, C. Kegan-Paul & Co., 1 Paternoster Square, London, England, 1881.

(7) Pareto, V.: *Manuale di Economia Politice,* Societa Editrice Libraria, Milano Italy, 1906. Translated into French, with revised Mathematical Appendix, by Girard and Briere, as *Manuel D'Economie Politique,* Graid, Paris, France, First Edition, 1909 and Second Edition, 1927. Translated into English by A.S. Schwier as *Manual of Political Economy*, The MacMillan Company, New York, New York, 1971.

(8) Cantor, G.: Contributions to the Foundation of Transfinite Set Theory (in German), *Mathematische Annalen*, Vol. 46, pp. 481-512, 1895, and Vol. 49, pp. 207-246, 1897.

(9) Borel, E.: The Theory of Play and of Integral Equations with Skewsymmetric Kernels (in French), *Comptes Rendus Academie des Sciences*, Vol. 173, pp. 1304-1308. English translation in *Econometrica*, Vol. 21, No. 1, pp. 97-100, January 1953.

(10) von Neumann, J.: On the Theory of Parlor Games (in German), *Mathematische Annalen,* Vol. 100, pp. 295-320, 1929.

(11) von Neumann, J. and O. Morgenstern: *Theory of Games and Economic Behavior,* 2nd Edition, Princton University Press, Princeton, N.J., 1947.

(12) McKinsey, J.: *Introduction to the Theory of Games,* The RAND Series, McGraw-Hill, New York, New York, 1952.

(13) Bergson, A.: Socialist Economics, in *A Survey of Contemporary Economics,* Vol. 1, edited by H.S. Ellis, Irwin, Homewood, Illinois, pp. 412-447. 1948.

(14) Stadler, W.: Fundamentals of Multicriteria Optimization, Chapter 1, in *Applications of Multicriteria Optimization in Engineering and in the Sciences,* edited by W. Stadler. (Series on Mathematical Methods in the Sciences and Engineering, edited by A. Miele), Plenum Press, New York, New York,1988.

(15) Dantzig, G.B.: Maximization of a Linear Function of Variables Subject to Linear Inequalities, Chapter 21 in *Activity Analysis of Production and Allocation,* Cowles Commission Monograph, No. 13, edited by. T.C. Koopmans, John Wiley & Sons, New York, New York, pp.

(16) Koopmans, T.C.: Analysis of Production as An Efficient Combination of Activities, *Activity Analysis of Production and Allocation,* Cowles Commission Monograph, No. 13, edited by T.C. Koopmans, John Wiley & Sons, New York, New York, pp. 33 – 97, 1951.

(17) Karush, W.: *Minima of Functions of Several Variables with Inequalities as Side Conditions,* MS Thesis, Department of Mathematics, University of Chicago, Chicago, Illinois, 1939. (For a summary of this work see: H.W. Kuhn, Nonlinear Programming: A Historical View, in *Nonlinear Programming*, SIAM_AMS Proceedings 9, edited by R.W. Cottle and C.E. Lempke, pp.1 – 26).

(18) John, F.: Extremum Problems with Inequalities as Subsidiary Conditions, *Studies and Essays*, Presented to R. Courant on his 60th Birthday, January 8, 1948, Interscience, New York , New York,1948, pp. 187 – 204..

(19) Kuhn, H.W. and A.W. Tucker: Nonlinear Programming, *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability,* edited by J. Neyman, Berkeley, California, pp. 481 – 492, 1951.

(20) Stadler, W.: A Survey of Multicriteria Optimization or the Vector Maximum Problem, Part I, 1776 – 1960, *Journal of Optimization Theory and Applications,* Vol. 29, No. 1, September 1979, pp. 1 – 52.

(21) Stadler, W.: A Survey of Vector Optimization in Infinite Dimensional Spaces, Part II, *Journal of Optimization Theory and Applications,* Vol. 51, No. 2, November 1986, pp. 205 – 242.

(22) Stadler, W.: Initiators of Multicriteria Optimization, in *Recent Advances and Historical Development of Vector Optimization,* (Proceedings of an International Conference on Vector Optimization, held at the Technical University of Darmstadt, Germany, August 4 – 7, 1986), edited by J. Jahn and W. Krabs, Lecture Notes in Economics and Mathematical Systems, No. 294, Springer Verlag, Berlin, Germany, 1987.

(23) Kuhn, H.W.: Steiner's Problem Revisited, *Studies in Optimization,* Vol. 10 of *Studies in Mathematics*, edited by G.B. Dantzig and B.C. Eaves, The Mathematical Association of America, 1974.

(24) Jarnik, V. and O. Kössler: On Minimal Graphs Containiing n Given Points (in Czech), *Casopis Pro Pestovani Matematiky a Fysiky,* 63 (1934) pp. 223 – 235.

(25) Hwang, F.K., D.S. Richards and P. Winter: *The Steiner Tree Problem,* Elsevier Science Publishers, The Netherlands, 1992.

(26) Stadler, W.: Cooperative Games for the Experimentalist (preliminary results), *Proceedings of the 5th International Conference on Systems Science*, University of Hawaii, Hawaii, 1972

(27) Eschenauer, H., J. Koski and A. Osyczka (Eds.): *Multicriteria Design Optimization Procedures and Applications,* Springer Verlag, Berlin, Germany, 1990.

(28) Statnikov, R.B.: *Multicriteria Design (Optimization and Identification),* Kluwer Academic Publishers, Dordrecht, The Netherlands, 1999.

(29) Stadler, W.: Natural Structural Shapes (The Static Case), *Quarterly Journal of Mechanics and Applied Mathematics,* Vol. 31, Part 2, May 1978, pp. 169 - 217.

(30) Stadler, W.: A Multicriteria View of the Rayleigh Quotient, *Proceedings of the Fourth World Congress of Structural and Multidisciplinary Optimization,* [Goslar, Germany , May 28 – June 2, 1995], edited by N. Olhoff and G.I.N. Rozvany, Pergamon Press, New York, New York,1995, pp. 789 – 796.

PECULIAR FEATURES OF AXIALLY MOVING MATERIALS

A. PRAMILA
Engineering Mechanics Laboratory
Department of Mechanical Engineering
P.O. Box 4200
FIN-90014 University of Oulu, FINLAND

ABSTRACT

Basic characteristics of vibration of flexible structures where the structure itself or fluid moving along it are considered. The motivation of the research of such systems, called axially moving materials, is discussed and brief historical notes are given indicating success and shortcomings of the work done. Many complex engineering applications involve as a part axially moving materials, e.g. fiber winding, paper sheet in a paper mill, magnetic tapes, band saw blades and pipes transporting fluids. A prototype model, axially moving string or narrow band, is chosen in order to be able to illustrate analytically the peculiar features, e.g. transport velocity dependent natural frequencies, and the existence of a critical speed at which instability occurs.

INTRODUCTION

The problem area of axially moving materials involves flexible structures where the structure itself or fluid moving along it has prescribed translational motion and that can vibrate about the state of steady translation. Typical engineering applications are paper sheets in paper machines, band saw blades, magnetic tapes, power transmission belts and chains, fluid conveying pipes and aerial cable tramways. The stimulus of the present author has been the first mentioned application. The engineering applications include many kind of complexities like contact with surrounding structure and interaction with surrounding fluid which make the formulation of the problem difficult and the solution even more difficult. Therefore the fundamental research has considered mainly isolated prototype structures like travelling strings, narrow bands, beams, plates, fluid conveying pipes and shells.

A detailed historical review is not possible nor necessary here. The work done within the field of fluid conveying pipes until 1991 has been reviewed by Paidoussis and Li[1] and and within the field of travelling strings and beams until 1987 by Wickert and Mote [2]. The first mentioned contains over 200 references and the last mentioned over 100 references (sic!).

The research of the latter topic was at the beginning curiosity driven. Skutsch, who derived more than hundred years ago the first correct equation for the natural frequencies of an axially moving string, wrote at the beginning of his article [3] "Das hier vorliegende Problem, welches

neben theoretischem Intresse vielleich auch die Möglickeit einer technischen Anwendung bietet, ... ". However, already at he end of fifties there appeared practical reasons to study the dynamical behaviour of an axially moving tape in Bell Telephone Laboratories [4]. The interest into this, perhaps simplest, problem has remained until these days [5] even in the linear case. Many fundamental issues, like the kinematics of large deflections, have been clarified not until recently [6]. Depending on simplifications made, the solution methods have been between exact closed form solutions [7] and FEM approximations [8], [9].

On the other hand, the early research on vibrations of fluid conveying pipes was linked to practical applications. The first attempt [10] associated with the vibration problems of Trans-Arabian oil pipeline omitted, however, the transverse force due to curvature of the pipe and only half of the Coriolis effect was taken into account. Next approach [11] led to the correct equation of motion, but shortcomings in the boundary conditions due to inappropriate use of Hamilton's principle. The probably first author giving correct formulation and solutions for the problem was Niordson [12]. Also here the interest has remained until recent days [13].

The purpose of the present paper is to clarify as simply as possible some basic, and slightly peculiar characteristics of axially moving materials including transport velocity dependent natural frequencies, eigenmodes with spatially distributed phase, and the existence of a critical velocity at which divergence or flutter instability occur.

AXIALLY MOVING STRING

Equation of motion and its solution

Consider a uniform, flexible, string or narrow band of mass per unit length m, tension P translating along x-axis with constant speed v in vacuum for simplicity. The transverse displacement of the material point located instantaneously at point x is denoted by $w(x, t)$, i.e. $w(x, t)$ describes the configuration of the string as a function of time. The transverse velocity of the material point located instantaneously at x is

$$\frac{dw}{dt} = \frac{\partial w}{\partial t} + v\frac{\partial w}{\partial x} \tag{1}$$

where the first term is due to the change of the configuration at point x and the second due to the axial velocity and nonzero slope.

Here, small displacements are assumed. Kinematics of axially moving strings with large displacements has been recently considered by Koivurova&Salonen, [6].

The operator of the material time derivative is according to equation (1)

$$\frac{d\square}{dt} = \frac{\partial \square}{\partial t} + v\frac{\partial \square}{\partial x} \tag{2}$$

By applying it to the velocity (1) we obtain the acceleration of the material point

$$\frac{d^2w}{dt^2}=\frac{\partial^2 w}{\partial t^2}+2\,v\frac{\partial^2 w}{\partial x\partial t}+v^2\frac{\partial^2 w}{\partial x^2} \tag{3}$$

Thus, the equation of motion reads

$$m\frac{\partial^2 w}{\partial t^2}+2mv\frac{\partial^2 w}{\partial x\partial t}+mv^2\frac{\partial^2 w}{\partial x^2}-P\frac{\partial^2 w}{\partial x^2}=0 \tag{4}$$

if there are no transverse forces present.

From equation (4) we immediately see that the net restoring force vanishes when $v=(P/m)^{1/2}$, i.e. this is the critical velocity causing divergence type instability.

The last term in equation (4) is the restoring force due to the tension of the string and the first one is the transverse "inertia force" due to the change of w at point x which, however, is occupied by different material points at different time instants, therefore the citation marks. The third term is the "centrifugal force" due to the curvature of the string. The second term is due to the Coriolis effect, because the material elements moving with constant velocity v along the configuration are turning with an angular velocity $\partial/\partial t\ (\partial w/\partial x)$. Only the two first mentioned terms are present in the equation of motion of a stationary string. The two last mentioned depend on axial velocity and are the salient feature of axially moving material systems. They are also guilty for the peculiar features of the dynamic behaviour of such systems.

Except the Coriolis term the terms are in phase or in opposite phase with w. The Coriolis term is in phase with the angular velocity and thus lags w by 90°. If we assume harmonic variation of w with respect to time and x, all terms except the Coriolis term will equal to zero at the straight equilibrium position. The Coriolis term will be at its maximum at that time. Hence, the system can not have classical eigenfrequencies and eigenmodes where all points move harmonically in phase passing the equilibrium position at the same time.

This can be seen also by considering the possible solutions for the case where both ends are supported. If the wave propagation velocity of a stationary string is denoted by a, $a^2=P/m$, the disturbances in an axially moving string propagate with velocity $(a+v)$ into the upstream direction and with velocity $(a-v)$ into the downstream direction. It can be shown by simple substitution that

$$w=g_1(x-(a+v)t)+g_2(x+(a-v)t) \tag{5}$$

is a solution of equation (4), g_1 and g_2 being arbitrary functional relationships of the parameters $x-(a+v)t$ and $x+(a-v)t$. The fundamental frequency for an axially moving string supported at

both ends, distance L apart from each other can be deduced as follows. A disturbance caused at the left end propagates with velocity $(a+v)$ towards the right end and arrives to the right end at time $L/(a+v)$. The disturbance reflects there, propagates towards the left end with velocity $(a-v)$ and arrives back there at time $L/(a+v)+L/(a-v)$. Thus, the string is after immediate reflection again in the same state after a period

$$T = \frac{L}{a+v} + \frac{L}{a-v} = \frac{2aL}{a^2 - v^2} \tag{6}$$

which means that the fundamental frequency is

$$f = \frac{a}{2L}\left(1 - \frac{v^2}{a^2}\right) \tag{7}$$

This result was obtained by Skutsch already 1897 [3]. The natural vibration of a stationary string can be regarded as standing waves produced by travelling waves of equal velocity, wavelength and amplitude, but proceeding in opposite directions. Now, however the upstream and downstream velocities are different. Thus, there are no eigenmodes in classical sense.

Equation (7) reveals the peculiar feature of axially moving material systems that the natural frequencies depend on velocity and are the smaller the higher is the axial velocity. In this case, when the axial velocity v equals a, f vanishes and the system becomes unstable by divergence.

Equation (4) can be rewritten as

$$\frac{\partial^2 w}{\partial t^2} + 2\, v\frac{\partial^2 w}{\partial x \partial t} + v^2 \frac{\partial^2 w}{\partial x^2} - a^2 \frac{\partial^2 w}{\partial x^2} = 0 \tag{8}$$

If a separable solution

$$w(x,t) = W(x)e^{i\omega t} \tag{9}$$

is assumed and substituted for w in equation (8) we obtain an ordinary differential equation

$$-\omega^2 W + 2i\omega v \frac{dW}{dx} - (a^2 - v^2)\frac{d^2 W}{dx^2} = 0 \tag{10}$$

By using the traditional approach, i.e. by assuming

$$W = e^{kx} \tag{11}$$

we are led to the characteristic equation

$$-\omega^2 + 2i\omega vk - (a^2 - v^2)k^2 = 0 \tag{12}$$

the roots of which are

$$k_1 = \frac{i\omega}{a - v} \quad , \quad k_2 = -\frac{i\omega}{a + v} \tag{13}$$

Thus, the solution of equation (10) is of the form

$$W(x) = Ae^{-\frac{i\omega x}{a+v}} + Be^{\frac{i\omega x}{a-v}} \tag{14}$$

where A and B depend on boundary conditions. Because $W(0)=0$, $B=-A$ and because $W(L)=0$, we must satisfy condition

$$e^{-\frac{i\omega L}{a+v}} - e^{\frac{i\omega L}{a-v}} = 0 \tag{15}$$

which in turn requires that

$$\frac{\omega L}{a - v} = -\frac{\omega L}{a + v} + 2n\pi \tag{16}$$

Thus, the natural angular frequencies are

$$\omega_n = \frac{n\pi a}{L}(1 - \frac{v^2}{a^2}) \tag{17}$$

With $n=1$ we obtain the same fundamental frequency as given already by equation (7).

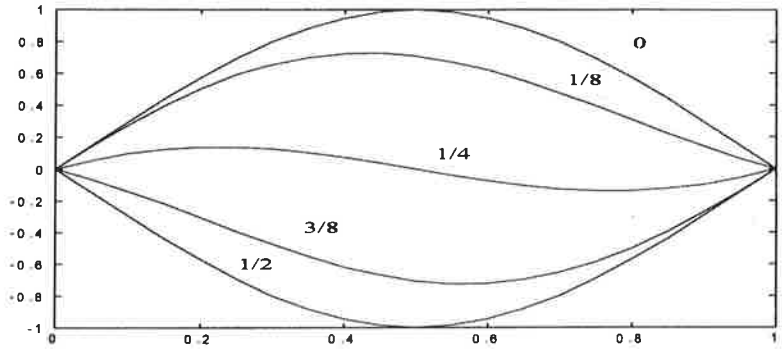
The general solution of equation (8) is

$$w(x,t) = \sum_{n=1}^{\infty} (A_n e^{-in\pi\frac{(a-v)x}{aL}} - A_n e^{in\pi\frac{(a+v)x}{aL}}) e^{i\frac{n\pi a}{L}(1-\frac{v^2}{a^2})t} \tag{18}$$

where the complex coefficients A_n are determined by the initial conditions. Details of derivation of the solution for arbitrary initial conditions was presented as late as 1990 by Wickert and Mote [7]. The form of equation (18) clearly shows that there is a nonzero phase difference between w at different points whenever v is not equal to zero. This is illustrated for

the fundamental frequency during a half cycle in Figure 1.

a)



b)

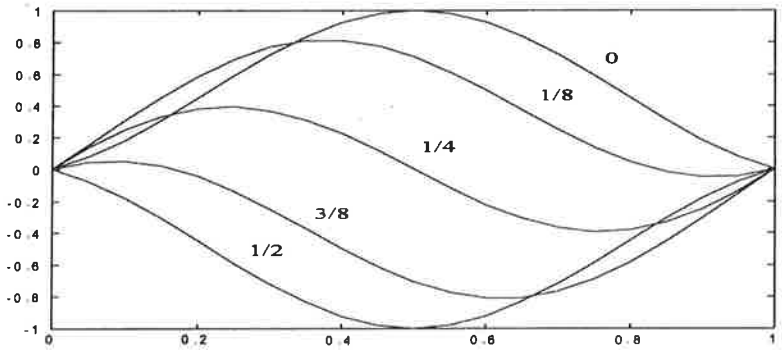


Figure 1. Shape of the string vibrating at its fundamental frequency a) *v/a=1/4* b) *v/a=3/4*. The numbers indicate the fraction of the period, *t/T*.

When both ends are supported there are no peculiarities with the boundary conditions. However, when either of the ends is free or forced to vibrate we are led to different conditions depending on the assumption how the string arrives to the domain and leaves the domain. If the string is forced to vibrate by using transverse force F at the right end and the left end is supported, the boundary condition

$$F = P\frac{\partial w}{\partial x} - mv\frac{\partial w}{\partial t} - mv^2\frac{\partial w}{\partial x} \tag{19}$$

applies at the right end if the string leaves the domain horizontally and

$$F = P\frac{\partial w}{\partial x} \tag{20}$$

if it leaves tangentially the domain. The two last terms in equation (19) are due to the change of the direction of the momentum flux. Equation (19) has been used by Le-Ngoc and McCallion [5] in deriving the dynamic stiffness matrix for an axially moving string. However, when the extended Hamilton's principle is used for the derivation of the equation of motion and boundary conditions we arrive to natural boundary condition at right end corresponding to equation (20) with $F=0$.

For an axially moving string with boundary conditions

$$w(0,t) = 0, \quad P\frac{\partial w(L,t)}{\partial x} = 0 \tag{21}$$

it is more instructing to assume that

$$w(x,t) = W(x)e^{\lambda t} \tag{22}$$

leading to solution of the form

$$W(x) = Ae^{-\frac{\lambda x}{a+v}} + Be^{\frac{\lambda x}{a-v}} \tag{23}$$

Application of boundary conditions reveals that the eigenvalues are now complex, $\lambda=\sigma+i\omega$, the real parts of which are

$$\sigma = -\frac{a}{2L}(1-\frac{v^2}{a^2})\ln\left|\frac{1+v/a}{1-v/a}\right| \tag{24}$$

and the imaginary ones are the natural angular frequencies

$$\omega_n = \frac{(2n-1)\pi a}{2L}(1-\frac{v^2}{a^2}) \tag{25}$$

Thus, with any initial conditions the motion is decaying when $0<v<a$, because the real parts of the eigenvalues are negative. Note that the damping factor (24) is independent of mode number n.

If we use boundary conditions (19) purely imaginary eigenvalues are obtained, the imaginary parts being the ones given by equation (25).

Forced vibrations

If a harmonically varying transverse force

$$F(t) = F_0 e^{i\Omega t} \tag{26}$$

is assumed to the right end we obtain (by repetition of the steps from equation (9) to (14) but now Ω replacing ω) a solution of the type

$$w(x,t) = W(x) e^{i\Omega t} \tag{27}$$

where

$$W(x) = A(e^{-\frac{i\Omega x}{a+v}} - e^{\frac{i\Omega x}{a-v}}) \tag{28}$$

Depending whether (19) or (20) is used the equation for determination of A is

$$F_0 = P\frac{dW(L)}{dx} \tag{29}$$

or

$$F_0 = P\frac{dW(L)}{dx} - mv\, i\Omega W(L) - mv^2 \frac{dW(L)}{dx} \tag{30}$$

respectively.

The resulting equations for transfer function $Y(x,\Omega)=W(x)/F_0$ become quite lengthy but can be dealt with easily by symbolic mathematical software like Derive [14].

The absolute value of the dimensionless transfer function $PW(L)/F_0L$ is shown in figure 2. We can see the peculiar feature that the absolute value of the frequency response function remains finite at natural frequencies when the natural choice of tangential exit, equation (20), is used in spite of the fact that we do not have damping present. The choice (19) leads to a dimensionless transfer function which does not equal to 1 at Ω=0. The reason is the third term in equation (30).

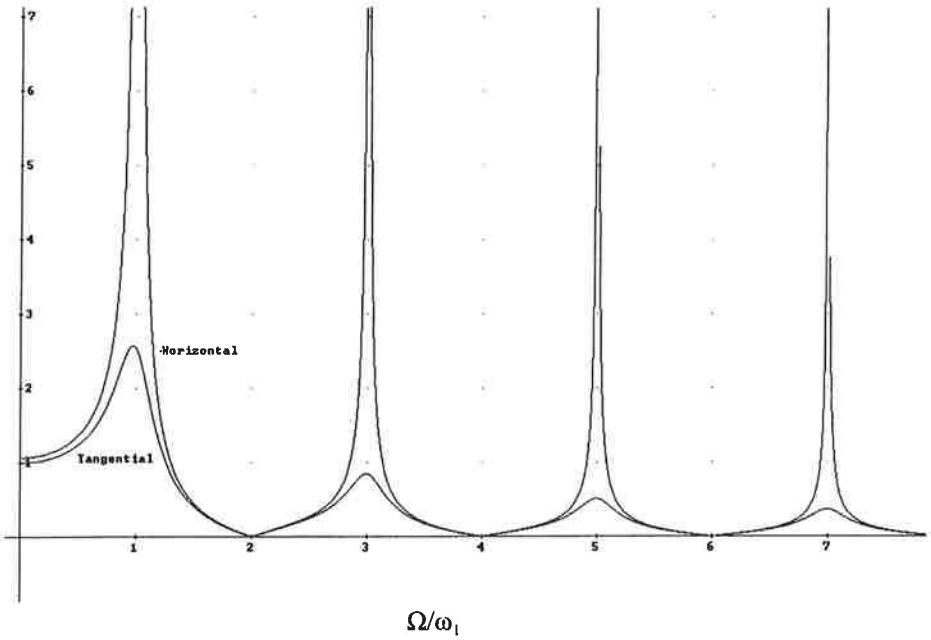


Figure 2. The absolute value of the dimensionless transfer function as function of the dimensionless angular frequency, *v/a=1/4*.

INFLUENCE OF SURROUNDING FLUID

Axially moving narrow band vibrating in air

The first comparisons of result (17) with experimental results obtained in a pilot paper mill [15] showed considerable differences (over 400%) as expected even with zero axial velocity. However, when the added mass effect of the surrounding fluid was taken into account in a traditional manner, the results agreed quite well. If ideal fluid assumption is made, the present case does not differ from the vibrations of stationary band, because of the slip between the band and the air in longitudinal direction. Thus, the added mass would appear just in the first term.

In practice the flow is viscous, and there is a thin boundary layer above and below the band where the velocity components of the flow in x and y directions should equal to v and $\partial w/\partial t + v\,\partial w/\partial t$ on the surface of the band, respectively. Far from the band the velocities should approach zero. Thus, one possibility could be to include the effect of surrounding fluid to second and third terms in equation (4) based on the displacement thickness and momentum thickness of the boundary layer, respectively. The present purpose is to deal with the characteristic features of the problem by using as simple as possible models amenable to analytical solution and that is why we deliberately assume these thichknesses equal to zero.

Thus, the equation of motion is based on the ideal fluid assumption and reads

$$(m+m_a)\frac{\partial^2 w}{\partial t^2}+2mv\frac{\partial^2 w}{\partial x \partial t}+mv^2\frac{\partial^2 w}{\partial x^2}-P\frac{\partial^2 w}{\partial x^2}=0 \tag{31}$$

By proceeding in a similar way as with the string (*vide,* equations (9)...(16)) we arrive at the equation for natural angular frequencies

$$\omega_n=\frac{n\pi a(1-v^2/a^2)}{L\sqrt{1+r-rv^2/a^2}} \tag{32}$$

where $r=m_a/m$. The result agrees with the one obtained in [16] in a slightly different way. The first natural frequency as a function of velocity is shown in figure 3 together with experimental results from [16]. The non dimensional frequency and velocity are defined as follows

$$V=v\sqrt{\frac{m}{P}} \quad ,F=2f_1L\sqrt{\frac{m}{P}} \tag{33}$$

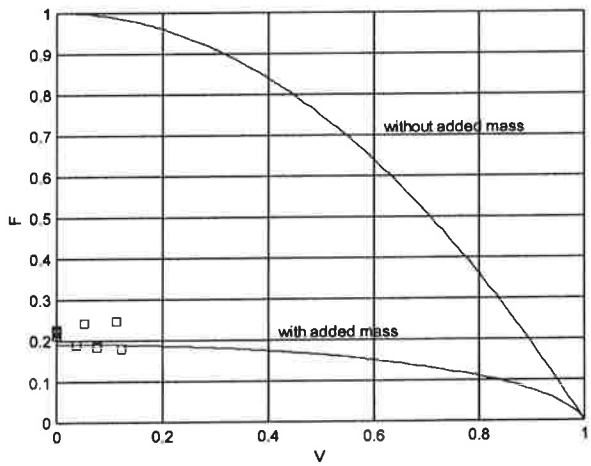


Figure 3. First natural frequency as a function of velocity. Mass per unit area 35.5g/m^2, length 2.4m and width 0.47m. Experimental results [15] are denoted by squares.

As can be seen from the figure the surrounding fluid plays an important role.

CONCLUSIONS

Peculiar characteristics of axially moving materials are clarified by using an axially moving string or narrow band as an idealized prototype structure. The system has velocity dependent

natural frequencies and the corresponding modes of vibration have spatially varying phase. The system loses its stability by divergence at the critical velocity. The transfer functions may remain finite at natural frequencies even in undamped cases depending on boundary conditions applied. When the axially moving narrow band is light, the surrounding air has a considerable effect on the eigenfrequencies.

REFERENCES

1. Paidoussis, M.P. and Li, G.X (1993) Pipes conveying fluid: a model dynamical problem. *Journal of Fluids and Structures* **7**, 137-204.

2. Wickert, J.A. and Mote Jr, C.D. (1988) Current research on axially moving materials. *Shock and Vibration Digest* **20**(5), 3-13.

3. Skutsch, R. (1897) Über die Bewegung eines gespannten Fadens, welcher gezwungen ist, durch zwei feste punkte mit einer constanten Geschwindigkeit zu gehen und zwischen denselben in Transversalschwingungen von geringer Amplitude versetzt. *Annalen der Physik und Chemie* **61**, 190-195.

4. Miranker, W.L. (1960) The wave equation in a medium in motion. *IBM Journal of Research and Development* **4**(1), 36-42.

5. Le-Ngoc, L. and McCallion (1999) H., Dynamic stiffness of an axially moving string. *Journal of Sound and Vibration* **220**(4), 749-756.

6. Koivurova, H. and Salonen, E.-M. (1999) Comments on non-linear formulations for travelling string and beam problems. *Journal of Sound and Vibration* **225**(5), 845-856.

7. Wickert, J.A. and Mote Jr, C.D. (1990) Classical vibration analysis of axially moving continua. *ASME Journal of Applied Mechanics* **57**, 738-744.

8. Pramila, A., Laukkanen, J. and Pautamo, M. (1983) Vibration of axially moving material by the FEM. *ASME 1983 Vibrations Conference,* Dearborn, Michigan, USA, September 1983.

9. Koivurova, H. (1998) Dynamic behaviour of an axially moving membrane interacting with the surrounding air and making contact with supporting structures. *Doctoral Thesis*. University of Oulu.

10. Ashley, H. and Haviland, G. (1950) Bending vibrations of a pipe line containing flowing fluid. *ASME Journal of Applied Mechanics* **17**, 229-232.

11. Housner, G.W. (1952)Bending vibrations of a pipe line containing flowing fluid. *ASME Journal of Applied Mechanics* **19**, 205-208.

12. Niordson, F. (1953) Vibration of a cylindrical tube containing flowing fluid. *Transactions*

of Royal Institute of Technology **73**, 1-27.

13. Panussis, D.A. and Dimarogonas, A.D. (2000) Linear in-plane and ou-of-lane lateral vibrations of a horizontally rotating fluid-tube cantilever. *Journal of Fluids and Structures* **14**(1), 1-24.

14. Kuzler, B. (1996) Introduction to DERIVE® for Windows. Gutenberg-Werbering GmbH, Linz, Austria.

15.Pramila, A. (1986) Sheet Flutter and the interaction between sheet and air. *TAPPI Journal* **69**(7),70-74.

16. Pramila, A. (1987) Natural frequencies of a submerged axially moving band. *Journal of Sound and Vibration* **113**(1), 198-203.

COMPUTATION OF PARETO OPTIMA FOR COMPOSITE LAMINATES SUBJECTED TO MULTIPLE LOADING CONDITIONS

Petri Kere [1†] and Juhani Koski [2]
[1] School of Engineering, National University of Uruguay
J. Herrera y Reissig 565, 11300 Montevideo, Uruguay
[2] Laboratory of Applied Mechanics, Tampere University of Technology
P.O. Box 589, 33101 Tampere, Finland

ABSTRACT

A method for computing Pareto optima for composite laminates subjected to multiple loading conditions is introduced. Laminate failure margins with respect to the applied loading conditions are treated as criteria. The original problem is reduced to a bicriterion problem and solved through an interactive iterative procedure. Improvement of the criterion values through the iterative procedure, the iteration history in the reduced criterion space, convergence of the procedure with respect to the used metric, and maximal points in the reduced criterion space corresponding to the Pareto optimal laminate lay-up configurations are illustrated with an example and discussed.

1. INTRODUCTION

Structures composed of fiber-reinforced composites are typically laminated composite plates and shells subjected to external mechanical loads, and internal loads caused by the changes in the operating temperature and moisture content. Laminated composites are frequently analyzed and designed by using the so-called point analysis methods, that is, the laminate lay-up configuration is analyzed and optimized for the loading applied at a given point. When the loading varies with the point position, the designer needs to consider laminate designs that maximize laminate strength under several different loading conditions. Furthermore, different loading conditions usually occur at different times under in-service conditions, which calls for an optimization formulation that is capable of taking into account the various loading conditions. A natural formulation for a problem with several competing criteria is the multicriterion (multiobjective, vector) optimization formulation, where the conflicting design objectives are introduced in a vector objective function and so-called Pareto optima (efficient solutions, vector optima) are sought. In the

† On leave of absence from Helsinki University of Technology. Correspondence should be addressed to petri.kere@hut.fi.

present work, the objective is to find such symmetric and balanced laminate lay-up configurations with constant ply properties and given initial allowable angles for layer orientations that the laminate failure margins with respect to the applied loading conditions are maximized with minimum feasible number of layers. Thus, we consider a multicriterion optimization problem where laminate failure margins with respect to the applied loading conditions are treated as criteria.

We describe an interactive iterative procedure, where some parameters are introduced to reduce the number of the criteria. The decision-maker can interact in the computation by setting the parameters on the basis of his preferences and the past computations. The procedure is defined through a point-to-set mapping, where at each cycle a new set of discrete design points, i.e., laminate lay-up configurations, and their mapping to the criterion space are generated from one design point. At each cycle, the generated sets depend on the choice of the *a priori* selected parameters. The decision-maker can thus control the evolution of the sequence generated by the optimization procedure toward a desired objective.

A reduced problem with two strongly conflicting criteria is first formulated. One criterion is chosen for the single criterion and others are combined linearly for the new bicriterion problem formulation. The Pareto optimal stacking sequence corresponding to the user defined parameters, i.e., initial allowable angles for layer orientations and weighting factors for the combined criteria, is determined through the iterative procedure. Next, laminate lay-up configurations based on the best-compromise stacking sequence with variation of the $\pm\theta$ layers are generated and maximal solutions corresponding to the Pareto optimal laminates determined with the constraint method.

2. STATEMENT OF THE PROBLEM

In this paper, a composite laminate design problem is considered as follows. Let constant ply properties be given. The objective is to find a symmetric and balanced laminate lay-up configuration with the minimum feasible total number of layers N such that laminate failure margins measured in terms of laminate initial failure (First Ply Failure) Reserve Factors $RF_i \in R$, $RF_i > 0$ due to $i=1, 2,\ldots, m$ loading conditions are maximized. The laminate thickness can only be multiples of the layer thickness.

Formally, we consider a multicriterion optimization problem in discrete form

$$\max_{x \in \Omega} f(x), \tag{1}$$

where the components $f_i : \Omega \to R, i = 1, 2, \ldots, m$ of the objective function are called criteria and they represent the design objectives by which the performance of the laminate is measured.

In the present work, we treat laminate initial failure reserve factors with respect to the different loading conditions as criteria and define the vector objective function as

$$f(x) = \{RF_1(x), RF_2(x), \ldots, RF_m(x)\}. \tag{2}$$

We employ the laminate initial failure analysis based on the laminate load response computed with the Classical Lamination Theory [4], and the constant and variable load approach [5], [6]. We solve laminate initial failure reserve factors by using a derivative-free line search method [7]. Since the failure criterion internal formulation has no influence on the solution procedure, the method is suitable to be used also with complicated failure criterion formulations.

Alternative laminate lay-up configurations are represented by zero-one design variables ($x_{ks} = 1$ if the allowable angle $\Theta(s)$=(0,90,+θ,–θ), s=1,2,3,4 occurs and $x_{ks} = 0$ if the allowable angle does not occur to the kth layer)

$$x = \{x_{ks} \mid x_{ks} \in \{0,1\}, \sum_s x_{ks} = 1, \sum_k (x_{k3} - x_{k4}) = 0, k = 1,2,\ldots,N/2, s = 1,2,3,4\} \tag{3}$$

that belong to the feasible set defined as

$$\Omega = \{x \mid g_i(x) = 1 - RF_i(x) \le 0, RF_i \in R, i = 1,2,\ldots,m\}. \tag{4}$$

The image of the feasible set in the criterion space, i.e., the attainable set, is defined by

$$\Lambda = \{z \in R^m \mid z = f(x), x \in \Omega\}. \tag{5}$$

Usually, there exists no unique solution which would maximize all m criteria simultaneously. For the definition of optimal solution of a multicriterion optimization problem we apply the optimality criterion introduced by Pareto [1], [2], [3].

3. THE REDUCED PROBLEM FORMULATION

Since the set of Pareto optima may be large, it is beneficial to formulate the multicriterion problem computationally as economically as possible. Instead of computing all Pareto optima, it is usually sufficient to determine only a relevant subset of Pareto optima. The more appropriately the subset of Pareto optima is determined, the easier it is for the decision-maker to deal with the results. For that reason, a reduced problem, where parameters are introduced to combine linearly some criteria, is introduced. The decision-maker can interact in computation by setting the parameters on the basis of his preferences and the information available on the past computations. The reduced problem is solved such that in nonconvex cases none of the Pareto optima are missed.

The objective is to formulate a bicriterion problem with strongly conflicting criteria. However, it might be difficult to give *a priori* knowledge on the relative importance of different criteria. The judgement can be made on the basis of the approximated criterion

sensitivities for varying the laminate lay-up configuration. For analyzing criterion sensitivities we apply the pay-off table introduced by Benayoun *et al.* [8].

An initial set of design points, i.e., symmetric and balanced laminate lay-up configurations with permutations on the allowable angles Θ for layer orientations is first generated. Failure margins to the given loading conditions are determined and results gathered into the pay-off table (Table 1). In the pay-off table, the first row corresponds to the vector maximizing the first criterion, the second row maximizing the second criterion, *et cetera.* Thus, the in general infeasible *ideal point* $\hat{f}$ maximizing all criteria simultaneously is on the diagonal of the pay-off table. The minimum value in each column of the pay-off table is the *nadir point* denoted by n_i.

Table 1. Pay-off table for m criteria.

x	$f_1(x)$	$f_2(x)$	...	$f_m(x)$
x^1	$RF_1(x^1)$	$RF_2(x^1)$	...	$RF_m(x^1)$
x^2	$RF_1(x^2)$	$RF_2(x^2)$	...	$RF_m(x^2)$
⋮	⋮	⋮	⋮	⋮
x^m	$RF_1(x^m)$	$RF_2(x^m)$	...	$RF_m(x^m)$

If the criterion value does not vary much from the best value for varying the laminate lay-up configuration, the corresponding criterion will not be sensitive to a variation in the parameter values that will be assigned to the criteria. To find the least sensitive criterion, normalization coefficients

$$\mu_i = \frac{\eta_i}{\sum_i \eta_i}, \text{ where } \eta_i = \frac{\hat{f}_i - n_i}{\hat{f}_i}, i = 1, 2, \ldots, m \tag{6}$$

are determined. A well-behaved formulation can usually be obtained by choosing the criterion producing minimum μ_i for the single criterion. Other criteria are combined linearly by introducing parameters called weighting factors. A new problem with a reduced vector objective function is defined as

$$\max_{x \in \Omega} \{f_1(x), f_2(x)\}, \tag{7}$$

$$\begin{aligned} f_1(x) &= RF_i(x), i = 1, 2, \ldots, m, \\ f_2(x) &= \sum_i \lambda_i RF_i(x), \lambda_i > 0, \sum_i \lambda_i = 1, i = 1, 2, \ldots, m-1, \end{aligned} \tag{8}$$

where the weighting factors $\lambda_i \in [0, 1]$ are arbitrarily chosen by the decision-maker.

Before grouping the criteria, the decision-maker should also compare the lay-up configurations of the ideal point and ensure that the most conflicting criteria will not be combined linearly. In some cases, the normalization coefficients of the most conflicting criteria are equal and combining these criteria would lead to a badly behaved problem formulation.

4. DETERMINING PARETO OPTIMAL SOLUTIONS

We are interested in finding Pareto optimal solutions through an iterative discrete procedure defined by a sequence $\{x^{(j)}\}, j=1,2,\ldots,n$. The state of a sequence $x^{(j)}$ corresponds to the selected design point, i.e., laminate lay-up configuration, at that cycle. The aim is to find a procedure converging toward a Pareto optimum, i.e., $x^{(j)} \to x^{*}$.

Let a positive orthant of the criterion space be defined by the convex cone

$$Z_{+}^{m} = \{f(x) \in R^{m} \mid f_i(x) > 0, i = 1,2,\ldots,m\} \tag{9}$$

with nonempty interior. In order to find a monotone path from an initial infeasible state to a Pareto optimal one, we seek the next state of the sequence through a point-to-set mapping $x^{(j)} \mapsto \varphi(x^{(j)})$

$$x^{(j+1)} \in \varphi(x^{(j)}) = \begin{cases} (x^{(j)},0,0),(x^{(j)},0,90),(x^{(j)},90,0),(x^{(j)},90,90), \\ (0,0,x^{(j)}),(0,90,x^{(j)}),(90,0,x^{(j)}),(90,90,x^{(j)}), \\ (x^{(j)},+\theta,-\theta),(x^{(j)},-\theta,+\theta),(+\theta,-\theta,x^{(j)}),(-\theta,+\theta,x^{(j)}) \end{cases} \tag{10}$$

where the layer orientations of the additional layers are for clarity denoted in deg. The generated set of design points is mapped into the reduced criterion space $\varphi(x^{(j)}) \mapsto f(\varphi(x^{(j)}),\lambda) \in Z_{+}^{2}$ and the ideal point

$$\hat{f}^{(j)} = \{\max_{x \in \varphi(x^{(j)})} f_1(x), \max_{x \in \varphi(x^{(j)})} f_2(x)\} \tag{11}$$

is determined. The nearest point $x \in \varphi(x^{(j)})$ minimizing the distance to the ideal point in the sense of Tchebycheff norm [11] (minimax approach) as

$$\min_{x \in \varphi(x^{(j)})} \| \hat{f}^{(j)} - f(x) \| = \min_{x \in \varphi(x^{(j)})} \max\{\hat{f}_1^{(j)} - f_1(x), \hat{f}_2^{(j)} - f_2(x)\} \tag{12}$$

is sought as the next state of the sequence $x^{(j+1)} \in \varphi(x^{(j)})$, such that in the reduced criterion space the condition

$$f(x^{(j+1)}) - f(x^{(j)}) \in Z_{+}^{2} \tag{13}$$

is satisfied. We thus require the monotone improvement of the criteria as the procedure goes on [9], [10]. The design point $x^{(J+1)}$ is selected by the procedure and depends on the choise of the weighting factors of the combined criteria.

We study the convergence of the procedure with respect to the metric

$$d(y, x^{(J)}) = \| f(y) - f(x^{(J)}) \| = \max_{i=1,\ldots,m} \{ f_i(y) - f_i(x^{(J)}) \}, \tag{14}$$

where $y \in \Omega$ and $f_i(y) = 1$, and

$$f_i(x^{(J)}) = \begin{cases} f_i(x^{(J)}), & 0 < f_i(x^{(J)}) \le 1 \\ 1, & f_i(x^{(J)}) > 1. \end{cases} \tag{15}$$

The termination condition is reached when the sequence $\{x^{(J)}\}$ generated by the optimization procedure has converged to the limit $d(y, x^{(J)}) \to 0$.

As the feasible lay-up configuration is reached through the iterative procedure, the number of layers is tried to reduce, if possible, one layer at a time such that the feasible laminate design is still maintained. Thus, for instance from a laminate having the sub-laminate (0/90)2 with 4 layers, laminates including the blocks (90/0/90), (0/0/90), (0/90/90), and (0/90/0) with 3 layers a block are generated. The $(\pm\theta)$ blocks are kept constant since the balanced laminate structure is required. Feasible laminate lay-up configurations are mapped into the reduced criterion space and a solution corresponding to the closest point to the ideal point is chosen for the best-compromise solution.

Finally, laminate lay-up configurations based on the best-compromise stacking sequence with variation of the $(\pm\theta)$ blocks are generated. Pareto optimal solutions are sought in the set $\theta \in \Theta^0 = \{0, 1, 2, \ldots, 90\}$ deg. Feasible laminate lay-up configurations are mapped into the reduced criterion space and maximal points determined with the constraint method [1], [2], [3]. The constraint method can generate the maximal points corresponding to a particular weighting factor also in nonconvex cases.

5. NUMERICAL EXAMPLE

To illustrate the optimization technique, we consider a laminate that is composed of layers having the mechanical properties of AS4 Carbon/epoxy ply (t=0.25 mm, E_1=126 GPa, E_2=11 GPa, G_{12}=6.6 GPa, ν_{12}=0.28, α_1=−1 10^{-6}/°C, α_2=26 10^{-6}/°C, X_t=1950 MPa, X_c=1480 MPa, S_{12}=79 MPa, Y_t=48 MPa, Y_c=200 MPa) and subjected to three conflicting loading conditions (Table 2).

The allowable angles for layer orientations are $\Theta = (0, 90, +\theta, -\theta)$ deg and the initial number of layers is $N/2$=4. The temperature difference between the constant stress-free temperature and the operating temperature is assumed to vary linearly through the

thickness of a laminate. Factors of safety for constant and variable loads are $FoS^c = 1.5$ and $FoS^v = 1.5$. Tsai-Hill failure criterion is used. Margin to failure is studied with respect to the increase of the variable load. The laminate is assumed to have no free edges, i.e., there are no interlaminar stresses which may cause failure.

Table 2. The three conflicting loading conditions subjected to the laminate.

	LC1	LC2	LC3
Constant	$\Delta T' = -100$ °C, $\Delta T^b = -80$ °C		
Variable	$N_x = 2$ MN/m $N_y = 0.37$ MN/m $N_{xy} = 0$ $M_x = -10$ Nm/m $M_y = 0$ $M_{xy} = 0$	$N_x = 0$ $N_y = 0$ $N_{xy} = 0.6$ MN/m $M_x = 0$ $M_y = 0$ $M_{xy} = 10$ Nm/m	$N_x = -1.4$ MN/m $N_y = -0.02$ MN/m $N_{xy} = 0$ $M_x = 0$ $M_y = 10$ Nm/m $M_{xy} = 0$

An initial set of symmetric and balanced laminate lay-up configurations with $N/2=4$ and permutations on the allowable angles $\Theta=(0, 90, +23, -23)$ deg for layer orientations is first generated. The number of stacking sequence permutations at this step is 70. On the basis of the pay-off table, RF_1 producing the minimum normalization coefficient is chosen for the single criterion and the problem with the reduced vector objective function is formulated as

$$\begin{aligned} &\max_{x\in\Omega}\{RF_1(x), \lambda_2 RF_2(x) + (1-\lambda_2)RF_3(x)\} \\ &\Omega = \{x \mid g_i(x) = 1 - RF_i(x) \le 0, RF_i \in R, i = 1,2,3\}. \end{aligned} \tag{16}$$

A solution to the design problem is found with $\lambda_2=0.74$ resulting after 5 cycles to the $[+\theta/-\theta/(-\theta/+\theta)5/+\theta/-\theta]$SE stacking sequence with $N=28$ (Figure 1).

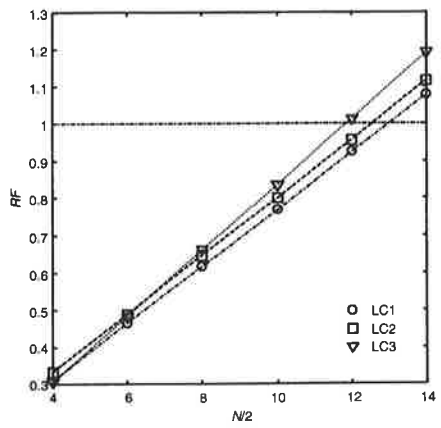


Figure 1. Improvement of the three criterion values as a function of the number of layers in a half laminate, $\Theta=(0, 90, +23, -23)$ deg and $\lambda_2=0.74$.

Improvement of the criterion values is monotone as illustrated in Figure 1. The iteration history in the reduced criterion space is shown in Figure 2.

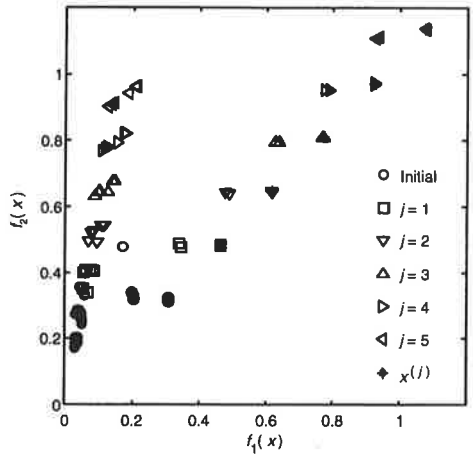


Figure 2. The iteration history in the reduced criterion space, $\Theta=(0, 90, +23, -23)$ deg and $\lambda_2=0.74$. Points corresponding to the state of the sequence $x^{(j)}$ are marked with an asterix.

Figure 3 shows the convergence of the procedure with respect to the used metric (14).

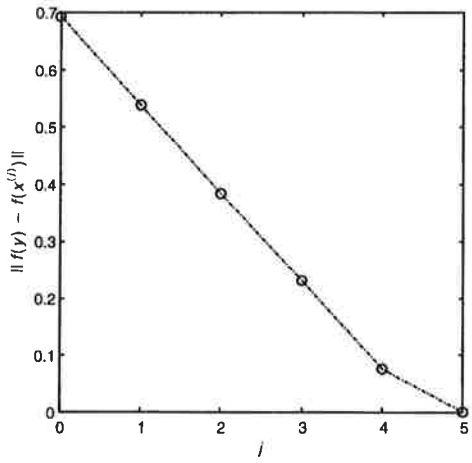


Figure 3. Convergence of the procedure with respect to the used metric.

Solution to the laminate design problem with different initial angles for layer orientations and $\lambda_2=0.74$ is presented in Table 3.

Table 3. Solution to the laminate design problem with different initial allowable angles for layer orientations, λ_2=0.74.

Θ / deg	Laminate stacking sequence	N
(0, 90, +22, −22)	[(−θ/+θ)6/+θ/−θ]SE	28
(0, 90, +23, −23)	[+θ/−θ/(−θ/+θ)5/+θ/−θ]SE	28
(0, 90, +24, −24)	[(+θ/−θ)2/(−θ/+θ)4/+θ/−θ]SE	28
(0, 90, +25, −25)	[(+θ/−θ)2/(−θ/+θ)5]SE	28

The maximal points corresponding to the laminate stacking sequences in Table 3 are shown in Figure 4.

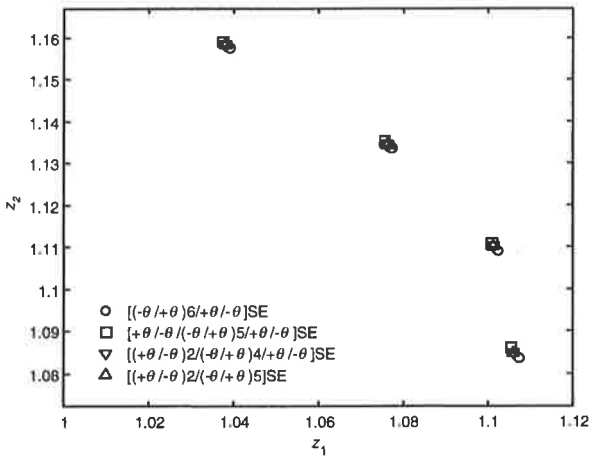


Figure 4. Maximal points corresponding to the laminate stacking sequences in Table 3. In each case, values of θ corresponding to the maximal points are 22, 23, 24, and 25 deg, λ_2=0.74.

In each case, the maximal points are reached with $\theta \in \Theta^* = \{22, 23, 24, 25\}$ deg. Criterion values and laminate failure margins with respect to the different loading conditions for the Pareto optimal laminates corresponding to the [+θ/−θ/(−θ/+θ)5/+θ/−θ]SE stacking sequence are given in Table 4.

Table 4. Pareto optimal laminates for the [+θ/−θ/(−θ/+θ)5/+θ/−θ]SE stacking sequence, λ_2=0.74.

θ / deg	$f_1(x)$	$f_2(x)$	$RF_1(x)$	$RF_2(x)$	$RF_3(x)$
22	1.0375	1.1590	1.0375	1.1096	1.2997
23	1.0758	1.1353	1.0758	1.1153	1.1922
24	1.1008	1.1112	1.1008	1.1160	1.0975
25	1.1055	1.0860	1.1055	1.1114	1.0135

6. CONCLUSIONS

Computing Pareto optima for composite laminates subjected to multiple loading conditions is considered. Laminate failure margins are treated as criteria and determined with an iterative line search method. The formulated bicriterion optimization scheme is solved through an interactive iterative procedure. The internal formulation of the used failure criterion has no effect on the solution procedure. The method is computationally efficient and enables the designer to obtain results in any accuracy that is relevant.

7. ACKNOWLEDGEMENTS

The work of the first author was supported by the Academy of Finland Grant 65375 and the Sanoma Oy foundation (Helsingin Sanomain 100-vuotissäätiö).

8. REFERENCES

1. Cohon J L. Multiobjective Programming and Planning. Mathematics in Science and Engineering, Vol. 140. New York: Academic Press, 1978.
2. Koski J, Silvennoinen R. Norm Methods and Partial Weighting in Multicriterion Optimization of Structures. Int. J. Numer. Methods Eng. 1987: 24: 1101-1121.
3. Koski J. Multicriterion Structural Optimization. In: Adeli H, editor. Advances in Design Optimization. London: Chapman & Hall, 1994.
4. Jones R M. Mechanics of Composite Materials. New York: Hemisphere Publishing Corporation, 1975.
5. Palanterä M, Klein M. Constant and Variable Loads in Failure Analyses of Composite Laminates. In: Computer Aided Design in Composite Material Technology IV (CADCOMP/94), Southampton, 1994. p. 221-228.
6. Palanterä M, Karjalainen J P. Failure Margins of Composite Laminates with Constant and Variable Load Approach. In: European Conference on Composite Materials (ECCM/8), Naples, 3-6 June 1998. p. 51-58.
7. Kere P, Palanterä M. A Method for Solving Margins of Safety in Composite Failure Analysis. In: Proceedings of the 6th Finnish Mechanics Days, Oulu, 1997. p. 187-197.
8. Benayoun R, de Montgolfier J, Tergny J, Larichev O. Linear Programming with Multiple Objective Functions: STEP Method (STEM). Mathl. Program. 1971: 1: p. 366-375.
9. Pascoletti A, Serafini P. An iterative Procedure for Vector Optimization. Journal of Mathematical Analysis and Applications 1982: 89: p. 95-106.
10. Serafini P. Reachability of Vector Optima through Dynamic Processes. In: French S, Hartley R, Thomas L C, White D J, editors. Multi-Objective Decision-Making. London: Academic Press, 1983.
11. Vanderbooten D, Vincke P. Description and Analysis of Some Representative Interactive Multicriteria Procedures. Mathl. Comput. Modelling 1989: 12: p. 1221-1238.

ANALYSIS AND OPTIMIZATION OF PLASTIC CONICAL SHELLS

J. LELLEP and E. PUMAN
Institute of Applied Mathematics,
Tartu University, 51014 Tartu, ESTONIA

ABSTRACT

Optimization of conical shells made from ideal rigid plastic materials is studied. The designs of minimum weight for given load carrying capacity and the designs of maximum load carrying capacity for given weight are established for shells of piece wise constant thickness.

1. INTRODUCTION

Conical shells have many applications in engineering. Due to the need for methods of analysis and design P. Hodge (1963) and his associates have studied limit analysis of conical shells made of a material which obeys the Tresca yield condition and associated flow law. Later R. H. Bryant, S. L. Lee and T. Mura (1969) established the safety factor of full conical shells made of a von Mises material.

It is somewhat surprising that there exists only a few papers on optimal design of plastic conical shells. J. Lellep and E. Puman (1994, 1999, 2000) developed methods on optimal design of conical shells made of perfectly plastic materials obeying piece wise linear yield conditions.

2. FORMULATION OF THE PROBLEM AND BASIC EQUATIONS

Let us consider a thin-walled conical shell subjected the uniformly distributed lateral pressure of intensity P. Assume that the shell is simply supported or clamped at the outer edge and absolutely free at the inner edge of radius a (Fig. 1).
The thickness of the shell wall is assumed to be piece wise constant, e.g.

$$h = h_j, \quad r \in (a_j, a_{j+1}), \tag{1}$$

where $j = 0, ..., n$ and $a_0 = a$, $a_{n+1} = R$.

We are looking for the design of the shell which weight (material volume) attains the

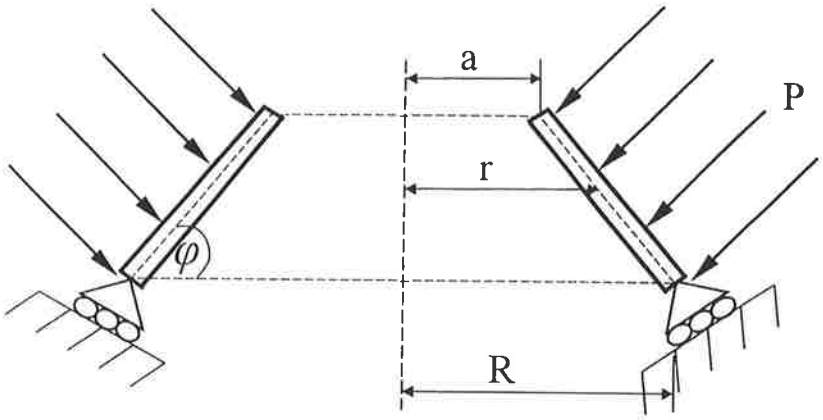


Figure 1: Conical shell subjected the uniformly distributed lateral pressure

minimum value for a given load carrying capacity. Evidently, the volume of the shell material may be presented as

$$V_0 = \sum_{j=0}^{n} h_j (a_{j+1}^2 - a_j^2) \cdot \frac{\pi}{\cos\varphi}, \tag{2}$$

where a_0, a_{n+1} and φ are considered as given constants.

In order to establish the minimum weight design of the shell one has to minimize (2) so that the equilibrium equations, the associated flow law with geometrical relations and plasticity conditions, are satisfied at each point of the shell.

Due to symmetry the stress state of the shell is defined by membrane forces N_1, N_2 and moments M_1, M_2. Equilibrium equations of a shell element may be presented as (Hodge, 1963)

$$\begin{cases} \dfrac{d}{dr}(rN_1) - N_2 = 0, \\ \dfrac{d}{dr}\left[\dfrac{d}{dr}(rM_1) - M_2\right] + N_2 \dfrac{\sin\varphi}{\cos^2\varphi} + \dfrac{Pr}{\cos^2\varphi} = 0. \end{cases} \tag{3}$$

The corresponding strain rate components are (Kuech et al., 1965)

$$\dot{\varepsilon}_1 = \frac{d\dot{U}}{dr}\cos\varphi, \qquad \dot{\varepsilon}_2 = \frac{1}{r}(\dot{U}\cos\varphi + \dot{W}\sin\varphi) \tag{4}$$

$$\dot{\kappa}_1 = -\frac{M_0}{N_0}\frac{d^2\dot{W}}{dr^2}\cos^2\varphi, \quad \dot{\kappa}_2 = -\frac{M_0}{N_0}\cdot\frac{1}{r}\frac{d\dot{W}}{dr}\cos^2\varphi.$$

In (4) $\dot{U}$ and $\dot{W}$ stand for the displacement rates in the normal and circumferential directions, respectively, whereas $M_0 = \sigma_0 h^2/4$, $N_0 = \sigma_0 h$, σ_0 being the yield stress of the material.

In the case of a distributed loading it is reasonable to introduce the following nondimensional quantities:

$$\begin{aligned} &\varrho = \frac{r}{R}, \qquad \alpha_j = \frac{a_j}{R}, \quad \gamma_j = \frac{h_j}{h_*}, \qquad v = \frac{h}{h_*}, \\ &w = \frac{W}{R}, \quad u = \frac{U}{R}, \quad n_{1,2} = \frac{N_{1,2}}{N_*}, \quad m_{1,2} = \frac{M_{1,2}}{M_*}, \\ &p = \frac{PR}{N_* \sin\varphi}, \quad k = \frac{M_* \cos^2\varphi}{RN_* \sin\varphi}. \end{aligned} \tag{5}$$

In (5), M_* and N_* stand for the limit moment and limit load for the reference shell of constant thickness h_*. Thus $M_* = \sigma_0 h_*^2/4$, $N_* = \sigma_0 h_*$.

In variables (5) the equilibrium equations (3) take the form

$$\begin{cases} (\varrho n_1)' - n_2 = 0, \\ k[(\varrho m_1)' - m_2]' + n_2 + p\varrho = 0, \end{cases} \tag{6}$$

where the primes denote differentiation with respect to ϱ.

Making use of (5) the strain rate components (4) may be presented as

$$\begin{aligned} &\dot{\varepsilon}_1 = \dot{u}'\cos\varphi, \qquad \dot{\varepsilon}_2 = \frac{1}{\varrho}(\dot{u}\cos\varphi + \dot{w}\sin\varphi), \\ &\dot{\kappa}_1 = -k_0\dot{w}''\sin\varphi, \quad \dot{\kappa}_2 = -\frac{k_0}{\varrho}\dot{w}'\sin\varphi, \end{aligned} \tag{7}$$

where

$$k_0 = \frac{M_0\cos^2\varphi}{RN_0\sin\varphi}. \tag{8}$$

Boundary conditions for stress components are given below. Since the inner edge is free one has

$$n_1(\alpha_0) = 0, \quad m_1(\alpha_0) = 0, \quad s(\alpha_0) = 0, \tag{9}$$

where s stands for the nondimensional shear force. Evidently,
$s = (m_1\varrho)' - m_2$. At the outer edge

$$m_1(1) = 0, \tag{10}$$

in the case of a simply supported shell and

$$m_1(1) = -\gamma_n^2, \tag{11}$$

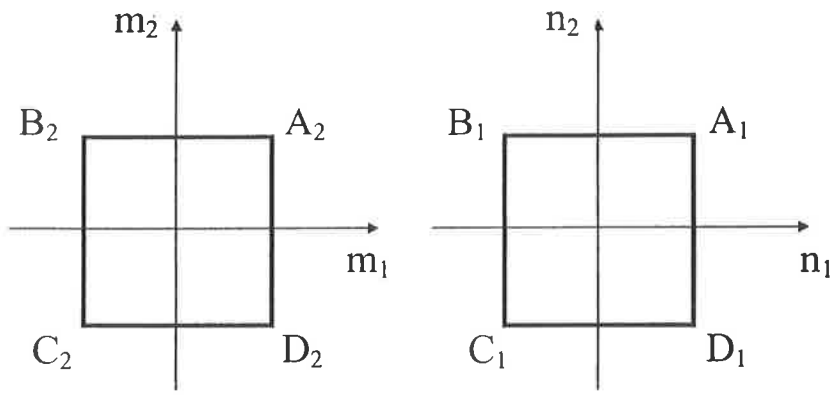


Figure 2: Generalized square yield condition

if the edge is clamped. In the both cases

$$\dot{w}(1) = \dot{u}(1) = 0. \tag{12}$$

In the case of a conical shell loaded by the rigid central boss (see J. Lellep and E. Puman, 2000) the equilibrium equations for a shell element have the form

$$\begin{cases} \dfrac{d}{dr}(rN_1) - N_2 = 0, \\ \dfrac{d}{dr}(rM_1) - M_2 - rN_1\dfrac{\sin\varphi}{\cos^2\varphi} + \dfrac{P}{2\pi\cos^2\varphi} = 0. \end{cases} \tag{13}$$

3. SHELLS OF PIECE WISE CONSTANT THICKNESS MADE OF A TRESCA MATERIAL

Let us consider conical shells subjected to uniformly distributed lateral loadings. Assume that the material of shells obeys the generalized square yield condition (Fig. 2) and associated flow law. In this case the governing equations are presented by (6) - (12).

It appears that the stress regime corresponds to the sides C_1D_1 and A_2B_2 of squares presented in Fig. 2.

Let $v = \gamma_j$ for $\varrho \in D_j$, where $D_j = (\alpha_j, \alpha_{j+1})$ for $j = 0, \ldots, n$. Thus

$$n_2 = -\gamma_j \tag{14}$$

and

$$m_2 = \gamma_j^2 \tag{15}$$

for $\varrho \in D_j$ $(j = 0, \ldots, n)$.
Equation (14) and the first equation in the set (6) give

$$(\varrho n_1)' = -\gamma_j, \qquad j = 0, \ldots, n. \tag{16}$$

Integrating (16) and satisfying (9) and the continuity condition of the quantity n_1 at $\varrho = \alpha_j$ $(j = 1, \ldots, n)$ leads to the relation

$$n_1 = \begin{cases} \gamma_0 \left(\dfrac{\alpha_0}{\varrho} - 1 \right), & \varrho \in (\alpha_0, \alpha_1), \\ \gamma_0 \left(\dfrac{\alpha_0}{\varrho} - \dfrac{\alpha_1}{\varrho} \right) + \displaystyle\sum_{i=1}^{j-1} \gamma_i \left(\frac{\alpha_i}{\varrho} - \frac{\alpha_{i+1}}{\varrho} \right) + \gamma_j \left(\frac{\alpha_j}{\varrho} - 1 \right), & \varrho \in (\alpha_j, \alpha_{j+1}), \end{cases} \tag{17}$$

for $\varrho \in D_j$, $j = 1, \ldots, n$. Inserting equations (14) and (15) in the second equation of the set (6) one obtains

$$k[(\varrho m_1)' - \gamma_j^2]' = \gamma_j - p\varrho, \tag{18}$$

for $\varrho \in D_j$, $j = 0, \ldots, n$.
Integrating (18) leads to the bending moment distribution

$$m_1' + \frac{1}{\varrho} m_1 = \frac{\gamma_j^2}{\varrho} + \frac{1}{k} \left(\gamma_j - \frac{p\varrho}{2} \right) + \frac{c_j}{k} \tag{19}$$

for $\varrho \in D_j$; where $D_j = (\alpha_j, \alpha_{j+1})$, $j = 0, \ldots, n$.
For determination of the constants of integration $c_0, \ldots, c_n$ one has the requirements $s(\alpha_0) = 0$ and $s(\alpha_j -) = s(\alpha_j +)$, $j = 1, \ldots, n$, which give

$$\begin{aligned} c_0 &= \frac{p}{2}\alpha_0^2 - \gamma_0 \alpha_0, \\ c_j &= \frac{p}{2}\alpha_0^2 - \gamma_0 \alpha_0 + \sum_{i=1}^{j} \alpha_i (\gamma_{i-1} - \gamma_i), \end{aligned} \tag{20}$$

for $j = 1, \ldots, n$. Subtituting equations (20) in (19) one obtains

$$m_1' + \frac{1}{\varrho} m_1 = \frac{\gamma_j^2}{\varrho} + \frac{1}{k} \left(\gamma_j + \frac{p}{2}(\alpha_0^2 - \varrho) - \gamma_0 \alpha_0 + \sum_{i=1}^{j} \alpha_i (\gamma_{i-1} - \gamma_i) \right), \tag{21}$$

for $\varrho \in D_j$, $j = 1, \ldots, n$. Integrating (21) once more

$$m_1 = \gamma_j^2 + \frac{\gamma_j}{k} \left(\frac{\varrho}{2} - \alpha_j \right) + \frac{p}{6k} (3\alpha_j^2 - \varrho^2) + \frac{1}{k} \sum_{i=0}^{j} \gamma_i (\alpha_{i+1} - \alpha_i) + \frac{C_j}{k\varrho} \tag{22}$$

and satisfying the boundary conditions $m_1(\alpha_0) = 0$ and continuity requirements imposed on the bending moment m_1 at $\varrho = \alpha_i$, we get the constants of integration

$$\begin{aligned} C_0 &= -\frac{p}{3}\alpha_0^3 + \frac{\gamma_0}{2}\alpha_0^2 - k\gamma_0^2\alpha_0, \\ C_j &= -\frac{p}{3}\alpha_0^3 + \sum_{i=0}^{j}\left\{\frac{1}{2}\gamma_i\left(\alpha_i^2 - \alpha_{i+1}^2\right) + k\gamma_i^2\left(\alpha_{i+1} - \alpha_i\right)\right\} + \\ &\quad + \frac{1}{2}\gamma_j\alpha_j\left(\alpha_j - 2k\gamma_j\right) \end{aligned} \tag{23}$$

for $j = 1, \ldots, n$.

Finally, one has to satisfy the boundary condition for bending moment at the outer edge of the shell. Evidently, $m_1(1) = 0$, in the case of a simply supported shell and $m_1(1) = \gamma_n^2$ in the case of a clamped shell. Making use of equations (22) and (23), we can present the load carrying capacity p of the clamped shell of piece wise constant thickness as

$$\begin{aligned} p &= \frac{3}{3\alpha_j^2 - 1 - 2\alpha_0^3}\Big(2k\gamma_n^2 + \gamma_j(2k\gamma_j + 1) + \gamma_j\alpha_j(\alpha_j - 2 - 2k\gamma_j) + \\ &+ 2\sum_{i=0}^{j}\left(\gamma_i(\alpha_{i+1} - \alpha_i)(k\gamma_i + 1) - \gamma_i(\alpha_{i+1}^2 - \alpha_i^2)\right)\Big) \end{aligned} \tag{24}$$

and the limit load for a simply supported shell as

$$\begin{aligned} p &= \frac{3}{3\alpha_j^2 - 1 - 2\alpha_0^3}\Big(\gamma_j(2k\gamma_j + 1) + \gamma_j\alpha_j(\alpha_j - 2 - 2k\gamma_j) + \\ &+ 2\sum_{i=0}^{j}\left(\gamma_i(\alpha_{i+1} - \alpha_i)(k\gamma_i + 1) - \gamma_i(\alpha_{i+1}^2 - \alpha_i^2)\right)\Big). \end{aligned} \tag{25}$$

4. SHELLS OF VON MISES MATERIAL

Assume that the material of shells obeys the yield condition presented in the form

$$n_1^2 - n_1n_2 + n_2^2 + m_1^2 - m_1m_2 + m_2^2 = 1 \tag{26}$$

in the case of the shell of constant thickness. For determination of the load carrying capacity of the shell we can use the lower bound theorem of limit analysis.

According to this theorem (see P. Hodge, 1963) the limit load factor could be defined as the maximal load factor corresponding to statically admissible stress fields, e.g. as $min(-p)$. In order to minimize the quantity $-p(\alpha_0)$ so that equilibrium equations are satisfied let us introduce the following augmented functional

$$\begin{aligned} J_* &= -p(\alpha_0) + \int_{\alpha_0}^{1} \left\{ \psi_1 \left(n_1' + \frac{n_1}{\varrho} - \frac{n_2}{\varrho} \right) + \psi_2 \left(m_1' + \frac{m_1}{2\varrho} \mp \right. \right. \\ &\mp \left. \frac{1}{\varrho} \sqrt{1 - \frac{3}{4} m_1^2 - n_1^2 - n_2^2 + n_1 n_2} + \frac{n_1}{k} - \frac{p}{2k\varrho} \left(\alpha_0^2 - \varrho^2 \right) \right) + \\ &+ \left. \psi_3 (p' - 0) \right\} d\varrho, \end{aligned} \tag{27}$$

where ψ_1, ψ_2 and ψ_3 are the adjoint variables.

Minimization of the functional J_* leads to the set of equations

$$\left\{ \begin{array}{l} n_1' = -\dfrac{n_1}{\varrho} + \dfrac{n_2}{\varrho}; \\ \\ m_1' = -\dfrac{m_1}{2\varrho} \pm \dfrac{\sqrt{1 - \frac{3}{4} m_1^2 - n_1^2 - n_2^2 + n_1 n_2}}{\varrho} - \dfrac{n_1}{k} + \dfrac{p}{2k\varrho} \left(\alpha_0^2 - \varrho^2 \right); \\ \\ p' = 0; \\ \\ \psi_1' = \dfrac{\psi_1}{\varrho} \mp \dfrac{\psi_2}{2\sqrt{1 - \frac{3}{4} m_1^2 - n_1^2 - n_2^2 + n_1 n_2 \varrho}} (n_2 - 2n_1) + \dfrac{\psi_2}{k}; \\ \\ \psi_2' = \dfrac{\psi_2}{2\varrho} \mp \dfrac{\psi_2}{2\sqrt{1 - \frac{3}{4} m_1^2 - n_1^2 - n_2^2 + n_1 n_2 \varrho}} \left(-\dfrac{3}{2} m_1 \right); \\ \\ \psi_3' = -\dfrac{\psi_2}{2k\varrho} \left(\alpha_0^2 - \varrho^2 \right), \end{array} \right. \tag{28}$$

which is to be integrated by the use of boundary conditions

$$\begin{aligned} &n_1(\alpha_0) = 0, \quad m_1(\alpha_0) = 0, \quad m_1(1) = 0, \\ &\psi_1(1) = 0, \quad \psi_3(\alpha_0) = 0, \quad \psi_3(1) = 1. \end{aligned} \tag{29}$$

In the similar way one can solve the problem of optimization.

α_0	p	α_1	γ_0	γ_1	V_*	V	e
0,1	4,846	0,8733	0,8802	1,1661	0,9900	0,9393	0,9488
0,2	4,554	0,8571	0,8631	1,1513	0,9600	0,9051	0,9428
0,3	4,477	0,8495	0,8346	1,1401	0,9100	0,8489	0,9329
0,4	4,630	0,8278	0,7917	1,1291	0,8400	0,7712	0,9181
0,5	5,100	0,8235	0,7346	1,1147	0,7500	0,6733	0,8977
0,6	6,136	0,8351	0,6713	1,0949	0,6400	0,5579	0,8717
0,7	8,472	0,8635	0,6131	1,0712	0,5100	0,4293	0,8417
0,8	15,000	0,9038	0,5653	1,0459	0,3600	0,2915	0,8098
0,9	48,210	0,9505	0,5278	1,0213	0,1900	0,1479	0,7784

Table 1: Optimal values of parameters for a Tresca material ($k = 0,2$)

5. NUMERICAL RESULTS

The results of calculations are presented in Tables 1-2 and Fig. 3-4. Table 1 corresponds to the shells with one step in the thickness made of a Tresca material whereas Table 2 is associated with a von Mises shell. In the last columns of Tables 1, 2 corresponding values of coefficients of efficiency are accomodated. Table 1 corresponds to a minimum weight problem where $e = V/V_*$, V and V_* being volumes of the optimized shell and the reference shell, respectively. In Table 2 $e = p/p_*$ where p is the maximal load for the stepped shell and p_* stands for the limit load of the reference shell of constant thickness. In Fig. 3 and 4 the bending moments m_1 and m_2 are depicted for shells of von Mises material. The shell wall has a step at the point where m_2 changes rapidly (Fig. 4).

6. CONCLUDING REMARKS

An optimization tehnique has been suggested for plastic conical shells subjected to the uniformly distributed lateral loading. Calculations carried out showed that maximally about 20% of the material can be saved when using the design of stepped shell with one step in the thickness. Eventual material saving depends on geometrical parameters of the shell and on the yield stress of the shell.

ACKNOWLEDGEMENT

The support of the Estonian Science Foundation through the Grant No 4377 is gratefully acknowledged.

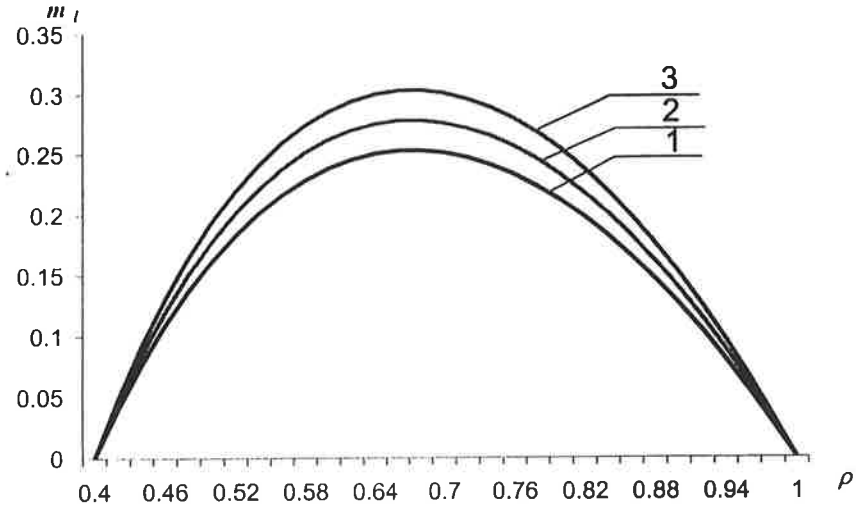


Figure 3: Bending moment distribution for the shell of von Mises material ($k = 0.9$ $\alpha_0 = 0.4$)

α_0	p	α_1	γ_0	γ_1	$n_1(1)$	V	e
0.4	2.0599	0.93099	1.05	0.73482	-0.53345	0.84	1.0363
0.4	2.1579	0.87568	1.10	0.73978	-0.55885	0.84	1.0857
0.4	2.2561	0.81598	1.15	0.77296	-0.58426	0.84	1.1350
0.4	2.3542	0.75699	1.20	0.80652	-0.60966	0.84	1.1844
0.5	1.8769	0.95485	1.05	0.62514	-0.42827	0.75	1.0394
0.5	1.9663	0.92045	1.10	0.60907	-0.44866	0.75	1.0889
0.5	2.0557	0.88714	1.15	0.62177	-0.46905	0.75	1.1384
0.5	2.1451	0.85348	1.20	0.64767	-0.48945	0.75	1.1879
0.6	1.7537	0.96695	1.05	0.55789	-0.32312	0.64	1.0422
0.6	1.8372	0.94225	1.10	0.52937	-0.33850	0.64	1.0918
0.6	1.9207	0.92354	1.15	0.49724	-0.35389	0.64	1.1414
0.6	2.0053	0.90472	1.20	0.49468	-0.36925	0.64	1.1917
0.7	1.7111	0.97634	1.05	0.50469	-0.22139	0.51	1.0449
0.7	1.7926	0.96296	1.10	0.39854	-0.23194	0.51	1.0947
0.7	1.8741	0.94959	1.15	0.37230	-0.24248	0.51	1.1445
0.7	1.9556	0.93489	1.20	0.39038	-0.25302	0.51	1.1942
0.8	1.8327	0.98455	1.05	0.46312	-0.12571	0.36	1.0476
0.8	1.9199	0.97682	1.10	0.31438	-0.13169	0.36	1.0975
0.8	2.0073	0.96868	1.15	0.27417	-0.13768	0.36	1.1471
0.8	2.0946	0.96090	1.20	0.26077	-0.14367	0.36	1.1973

Table 2: The shell of von Mises material with one step in the thickness ($k = 0.1$)

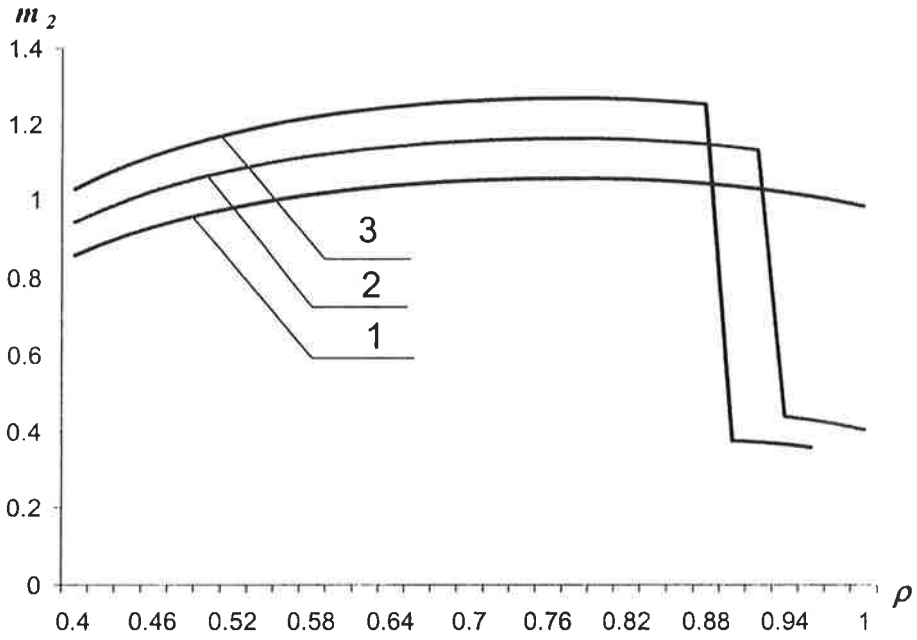


Figure 4: Bending moment m_2 distribution for the shell of von Mises material ($k = 0.9$, $\alpha_0 = 0.4$)

REFERENCES

1. Bryant, R. H.; Lee, S. L. and Mura, T., 1969. Mises limit load for simply supported conical sandwich shells. Ing. - Arch. 37, 5, 281-287.
2. Hodge, P. G., 1963. Limit Analysis of Rotationally Symmetric Plates and Shells. Englewood Cliffs: Prentice Hall.
3. Lellep, J. and Puman, E., 1994. Optimal design of rigid-plastic conical shells of piece-wise constant thickness. Tartu Ülik. Toim. 973, 21-39.
4. Lellep, J. and Puman, E., 1999. Optimization of plastic conical shells of piecewise constant thickness. Struct. Optim. 18, 74-79.
5. Lellep, J. and Puman, E., 2000. Optimization of plastic conical shells loaded by a rigid central boss. Int. J. Solids and Struct. 37 (17), 2695-2708.

ON AUTOMATIC DERIVATIVES IN SENSITIVITY ANALYSIS FOR SHAPE OPTIMIZATION PROBLEMS

R. A. E. MÄKINEN
Department of Mathematical Information Technology
University of Jyväskylä
P. O. Box 35
FIN-40351 Jyväskylä, FINLAND

ABSTRACT

A hybrid approach for shape design sensitivity analysis for a class of shape optimization is presented. Hand-coded derivatives and automatic derivatives are combined in such a way that the adjoint equation technique can be utilized. This approach yields significant reduction in the memory and time required to compute derivatives if compared to the approach where automatic differentiation is applied to the whole code. Numerical example is given.

1. INTRODUCTION

We consider the following abstract shape optimization problem in discrete form

$$\min \Phi(\boldsymbol{a}) = f(\boldsymbol{a}; \boldsymbol{q}(\boldsymbol{a})) \tag{1}$$

$$\text{subject to} \quad \boldsymbol{a} \in U^{ad} \subset \mathbb{R}^n \tag{2}$$

$$\boldsymbol{R}(\boldsymbol{a}; \boldsymbol{q}(\boldsymbol{a})) = 0. \tag{3}$$

Here $\boldsymbol{a}$ is the vector of design parameters defining the shape $\Omega(\boldsymbol{a})$ of the system. The state $\boldsymbol{q}(\boldsymbol{a})$ of the system is obtained by solving the state equation (3).

Gradient free global optimization methods like Genetic Algorithms are becoming more and more popular also in shape optimization [1], [2]. However, they cannot compete with gradient based methods like Sequential Quadratic Programming in efficiency when good initial guess is known and high accuracy is requested. Unfortunately, gradient based methods require the computation of the partial derivatives of the cost function with respect to design variables. Finite difference approximations for derivatives are often unreliable and are obtained too slowly. Therefore, obtaining analytic derivatives is an important step in the numerical optimization process.

By shape design sensitivity analysis we mean computing derivatives of the function Φ with respect to the geometrical parameter vector $\boldsymbol{a}$ defining the positions of nodal coordinates of the finite element mesh. Hand-coding of shape design sensitivity analysis has been considered extremely elaborate and difficult even for linear state problems.

Despite this, hand-coded sensitivity analysis can be done with a reasonable amount of work even in the case of quite complicated state problems [3], [4], [5]. However, fully hand-coded derivatives require a lot of man-hours and therefore automatic derivatives are now under active study [6], [7], [8].

2. ON AUTOMATIC DIFFERENTIATION OF COMPUTER PROGRAMS

Automatic differentiation (AD) is a technique for augmenting computer programs with derivative computations [9]. It exploits the fact that every computer program executes a sequence of elementary arithmetic operations. By applying the chain rule of derivative calculus repeatedly to these operations, accurate derivatives of arbitrary order can be computed automatically. Automatic differentiation has two basic modes the "forward" and the "reverse" modes. The running time and storage requirements of the forward mode are approximately proportional to the number of design variables. The reverse mode which is closely related to adjoint methods has a lower operations count for derivative computations, but potentially very large memory requirements.

Automatic differentiation can be implemented in two different ways: Existing analysis code written in e.g. Fortran 77 is precompiled using a precompiler (like ADIFOR [10]) into a new code that includes derivative calculations, or operator overloading technique available in e.g. Fortran 90 is used to produce sensitivity information. Modern programming languages such as Fortran 90, C++ make it possible to redefine the meaning of elementary operators. That is, we can define a new type for floating point numbers that has gradient information associated with it. For each elementary operation and standard function (`+`, `*`, `sin()`, `dot_product()`,...) we can define the meaning of the operation for variables of that new data type. The advantages of operator overloading is that it almost completely hides the AD tool from the user. If the implementation of the AD tool is changed the source code needs no modification.

AD techniques can be applied to add gradient computations to codes in their entirety. If forward mode is used, the computing time of one combined analysis and gradient evaluation is approximately n times the computing time of sheer analysis, i.e. almost the time needed for the forward difference approximation of the gradient. Moreover, the amount of memory used is also multiplied by n. Significant reduction in the memory and time required to compute derivatives is possible if some hand-coding is coupled with AD techniques.

We proceed with the following hybrid approach. First, we implement the forward mode of AD using operator overloading in Fortran 90. Secondly, we develop the geometric sensitivity analysis with respect to positions of nodal coordinates in general matrix form for a class of partial differential equations. As a by-product we also generate the Jacobian of $\boldsymbol{R}(\boldsymbol{a};\boldsymbol{q})$, needed when (3) is solved using Newton's method, using AD.

3. HYBRID APPROACH TO SENSITIVITY ANALYSIS

Many physical phenomena in solid and fluid mechanics can be modeled using the (set of) second order quasilinear partial differential equation(s) of type

$$\nabla \cdot \sigma(u) + g(u) = 0 \quad \text{in } \Omega(\boldsymbol{a}) \tag{4}$$

with suitable boundary conditions. For simplicity we assume that $u = 0$ on $\partial\Omega$ in what follows. After the finite element discretization the discrete analogue of (4) reads

$$\boldsymbol{R}(\boldsymbol{a};\boldsymbol{q}) \equiv \boldsymbol{K}(\boldsymbol{a};\boldsymbol{q})\,\boldsymbol{q} - \boldsymbol{g}(\boldsymbol{a};\boldsymbol{q}) = 0, \tag{5}$$

where $\boldsymbol{K}(\boldsymbol{a};\boldsymbol{q})$ and $\boldsymbol{g}(\boldsymbol{a};\boldsymbol{q})$ are the "stiffness" matrix and "force" vector, respectively.

Differentiating (5) with respect to a design variable a_k implicitly gives

$$\frac{\partial \boldsymbol{R}(\boldsymbol{a};\boldsymbol{q})}{\partial a_k} + \frac{\partial \boldsymbol{R}(\boldsymbol{a};\boldsymbol{q})}{\partial \boldsymbol{q}} \frac{\partial \boldsymbol{q}}{\partial a_k} = 0. \tag{6}$$

By introducing an adjoint state vector $\boldsymbol{p}$ we can eliminate the partial derivative of $\boldsymbol{q}$ from (6) and obtain the formulae needed for the evaluation of the gradient of the objective function $\Phi(\boldsymbol{a})$:

$$\left(\frac{\partial \boldsymbol{R}(\boldsymbol{a};\boldsymbol{q})}{\partial \boldsymbol{q}}\right)^{\mathrm{T}} \boldsymbol{p} = \nabla_{\boldsymbol{q}} f(\boldsymbol{a},\boldsymbol{q}) \tag{7}$$

$$\frac{\partial \Phi(\boldsymbol{a})}{\partial a_k} = \frac{\partial f(\boldsymbol{a},\boldsymbol{q})}{\partial a_k} - \boldsymbol{p}^{\mathrm{T}} \left(\frac{\partial \boldsymbol{R}}{\partial a_k}\right). \tag{8}$$

In [11], the case with hand-coded Jacobian $\partial \boldsymbol{R}/\partial \boldsymbol{q}$ and the right-hand side $\nabla_{\boldsymbol{q}} f$ was considered. The derivatives in (8) are more difficult due to the dependencies

$$\boldsymbol{a} \to \boldsymbol{X}(\boldsymbol{a}) \to \boldsymbol{R}(\cdot;\boldsymbol{q}),\ f(\cdot;\boldsymbol{q}) \tag{9}$$

which imply that one must first differentiate the nodal coordinate matrix $\boldsymbol{X}(\boldsymbol{a})$ with respect to design parameters. Hand-coding of this is error-prone and sometimes even impossible. In this work all derivatives in equations (7), (8) are computed using automatic differentiation.

We assume that every finite element has m degrees of freedom. The residual vector $\boldsymbol{R}(\boldsymbol{a};\boldsymbol{q}) \in \mathbb{R}^M$ is obtained using the standard assembly process

$$\boldsymbol{R}(\boldsymbol{a};\boldsymbol{q}) = \sum_{e}^{\text{ass.}} \boldsymbol{R}^e(\boldsymbol{a},\boldsymbol{q}^e), \tag{10}$$

where

$$\boldsymbol{R}^e(\boldsymbol{a};\boldsymbol{q}^e) = \boldsymbol{K}^e(\boldsymbol{a};\boldsymbol{q}^e)\boldsymbol{q}^e - \boldsymbol{g}^e(\boldsymbol{a};\boldsymbol{q}^e)$$

and $\boldsymbol{q}^e \in \mathbb{R}^m$ is the vector of element degrees of freedom. Moreover, we assume that the cost function is separable in the following sense:

$$f(\boldsymbol{a}; \boldsymbol{q}) = \sum_e f_e(\boldsymbol{a}; \boldsymbol{q}^e). \tag{11}$$

Now we can apply automatic differentiation to local contributions in (10), (11). Thus, we have to differentiate only "small" matrices and vectors with respect to "small" number of independent variables $a_1, ..., a_n, q_1^e, ..., q_m^e$. The global terms in (7), (8) are then obtained using the standard assembly process which contains no automatic differentiation.

In Table 1 a Fortran 90-style pseudo-code for calculating the value of the cost function Φ and its gradient at a given point $\boldsymbol{a}$ using the formulae

$$\begin{aligned} &\boldsymbol{J}^{\mathrm{T}}\boldsymbol{p} = \boldsymbol{h}, \\ &\frac{\partial \Phi(\boldsymbol{a})}{\partial a_k} = \frac{\partial f(\boldsymbol{a}; \boldsymbol{q})}{\partial a_k} - \boldsymbol{p}^{\mathrm{T}}\boldsymbol{s}^{(k)}, \quad k = 1, ..., n, \end{aligned}$$

is shown. Here $\boldsymbol{s}^{(k)} = \frac{\partial \boldsymbol{R}(\boldsymbol{a};\boldsymbol{q})}{\partial a_k}$, $\boldsymbol{J} = \frac{\partial \boldsymbol{R}(\boldsymbol{a};\boldsymbol{q})}{\partial \boldsymbol{q}}$, $\boldsymbol{h} = \nabla_{\boldsymbol{q}} f(\boldsymbol{a}; \boldsymbol{q})$.

We assume that there is m degrees of freedom per element. The Fortran 90-module `AD` contains the code needed to implement the forward mode of automatic differentiation using operator overloading. For a description of the technique (using C++), see [12]. A new data type `DVAR` containing derivative information is defined and the arithmetic operations and standard functions are overloaded for this new data type. Also mixed `DVAR/REAL` arithmetic is implemented. Moreover, module `AD` contains the functions `AD_indep_var`, `AD_value`, and `AD_pd`. Let `y0` and `y` be n-vectors of type `REAL` and `DVAR`, respectively. Then the call

```
y = AD_indep_var( y0 )
```

declares `y` as an independent variable vector with initial value `y0`. The calls

```
f0 = AD_value( f )
df = AD_pd( f, i )
```

return the `REAL` values of `f` and its partial derivative with respect to `y(i)`.

4. NUMERICAL EXAMPLE

We test our approach with the following simple shape optimization problem with nonlinear state problem:

$$\begin{aligned} \min \quad & \int_D |\nabla(u - u_d)|^2 \, dx \\ \text{subject to } & \boldsymbol{a} \in U^{ad} \\ & \nabla \cdot (\rho(u)\nabla u) = 0 \quad \text{in } \Omega(\boldsymbol{a}) \\ & u = u_0 \quad \text{on } \Gamma_0 \\ & \rho(u)\nabla u \cdot n = 0 \quad \text{on } \partial\Omega(\boldsymbol{a}) \setminus \Gamma_0 \end{aligned}$$

```
SUBROUTINE objfun( n, a, Phi, dPhi )
  USE AD
  INTEGER, INTENT(IN) :: n
  REAL,    INTENT(IN) :: a(n)
  REAL,    INTENT(OUT):: Phi, dPhi(n)
  !
  REAL:: Q(M), dQ(M), R(M), J(M,M), p(M), h(M), q_e(m), h_e(m),
         J_e(m,m), dRda_e(m), s(M,n)
  TYPE(DVAR):: X(Nnodes,2), R_e(m), f_e, z(n+m)
  !
  q_e = 0
  z = ad_indep_var( [a; q_e] )
  CALL generate_mesh( z, X )
  !
  ! Solution of the state equation:
  !
  Q = 0
  DO WHILE ( .not.converged )
     J = 0
     R = 0
     DO e=1,Nelems
        q_e = get_local_dofs( Q, e )
        z = ad_indep_var( [ a; q_e ] )
        CALL calc_elem_res( X, z, R_e )
        R = R   +   ad_value( R_e )
        DO j=1,m
           J_e(:,j) = ad_pd( R_e, n + j )
        END DO
        J = J   +   J_e
        DO k=1,n
           dRda_e = ad_pd( R_e, k )
           s(:,k) = s(:,k)   +   dRda_e
        END DO
     END DO
     CALL LinearSolve( J, dQ, -R )
     Q = Q + dQ
  END DO
  !
  ! Calculation of the cost function and adjoint r.h.s.:
  !
  Phi = 0;  h = 0; dPhi = 0
  DO e=1,Nelems
     q_e = get_local_dofs( Q, e )
     z = ad_indep_var( [ a; q_e ] )
     CALL calc_elem_cost( X, z, f_e )
     Phi = Phi + ad_value( f_e )
     DO k=1,n
        dPhi(k) = dPhi(k) + ad_pd( f_e, k )
     END DO
     DO j=1,m
        h_e(j) = ad_pd( f_e, n + j )
     END DO
     h = h   +   h_e
  END DO
  !
  ! Solve adjoint equation:
  !
  CALL LinearSolve( Transpose(J), p, h )
  !
  ! Calculate the gradient of the cost function:
  !
  DO k=1,n
     dPhi(k) = dPhi(k) - dot_product( p, s(:,k) )
  END DO
END SUBROUTINE objfun
```

Table 1: Subroutine for the calculation of the value and the gradient of Φ at $\boldsymbol{a}$

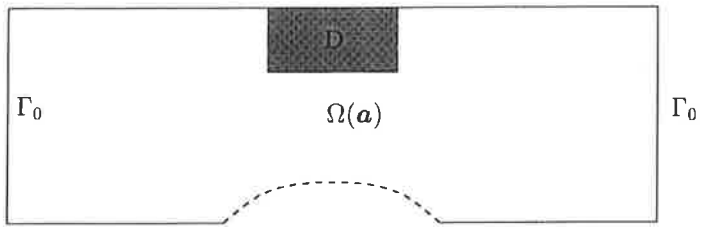


Figure 1: Problem geometry

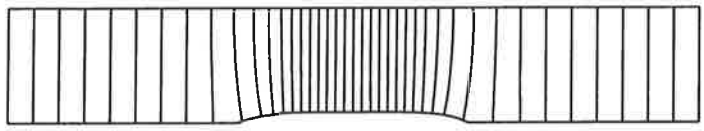


Figure 2: Final domain with equipotential contours

Here $\rho(u) = (1 + |\nabla u|^2)^{-1/2}$. The domain $\Omega(\boldsymbol{a})$, subdomain D, and parts of the boundary with Dirichlet boundary condition are depicted in Figure 1. The dashed part of the boundary is the object of the optimization. This problem can be considered as a simplification of an inverse design problem for a nozzle, for example.

The unknown part of the boundary was parameterized using a Bezier curve with six control points. Four of the control points are allowed to move between given "move limits". The state problem was discretized using 224 four-noded finite elements. Thus in this case $n + m = 4 + 4 = 8$. The correctness of the derivative computations was verified by comparing the AD derivatives of the cost function with (expensive) central difference approximations.

The initial domain, the fixed subdomain, and the functions u_0, u_d were chosen to be $\Omega(\boldsymbol{a}^0) =]0, 3[\times]0, \frac{1}{2}[$, $D =]\frac{1}{4}, \frac{3}{4}[\times]\frac{1}{4}, \frac{1}{2}[$, $u_0(x) = 3x_1$, $u_d(x) = 6x_1$, respectively. The final domain with selected equipotential contours of u is shown in Figure 2.

5. CONCLUSIONS

The proposed hybrid method for shape design sensitivity analysis is both easy to program and efficient in terms of computer time and memory. It is efficient as the differentiation of the (non)linear state solver is avoided making is possible to use standard software (LAPACK, for example) to solve the linearized state problem. Computed sensitivities are very accurate provided that the mesh topology remains fixed and the nonlinear state equation is solved with sufficiently strict stopping criterion. Our approach is general as it applies to multidisciplinary shape optimization problems and several finite elements. General purpose programs can be easily developed as the dependence on the specific application can be isolated into separate modules.

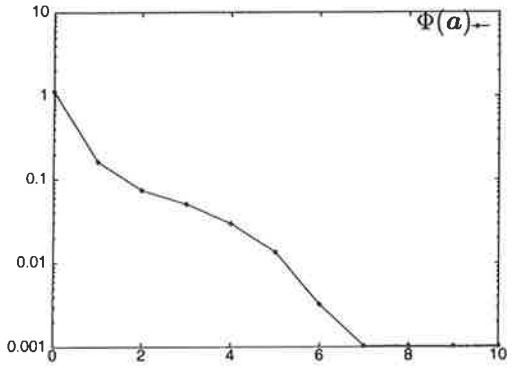


Figure 3: Cost vs. iterations

REFERENCES

[1] R. A. E. Mäkinen, J. Toivanen and J. Périaux, Multidisciplinary shape optimization in aerodynamics and electromagnetics using genetic algorithms, Int. J. Numer. Meth. Fluids, **30**, 149–159, 1999.

[2] M. Cerrolaza and W. Annicchiarico, Genetic Algorithms in Shape Optimization: Finite and Boundary Element Applications, Evolutionary Algorithms in Engineering and Computer Science, K. Miettinen, M. M. Mäkelä, P. Neittaanmäki, and J. Périaux (eds.), John Wiley & Sons, Chichester, 1999.

[3] J. Haslinger and R. A. E. Mäkinen, Shape optimization of elasto-plastic bodies under plane strains: sensitivity analysis and numerical implementation, Structural Optimization, **4**, 133–141, 1992.

[4] R. A. E. Mäkinen and J. Toivanen, Optimal shape design for Helmholtz/potential flow problem using fictitious domain method, Proceedings of 5th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, 529–536, 1994.

[5] P. Tarvainen, R. A. E. Mäkinen and J. Hämäläinen, Shape optimization for laminar and turbulent flows with applications to geometry design of paper machine headboxes, Proceedings of the Tenth International Conference on Finite Elements in Fluids, M. Hafez and J. C. Heinrich (eds.), The University of Arizona, 536–541, 1998.

[6] B. Mohammadi, J.-M. Malé, and N. Rostaing-Schmidt, Automatic Differentiation in Direct and Reverse Modes: Application to Optimum Shapes in Fluid Mechanics, Computational Differentiation: Techniques, Applications, and Tools, M. Berz, C. Bischof, G. Corlis, and A. Griewank (eds.), SIAM, 1996.

[7] I. Charpentier and M. Ghemires, Efficient adjoint derivatives: Application to the atmospheric model Meso-NH, Optimization Methods and Software, **13**, 35–63, 2000.

[8] C. Faure, P. Dutto, and S. Fidanova, Automatic differentiation and parallelism, Proceedings of The 3rd European Conference on Numerical Mathematics and Advanced Applications, Jyväskylä, Finland, July 26-30, 1999. World Scientific, 2000 (To appear).

[9] A. Griewank, On Automatic Differentiation, Mathematical Programming: Recent Developments and Applications, M. Iri and K. Tanabe (eds.), Kluwer Academic Publishers, Amsterdam, 83–108, 1989.

[10] C. Bischof, A. Carle, P. Khademi, and A. Mauer, ADIFOR 2.0: Automatic Differentiation of Fortran 77 Programs, IEEE Computational Science & Engineering, **3**, 18–32, 1996.

[11] R. A. E. Mäkinen, Combining automatic derivatives and hand-coded derivatives in sensitivity analysis for shape optimization problems, Short paper proceedings of 3rd World Congress of Structural and Multidisciplinary Optimization, C. L. Bloebaum, K. E. Lewis and R. W. Mayne (eds), State University of New York at Buffalo, 171–173, 1999.

[12] C. Bendtsen and O. Stauning, FADBAD, a flexible C++ package for automatic differentiation using the forward and backward methods, Technical report IMM-REP-1996-17, Technical University of Denmark, 1996.

KEHÄRAKENTEEN TOPOLOGIAN OPTIMOINNISTA

Timo Turkkila
Tampereen teknillinen korkeakoulu/ Teknillinen mekaniikka
PL 589
33101 Tampere

TIIVISTELMÄ

Tässä artikkelissa käsitellään kehärakenteen topologian ja mitoituksen optimointia. Optimoititehtävässä on kahdentyyppisiä suunnittelumuuttujia eli topologia- ja mitoitusmuuttujia. Topolgiamuuttujilla päätetään laitetaanko sauva, palkki vai poistetaanko kyseinen rakenneosa. Vastaavasti mitoitusmuuttujilla valitaan rakenteelle parhaat mahdolliset profiilit.

Optimointitehtävän ratkaisemiseksi on kehitetty kolme erilaista strategiaa. Relaksointimenetelmässä sekä topologia- että mitoitusmuuttujat relaksoidaan jatkuviksi, jonka jälkeen tehtävä ratkaistaan branch and bound algoritmilla. Hybridimenetelmässä topologiamuuttujien arvot ratkaistaan geneettisellä algoritmilla ja mitoitusmuuttujat relaksoidaan jatkuviksi. Suorassa menetelmässä sekä topologia- että mitoitusmuuttujat ratkaistaan geneettisellä algoritmilla.

JOHDANTO

Topologia määritellään geometriseksi ominaisuudeksi, jota ei voi muuttaa elastisilla muodonmuutoksilla rikkomatta kappaletta. Esimerkiksi nelikulmaisesta pinnasta voidaan muotoilla ympyrän tai kolmion muotoiset pinnat muttei rengasta. Kahdella eri ristikkorakenteella on sama topologia, jos ne voidaan muokata samanlaisiksi siten, ettei rakenteen nivelkiinnityksiä avata eikä rakenteeseen lisätä tai siitä poisteta sauvoja tai niveliä.

Ristikkorakenteiden topologiaa on optimoitu eri tekniikoilla jo kymmeniä vuosia ja aiheesta löytyy runsaasti artikkeleita ja kirjallisuutta. Asia on käsitelty jonkin verran esimerkiksi Kirschin [1] ja Bendsoen [2] oppikirjoissa. Geneettisen algoritmin soveltamista ristikoiden topologian optimointiin on esitelty esimerkiksi Hajelan ja Leen artikkelissa [3].

Tässä artikkelissa ei kuitenkaan käsitellä ristikoiden vaan tasokehien topologian optimointia. Kehärakenne muodostuu palkeista, joita voidaan kuormittaa sekä normaalivoimalla että taivuttamalla. Kehärakenteeseen voidaan asentaa myös niveliä, ja jos yksittäisen palkin molemmissa päissä on nivel, se käyttäytyy sauvana eli kantaa vain normaalivoimaa. Nivelen lisäys tai poisto muuttaa kehärakenteen topologian toiseksi.

Topologian optimointitehtävää yksinkertaistetaan muuntamalla se valintatehtäväksi. Tällöin suunnittelumuuttujilla päätetään, käytetäänkö rakenneosassa sauvaa vai palkkia. Jos valitaan sauva, sen molemmissa päissä on nivelet. Näin nivelpisteiden sijoittelu voidaan yksinkertaistaa palkkien ja sauvojen valinnoiksi. Samalla rakenteen sauvojen analysointi voidaan tehdä yksinkertaisemmalla nelivapausasteisella sauvaelementillä, kun palk-

kielementillä vapausasteita on kuusi. Sauvojen ja palkkien selkeä erottelu mahdollistaa myös eri profiilijoukkojen käytön. Niinpä esimerkiksi sauvat voidaan valita putkiprofiileista ja palkit I-profiileista.

Optimoinnin lähtökohtana on perusrakenne, jossa on n_s solmua ja n viivaa. Sanaa viiva on käytetty sekaannusten välttämiseksi ja sillä tarkoitetaan rakenneosaa, johon voidaan asentaa sauva tai palkki. Viiva voidaan myös poistaa rakenteesta, jolloin kyseisten solmujen väliin ei tule mitään. Kullekin perusrakenteen viivalle on kaksi diskreettiä suunnittelmuuttujaa: topologiamuuttuja b ja mitoitusmuuttuja s. Topologiamuuttujilla on kolme mahdollista arvoa, joista arvo $b = 0$ tarkoittaa kyseisen viivan poistoa, arvo $b = 1$ sauvaa ja arvo $b = 2$ palkkia. Mitoitusmuuttujien avulla etsitään rakenneosalle sopivin palkki- tai sauvaprofiili.

Kehärakenteesta on tehty joitakin yksinkertaistavia oletuksia:

- Tarkastellaan vain tasokehiä.
- Palkkielementti liittyy toiseen palkkielementtiin jäykällä liitoksella.
- Sauvaelementti liitetään nivelillä molemmista päistään.
- Oletustuentana on jäykkä tuenta palkkielementille ja niveltuenta sauvaelementille. Näitä tuentoja käyttäjä voi kuitenkin muuttaa.
- Rakennetta voidaan kuormittaa joko solmuihin kohdistuvilla pistevoimilla tai viivakuormituksilla ennalta valituille palkeille. Jos käytetään viivakuormitusta, kyseisen viivan topologiamuuttuja b saa arvon 2.

Rakenne analysoidaan elementtimenetelmällä. Jokaiselle rakenteen viivalle käytetään yhtä sauva- tai palkkielementtiä.

Optimointitehtävän kohdefuntio $f(\mathbf{x})$ voi olla esimerkiksi rakenteen massa, joku siirtymäkomponentti tai siirtymäkomponenttien painotettu summa, rakenteeseen sitoutunut kimmoenergia, rakenteen suurin normaalijännitys, alin ominaiskulmataajuus jne. Näitä samoja suureita voidaan käyttää myös rajoitusehtoina. Diskreetti optimointitehtävä on siis

$$\begin{aligned} &\min f(\mathbf{x}) \\ &\mathbf{g}(\mathbf{x}) \le \mathbf{0} \\ &\mathbf{x} = \begin{bmatrix} \mathbf{b} \\ \mathbf{s} \end{bmatrix} \\ &b_i \in \{0,1,2\} \qquad i = 1,2,\ldots,n \\ &s_i \in \{1,2,\ldots,n_p\} \; i = 1,2,\ldots,n, \end{aligned} \tag{1}$$

missä n_p on palkki- tai sauvaprofiilien lukumäärä ja i on rakenteen viivan numero.

OPTIMOINTIALGORIMEISTA

Tutkimuksessa käytetään branch and bound ja geneettistä algoritmia, jotka molemmat ratkaisevat diskreettejä optimointitehtäviä. Branch and bound algoritmia esitellään esimerkiksi Nemhauserin ja Wolseyn kirjassa [4]. Algoritmissa tarvitaan relaksoitu eli jatkuvaksi muunnettu optimointitehtävä. Aluksi relaksoitu tehtävä ratkaistaan ilman ylimääräisiä rajoitusehtoja. Jos tämän optimointitehtävän ratkaisun kaikilla suunnittelumuuttujilla ei ole diskreettiä arvoa, valitaan yksi suunnittelumuuttujista, jonka arvo on kahden diskreetin

arvon välissä. Tämän jälkeen muodostetaan kaksi uutta osatehtävää, joista toisessa kyseisen suunnittelumuuttujan arvon on oltava suurempi tai yhtäsuuri kuin ylempi diskreetti arvo ja toisessa pienempi tai yhtäsuuri kuin alempi diskreetti arvo. Nämä osatehtävät ratkaistaan ja, jos vieläkään ei saada kaikille muuttujille diskreettiä arvoa, muodostetaan uudet osatehtävät edellisen osatehtävän käyvästä alueesta. Prosessia jatketaan kunnes löydetään diskreetti ratkaisu, ei ole enää käypää aluetta tai osatehtävän kohdefunktion arvo on suurempi kuin parhaassa tunnetussa diskreetissä ratkaisussa.

Branch and bound algoritmi antaa globaalin minimin, jos relaksoidun tehtävän kohdefunktiossa ei ole lokaaleja minimejä ja käypä alue on konveksi. Yleensä rakenteiden optimointitehtävissä nämä ehdot eivät toteudu. Tällöin optimoinnin tulos saattaa riippua iteroinnin aloituspisteestä.

Geneettinen algoritmi perustuu luonnonvalinnan simulointiin ja siitä on varsin perusteellinen esitys esimerkiksi Michalewiczin kirjassa [5]. Algoritmissa suunnittelumuuttujat koodataan yksilön kromosomiksi, jolloin kukin yksilö vastaa yhtä suunnitteluavaruuden pistettä. Algoritmissa tutkitaan pisteiden joukkoa, ei yksittäistä pistettä. Pisteiden joukko ajatellaan eliöpopulaatioksi, jonka kehittymistä simuloidaan kymmenien, jopa satojen sukupolvien ajan.

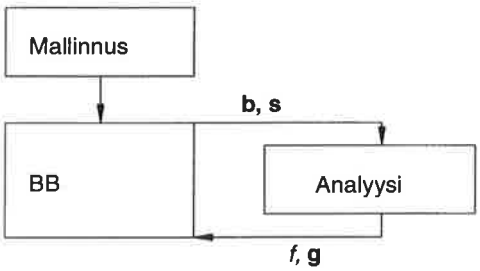
Uuden sukupolven luomiseen tarvitaan kolme proseduuria: vanhempien valinta, kromosomien risteytys ja geenimutaatio. Vanhemmat valitaan siten, että sukupolven parhailla yksilöillä on heikompia parempi mahdollisuus saada jälkeläisiä. Tässä yhteydessä on käytetty modifioitua järjestysvalintaa [6, s. 24], jossa kahden yksilön välisen paremmuusjärjestyksen määrää kohdefunktion arvo, jos kumpikin yksilö rikkoo käypää aluetta suunnilleen yhtä paljon. Jos toinen yksilö rikkoo selvästi enemmän, se häviää kohdefunktion arvosta riippumatta. Kun koko populaation yksilöiden keskinäinen järjestys on selvitetty, jälkeläisten saamistodennäköisyys riippuu paremmuusjärjestyksen sijaluvusta. Risteytyksessä vanhempien kromosomit katkaistaan ja loppuosat vaihdetaan keskenään, jolloin jälkeläisen kromosomin alkuosa on eri yksilöltä kuin loppuosa. Geenimutaatiossa jotkut kromosomin osat eli geenit saavat uuden arvon.

Koska geneettinen algoritmi parantaa pistejoukon keskimääräistä kohdefunktion arvoa, ei yksittäisen optimipisteen löytymisestä ole mitään takeita. Saatujen laskentakokemuksien perusteella geneettinen algoritmi antaa yleensä muutaman prosentin huonompia kohdefunktion arvoja kuin branch and bound algoritmi. Vaihtelu on kuitenkin varsin suurta ja joskus geneettinen algoritmi on selvästi branch and bound algoritmia parempi.

RATKAISUSTRATEGIAT

Relaksointimenetelmä

Kuvassa 1 on esitetty relaksointimenetelmän perusperiaate. Menetelmässä elementtimenetelmän laskentamalli pysyy koko ajan vakiona ja diskreetti optimointiongelma (1) ratkaistaan branch and bound algoritmilla. Opimointialgoritmi syöttää analyysimoduuliin suunnittelumuuttujavektorin ja analyysi palauttaa kyseissä pisteessä olevan kohdefunktion arvon sekä rajoitusehtovektorin.



Kuva 1. Relaksointimenetelmän periaatekaavio.

Koska branch and bound algoritmi tarvitsee jatkuvan optimointitehtävän, diskreetti optimointiongelma on relaksoitava jatkuvaksi. Mitoitusmuuttujat relaksoidaan interpolomalla kuutiosplinikäyrillä pintasuureiden eli poikkipinta-alan A, neliömomentin I ja taivutusvastuksen W arvoja. Tällöin esimerkiksi mitoitusmuuttujan arvo $s = 1{,}236$ tarkoittaa kuvitteelista profiilia, jonka pintasuureiden arvot ovat todellisten profiilien 1 ja 2 arvojen välissä. Interpoloitujen profiilien optimointi toimii luonnollisesti paremmin, jos todelliset profiilit voidaan järjestää siten, että niissä kaikkien pintasuureiden arvot ovat samassa suuruusjärjestyksessä. Pintasuureiden herkkyysanalyysi onnistuu derivoimalla splinikäyrän lauseketta kyseiseen viivaan liittyvän mitoitusmuuttujan s suhteen.

Topologiamuuttujan relaksointi on sitävastoin monimutkaisempi ongelma. Jokaiselle kohdefunktiossa tai rajoitusehdoissa olevalle suureelle on pystyttävä laskemaan osittaisderivaatat suunnittelumuuttujien suhteen kaikissa mahdollisissa suunnitteluavaruuden pisteissä. Lisäksi suunnittelumuuttujien ja rajoitusehtojen lukumäärän on oltava koko ajan vakio. Tämä ei ole aivan yksinkertaista, sillä pintasuureet ja koko rakenteen laskentamalli riippuvat topologiamuuttujien arvoista. Myös rakenteen mahdolliset mekanismit on pystyttävä hallitsemaan elgantisti ja poistettuihin viivoihin liittyvät rajoitusehdot eivät saa haitata optimointialgoritmia.

Topologiamuuttujat relaksoidaan käyttämällä laskentamallissa palkki- ja kiertojousielementtejä. Palkkielementit liitetään solmupisteissä nivelillä ja kiertojousilla toisiinssa. Jos jompikumpi kytkettävistä viivoista on sauva, kiertojousen jousivakio on nolla. Jos molemmat viivat ovat palkkeja, käytetään suurta jousivakion arvoa kuvaamaan jäykkää kiinnitystä. Poistettaville viivoille käytetään palkkielementtejä, joiden pintasuureiden arvot ovat hyvin pieniä todellisten profiilien arvoihin verrattuna. Tällöin viivan poisto ei voi tehdä rakenteesta mekanismia, mutta muuten näiden palkkien vaikutus rakenteen käyttäytymiseen on vähäinen.

Käytettävien palkkielementtien pintasuureiden arvot lasketaan sauva- ja palkkiprofiilien interpoloiduista käyristä edelleen interpoloimalla. Tällöin esimerkiksi poikkipinta-alalle A saadaan

$$A(b,s) = N_1(b)A_0 + N_2(b)A_s(s) + N_3(b)A_p(s), \qquad (2)$$

missä A_0 on poistettavan viivan pieni poikkipinta-ala, A_s sauvaprofiilien poikkipinta-alojen splinikäyrä, A_p palkkiprofiilien splinikäyrä ja N_1,N_2 sekä N_3 ovat kvadraattisia interpolointifunktioita

$$N_1 = (1-b)(2-b)/2$$
$$N_2 = b(2-b) \tag{3}$$
$$N_3 = b(b-1)/2.$$

Neliömomentin ja taivutusvastuksen arvot voidaan interpoloida samalla tavalla. Viivojen i ja j välissä olevan kiertojousen jousivakio lasketaan kaavalla

$$k_{ij} = a(b_i)a(b_j)k_0, \tag{4}$$

missä k_0 on suurehko jousivakio ja a funktio

$$a(b) = \begin{cases} 0, & 0 \le b < 1 \\ (b-1)^2, & 1 \le b \le 2. \end{cases} \tag{5}$$

Koska palkkielementin normaalijännitys lasketaan solmusiirtymien avulla, poistetulle viivalle voidaan saada näennäinen sallittua arvoa suurempi normaalijännitys. Rajoitusehtojen lukumäärää ei saa muuttaa ja kaikkien rajoitusehtojen on oltava derivoituvia kaikissa mahdollisissa suunnitteluavaruuden pisteissä. Tästä käyvän alueen rikkomisesta päästään eroon kirjoittamalla viivan i normaalijännitysrajoitusehto $g_j(\mathbf{x})$ muotoon

$$g_j(\mathbf{x}) = \begin{cases} b(2-b)(\sigma - \sigma_{sall}) \le 0, & 0 \le b < 1 \\ (\sigma - \sigma_{sall}) \le 0, & 1 \le b \le 2, \end{cases} \tag{6}$$

jolloin poistetulla viivalla rajoitusehdon arvo on aina 0.

Kaavojen (2) ... (6) avulla diskreetti optimointiongelma relaksoidaan jatkuvaksi. Optimointitehtävän herkkyysanalyysissä tarvittavat lausekkeet voidaan näistä kaavoista eteenpäin johtaa normaaliin tyyliin ketjuderivoinnilla käyttäen joko suoraa tai adjungoitua menetelmää.

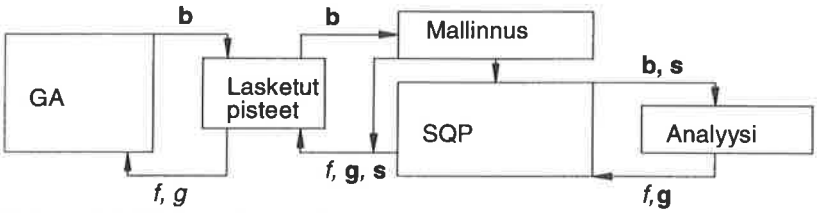
Hybridimenetelmä

Hybridimentelmässä topologian valinta ja rakenteen mitoitus erotetaan omiksi tehtävikseen. Topologiamuuttujavektori **b** ratkaistaan geneettisellä algoritmilla, jolloin topologian muutosta ei tarvitse relaksoida. Geneettinen algoritmi tarjoaa yhden topologian kerrallaan, jonka jälkeen siitä voidaan muodostaa laskentamalli ja ratkaista tälle topologialle optimaalinen mitoitusmuuttujavektori **s** jatkuvaksi relaksoidusta mitoitustehtävästä. Mitoitustehtävä on periaatteessa diskreetti ja se pitäisi ratkaista esimerkiksi branch and bound algoritmilla. Näin ei kuitenkaan kannata tehdä, sillä laskennan tarvitsema työmäärä kasvaa hyvin nopeasti kehärakenteen viivojen lukumäärän kasvaessa. Diskreetin ratkaisun sijaan tyydytään relaksoidun tehtävän ratkaisuun, joka lasketaan SQP-algoritmilla.

Hybridimenetelmässä ratkaistaan siis optimitopologia tapaukselle, jossa mitoitusmuuttujat on relaksoitu jatkuviksi eli jossa on käytettävissä ääretön määrä erilaisia profiileja. Optimitopologia ei välttämättä ole sama, jos käytetään jatkuvia tai diskreettejä mitoitusmuuttujia. Toisaalta geneettinen algoritmi ei läheskään aina löydä tarkkaa ratkaisua, joten

tässä tapauksessa ei käytännössä tehdä lopputuloksen kannalta kovin suurta virhettä. Kun geneettinen algoritmi on lopettanut laskentansa, saadulle optimitopologialle etsitään lopullinen diskreetti mitoitus käyttäen mitoitustehtävän ratkaisuun branch and bound algoritmia.

Topologia- ja mitoitustehtävien välissä on moduuli, joka tutkii onko kyseinen topologia jo aiemmin analysoitu. Näin voidaan jonkin verran pienentää kokonaistyömäärää. Hybridimenetelmän periaatekaavio on kuvassa 2.



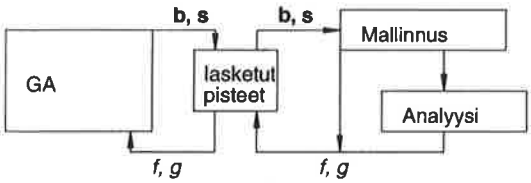
Kuva 2. Hybridimenetelmän periaatekaavio.

Mitoitustehtävän relaksoiminen ei ole kovin monimutkaista. Koska jokaiselle viivalle tiedetään, käytetäänkö siinä sauvaa, palkkia vai poistetaanko se, laskentamalli on helppomuodostaa. Ainoa asia, joka tarvitsee relaksointia, on profiilien pintasuureiden arvot. Ne saadaan kuten relaksointimenetelmässä interpoloimalla todellisten profiilien arvoja splinikäyrillä. Mitoitustehtävän suunnittelumuuttujien ja rajoitusehtojen lukumäärä riippuu tarkasteltavasta topologiasta. Tämä ei haittaa, sillä laskentaohjelmistossa käytettävä geneettisen algoritmin versio tarvitsee vain kohdefunktion ja pahiten rikkovan rajoitusehdon arvon. Jos ratkaisu on käypä, pahiten rikkovan rajoitusehdon arvo on 0.

Toinen huomiota vaativa seikka on rakenteeseen mahdollisesti syntyvät mekanismit. Ne etsitään tutkimalla jäykkyysmatriisin singulaarisuutta. Jos tarjottu topologia on mekanismi, rakennetta ei voi analysoida. Tällaiselle topologialle annetaan joku keskimääräinen kohdefunktion arvo ja suurehko pahiten rikkovan rajoitusehdon arvo, jolloin kyseinen rakenne joutuu populaation yksilöiden paremmuusjärjestyksessä hännille ja vähitellen häviää geneettisen algoritmin populaatiosta.

Suora menetelmä

Suorassa menetelmässä sekä topologia- että mitoitusmuuttujat ratkaistaan geneettisellä algoritmilla. Menetelmän periaate on esitetty kuvassa 3.



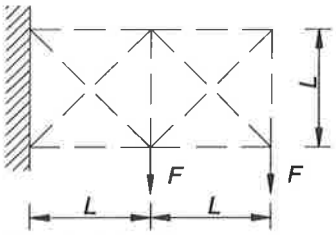
Kuva 3. Suoran menetelmän periaatekuva

Geneettinen algoritmi lähettää suunnittelumuuttujien arvot analyysiohjelmalle, joka muodostaa laskentamallin, ratkaisee sen ja palauttaa tuloksena kohdefunktion sekä pahiten rikkovan rajoitusehdon arvot. Menetelmässä ei tarvita mitään relaksointeja, joten se on

yksinkertaisin soveltaa. Kuten hybridimenetelmässäkin ennen analysoinnin aloittamista tutkitaan, onko kyseinen rakenne jo aiemmin analysoitu. Mahdolliset mekanismit käsitellään antamalla keskimääräinen kohdefunktion arvo ja suurehko käyvän alueen rikkominen.

ESIMERKKI

Testiesimerkkinä käytetään kehää, jonka kymmenviivainen perusrakenne on esitetty kuvassa 4. Rakenteen vasemmassa reunassa on palkkielementeille jäykkä ja sauvaelementille niveltuki. Samankaltaisen ristikkorakenteen, tosin hieman eri mitoilla ja kuormituksilla sekä jatkuvilla poikkipinta-aloilla, topologiaa on optimoitu useissa eri artikkeleissa.



Kuva 4. Esimerkin perusrakenne.

Kuvassa 4 oleva pituusmitta L = 3 m ja voima F = 10 kN. Valittavissa on viisi erilaista sauva- ja palkkiprofiilia. Tällöin optimointitehtävällä on yhteensä noin $5{,}77 \cdot 10^{11}$ erilaista vaihtoehtoa. Laskettaessa erilaisten rakennevaihtoehtojen lukumäärää jätetään poistettujen viivojen profiilit huomioimatta, jolloin päädytään noin $25{,}9 \cdot 10^{9}$ vaihtoehtoon.

Esimerkissä sauvaprofiilit ovat neliöputkia ja palkkiprofiilit IPE-palkkeja. Niiden pintasuureiden arvot on esitetty taulukossa 1.

Taulukko 1. Käytettävien profiilien pintasuureiden arvot

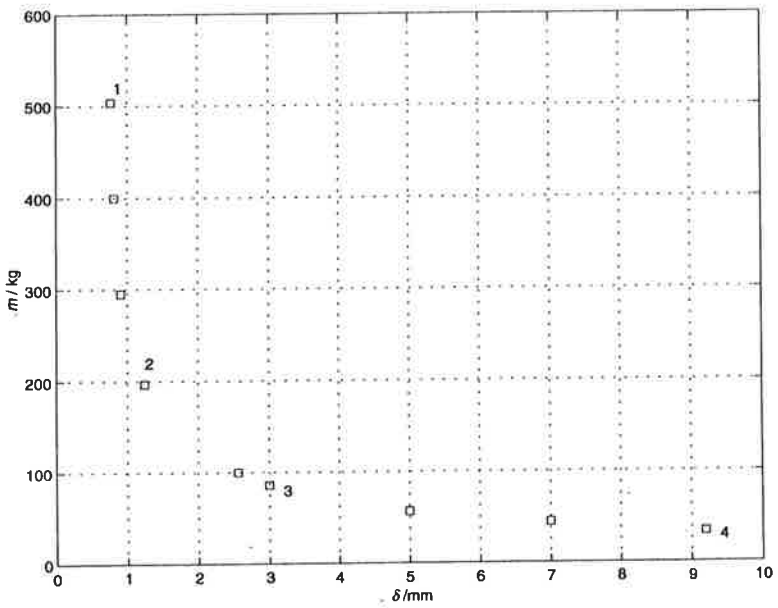
	Profiili	A / mm^2	I / 10^6 mm^4	W / 10^3 mm^3
1	20 x 20 / 2,0	142	0,0076	0,76
2	30 x 30 / 3,2	338	0,04	2,67
3	40 x 40 / 4,0	568	0,121	6,07
4	50 x 50 / 4,0	728	0,255	10,2
5	60 x 60 / 5,0	1090	0,544	18,1
1	IPE 80	764	0,801	20,0
2	IPE 100	1030	1,71	34,2
3	IPE 120	1320	3,18	53,0
4	IPE 140	1640	5,41	77,3
5	IPE 160	2010	8,69	109

Optimointitehtävässä minimoidaan rakenteen massaa m ja oikean alanurkan pystysiirtymää δ. Rajoitusehtona on, ettei minkään sauvan tai palkin normaalijännitys ylitä sallittua jännitystä σ_{sall} = 140 MPa. Optimointitehtävä on siis

$$\begin{aligned} &\min\left[m\ \delta\right]^{\mathrm{T}} \\ &\max\sigma \le \sigma_{\mathrm{sall}} \\ &b_i \in \{0,1,2\}\ i = 1,2,\ldots,10 \\ &s_i \in \{0,1,\ldots,5\}\ i = 1,2,\ldots,10. \end{aligned} \tag{7}$$

Optimointitehtävä (7) on monitavoitteinen, joten optimiratkaisut ovat Pareto-optimeja. Niiden määritelmä verbaalisti esitettynä on: "Käypä ratkaisu **x*** on Pareto-optimi, jos ei löydy toista käypää ratkaisua **x**, joka pienentäisi jonkun kriteerin arvoa aiheuttamatta samanaikaisesti vähintään yhden muun kriteerin arvon suurenemisen."

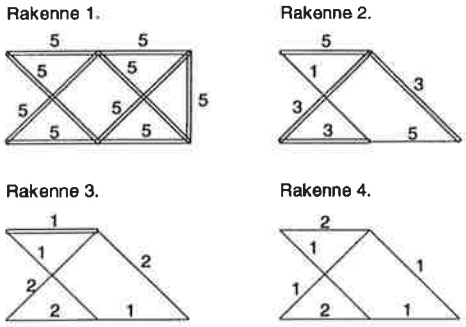
Monitavoitteisen optimointitehtävän tulokset voidaan esittää myös kriteeriavaruudessa. Tehtävän käypä alue kuvautuu sinne kuvajoukoksi ja Pareto-optimeja kutsutaan tässä avaruudessa minimaalisiksi pisteiksi. Kaksikriteerisen optimointitehtävän (7) tulokset voidaan siis esittää kompaktisti massa-siirtymä -koordinaatistossa. Joitakin laskettuja minimaalisia pisteitä on merkitty kuvaan 5 pienillä neliöillä.



Kuva 5. Minimaalisia pisteitä

Kuvan 5 minimaalistet pisteet on ratkaistu käyttäen rajoitusmenetelmää. Siinä yksi kriteereistä valitaan kohdefunktioksi ja loput parametrisoitaviksi rajoitusehdoiksi. Neljä kuvan pisteistä on numeroitu ja niitä esitellään tarkemmin. Piste 1 on siirtymän minimi 0,77 mm ja piste 4 massan minimi 33,5 kg. Siirtymän minimipisteen massa on 504 kg ja massan minimipisteessä siirtymä on 9,2 mm. Pisteet 2 ja 3 ovat niin sanottuja kompromissiratkaisuja. Pisteen 2 massa on 197 kg ja siirtymä 1,24 mm. Vastaavasti pisteen 3 massa on 86 kg

ja siirtymä 3,0 mm. Näitä neljää minimaalista pistettä vastaavat rakenteet on esitetty kuvassa 6. Tässä kuvassa palkit on esitetty kaksoisviivoilla ja sauvat yksinkertaisilla viivoilla. Viivan (tai kaksoisviivan) yläpuolella oleva luku tarkoittaa taulukossa 1 olevaa profiilin järjestysnumeroa. Tällöin esimerkiksi rakenteessa 1 kunkin kaksoisviivan päällä oleva luku 5 tarkoittaa järeintä palkkiprofiilia IPE 160. Rakenteiden vasemman reunan tuenta on palkkielementeillä jäykkä ja sauvaelementeillä niveltuenta.



Kuva 6. Pareto-optimaalisia rakenteita.

Näitä minimipisteitä on ratkottu kaikilla kolmella strategialla. Hybridimenetelmässä geneettisen algoritmin populaation suuruus oli 30 ja sukupolvien lukumäärä 80. Suorassa menetelmässä vastaavat luvut olivat 60 ja 600. Hybridi- ja suorassa menetelmässä kukin piste laskettiin kahteen kertaan. Pisteen 1 laskemiseen relaksointimenetelmä tarvitsi 30 rakenneanalyysiä, hybridimenetelmä noin 8000 analyysiä ja suora menetelmä noin 23000 analyysiä. Kaikki menetelmät päätyivät samaan rakenteeseen.

Pisteen 2 laskemiseen relaksointimenetelmä tarvitsi noin 3800 analyysia, hybridimenetelmä 5700 ja 12300 analyysiä ja suora menetelmä noin 22000 analyysiä. Relaksointi- ja hybridimenetelmä päätyivät samaan tulokseen. Suoralla menetelmällä saatiin aavistuksen huonompi tulos. Kolmannessa pisteessä relaksointimenetelmä ei löytänyt ratkaisua. Kuvan 6 ratkaisu saatiin hybridimenetelmällä. Molemmista yrityksistä saatiin sama tulos ja analyysejä tarvittiin 4300 ja 6300. Suora menetelmä tarvitsi noin 21000 analyysiä ja saatu tulos oli 700 g painavampi.

Piste 4 eli massan minimi osoittautui kaikille menetelmille vaikeaksi. Relaksointimenetelmä ei löytänyt ratkaisua. Hybridimenetelmällä saatiin kuvassa 6 esitetty rakenne (m = 33,5 kg), mutta toisen analyysin tulos oli huomattavasti painavampi rakenne (m = 41,5 kg). Analyysejä tarvittiin 3000 ja 2700. Suoralla menetelmällä päädyttiin rakenteisiin, joiden massat ovat 34,0 kg ja 40,7 kg. Näiden laskemiseen tarvittiin 21000 ja 22000 analyysiä.

Myös muissa kuvan 5 pisteissä, joita ei tässä yhteydessä esitellä tarkemmin, on esiintynyt samankaltaista käytöstä eri menetelmien välillä. Tässä esimerkissä relaksointimenetelmä näyttäisi olevan kilpailukykyinen, silloin kun rakenne on varsin jäykkä ja minimoidaan siirtymää massarajoituksella. Kun kohdefunktio vaihdetaan massaksi, relaksointimenetelmä joutuu vaikeuksiin. Kannattaa huomata, että minimoitaessa massaa ilman siirtymärajoituksia myös tyhjä rakenne, josta kaikki viivat on poistettu, on rajoitusehtojen puolesta käypä. Tämä ei toivottu tilanne voidaan kiertää käyttämällä siirtymärajoitusehtoa, joka branch and bound algoritmin osatehtävien lopputuloksissa ei ole aktiivinen. Valitettavasti

idea ei toimi kovin hyvin, sillä osatehtävien ratkaisemiseen käytettävä SQP-algoritmi ajautuu usein lisätystä rajoitusehdosta huolimatta tähän tyhjään ratkaisuun osaamatta sieltä enää takaisin käyvälle alueelle.

Suoralla menetelmällä rakenneanalyysien lukumäärässä ei ole ollut kovin suurta vaihtelua. Geneettisessä algoritmissa analyysien maksimimäärä on sukupolven koon ja sukupolvien lukumäärän tulo. Branch and bound algoritmissa ei tällaista maksimimäärää voi ainakaan näin helposti määrittää. Niinpä relaksointi- ja hybridimenetelmässä rakenneanalyysien lukumäärä vaihtelee melkoisesti.

Pelkkiä rakenneanalyysien lukumääriä vertailemalla saa helposti kuvan, että suoran menetelmän työmäärä olisi moninkertainen muihin menetelmiin verrattuna. Asia ei kuitenkaan ole näin yksinkertainen. Geneettinen algoritmi on hyvin yksinkertainen, joten rakenneanalyyseihin kulutettu laskenta-aika on likimain sama kuin koko optimointiin kulutettu laskenta-aika. Sitä vastoin SQP-algoritmissa tarvitaan matriisilaskentaa uuden hakusuunnan etsimiseen. Lisäksi rakenteelle on tehtävä herkkyysanalyysi, joten rakenneanlyyseihin käytetty laskenta-aika on vain osa kokonaisajasta. CPU-aikoja ei tässä yhteydessä ole tutkittu, mutta yleensä suora menetelmä on ollut laskennan nopein. Menetelmien nopeusjärjestys ei välttämättä vaihdu, vaikka tehtävän kokoa suurennettaisiin, sillä samalla lisääntyvät myös suunnittelumuuttujien ja rajoitusehtojen lukumäärät. Tätä kautta myös branch and bound algoritmin päätäntäpuu kasvaa, jolloin tarvitaan enemmän osatehtäviä ja enemmän anlyysejä. Myös SQP-algoritmin ja herkkyysanalyysin työmäärä kasvaa, joten menetelmien nopeusero todennäköisesti vain suurenee entisestään.

LÄHTEET

1. Kirsch, U., Structural Optimization: Fundamentals and Applications, Berlin 1993, Springer-Verlag, 302 s.

2. Bendsoe, M.P., Optimization of Structural Topology, Shape, and Material, Berlin 1995, Springer-Verlag, 271 s.

3. Hajela, P. and Lee, E., Genetic Algorithms in Truss Topological Optimization, Int. J. Solids and Structures, Vol 32 (1995), No. 22, ss. 3341 – 3357.

4. Nemhauser, G.L. and Wolsey L.A., Integer and Combinatorial Optimization, New York 1988, John Wiley & Sons, 763 s.

5. Michalewicz, Z., Genetic Algorithms + Data Structures = Evolution Programs, 2. painos, Berlin 1994, 340 s.

6. Turkkila, T. Kantavien rakenteiden diskreetti optimointi, Lisensiaatintutkimus, Tampere 1997, Tampereen teknillinen korkeakoulu, Konetekniikan osasto, 92 s.

MUODONMUUTOSNOPEUDEN VAIKUTUS ASE- JA AMMUS-TERÄSTEN MYÖTÖLUJUUTEEN

T. Erkkilä & S. Moilanen
Asetekniikkaosasto
Puolustusvoimien Teknillinen Tutkimuslaitos
PL 5
34 111 LAKIALA

TIIVISTELMÄ

Aseiden ja ampumatarvikkeiden lujuusopillisessa mitoituksessa voidaan hyödyntää laukausrasituksista aiheutuvan muodonmuutosnopeuden tuoma teräksen lujuuden mahdollinen lisäys. Edellytyksenä on, että rakenteen muodonmuutosnopeudet tunnetaan riittävällä tarkkuudella joko laskennallisesti tai mittauksellisesti. Mahdollinen materiaalin lujuuden kasvu voidaan määrittää kokeellisesti keskinopeilla dynaamisilla vetokokeilla. Lujalla nuorrutetulla aseteräksellä tehdyissä vetokokeissa saavutettiin ~5 % suhteellinen venymisrajan $R_{p0,2}$ lisäys, kun dynaamisen aineenkoetuksen venymänopeus oli ~10^4 -kertainen kvasistaattisen vetokokeen venymänopeuteen verrattuna. Kokeiden hajonta oli verrattavissa staattisten vetokokeiden ja kirjallisuudessa julkaistujen dynaamisten vetokokeiden tuloksiin. Niukkaseosteisella nuorrutetulla kranaattiteräksellä tehdyissä vetokokeissa myötö- ja murtolujuuksien keskimääräinen kasvu oli (7...9) %, mutta koetulosten hajonta oli lähes yhtä suuri, joten kranaattiteräksen osalta kokeita on jatkettava tilastollisen luotettavuuden saavuttamiseksi.

1. JOHDANTO

Aseen tuliputken lujuusopillinen mitoitus perustuu tavallisesti materiaalin myötölujuuteen, jonka ylitystä ei sallita normaaleilla laukauskuormituksilla. Materiaalin väsymiseen tai murtumien etenemiseen liittyvät ilmiöt eivät yleensä muodostu ongelmaksi, koska suurikaliiperisilla aseilla putkikohtaiset laukausmäärät ovat pieniä ja tuliputket ehtivät vaurioitua käyttökelvottomiksi kulumalla.

Tuliputken kannalta äärikuormitusolosuhteet esiintyvät tyypillisesti ampumatarvikkeiden epätavallisesta toiminnasta johtuvina ylikuormina. Ammus saattaa olla normaalia painavampi, ruutipanos ylisuuri, kuumentunut tai vaurioitunut, jolloin ruutikaasun paine ylittää ampumatarvikeyhdistelmälle suunnitellun painetason. Tällöin staattisen mitoituksen varmuuskertoimesta syntyvä pieni lisäkuormituksen mahdollisuus on nopeasti kulutettu ja putken vaurioitumiseen johtava painetaso saattaa ylittyä.

Ammukseen kohdistuvat kuormitukset ovat kertaluonteisia, joten sen materiaalin myötäminen ja pysyvät muodonmuutokset voidaan sallia, kunhan aseen ja ammuksen toiminnallisia rajoituksia ei loukata. Tyypillinen toimintarajoitus putkiaikana on, että ammuskuoren kosketusta tuliputken seinämään ei sallita muutoin kuin ohjauspintojen alueella, jotta putken "turha" kuluminen vältetään. Kun ammuskuoren materiaalin plastinen muodonmuutoskyky hyödynnetään kuoren lujuusopillisessa mitoituksessa, myötölujuuden ylityksen jälkeiset pienet kuormituslisät synnyttävät suuria muodonmuutoksia ja kuoren putkikosketuksen vaara kasvaa.

Suurikaliiperisen tykin laukauksessa ruutikaasun painepulssin kestoaika on t=(10...30) ms ja maksimipaine p_{max}= (200...600) MPa saavutetaan tyypillisesti $t_{p\ max}$=(4...10) ms aikana. Sekä ammuksen että tuliputken nopeasta kuormituksesta aiheutuvissa muodonmuutoksissa venymänopeudet $\dot{\varepsilon}_{dyn} \approx (0,1...10)\ 1/s$ ovat selvästi suurempia kuin perinteisen staattisen aineenkoetuksen venymänopeudet kimmoisalla alueella $\dot{\varepsilon}_{staat} \approx (10^{-4}...10^{-3})\ 1/s$ [1]. Muodonmuutosnopeuden kasvaessa myös lujuuden on todettu nousevan teräksillä. Pienikin lujuuslisä saattaa olla ratkaisevassa asemassa erityisesti silloin, kun vanhaa asetta modernisoidaan suorituskyvyn lisäämiseksi putkivaiheen painekuormitusta kasvattamalla.

2. MUODONMUUTOSNOPEUDEN VAIKUTUS METALLISTEN MATERIAALIEN LUJUUSARVOIHIN

Metallisissa materiaaleissa muodonmuutosta välittävät pääasiassa dislokaatiot ja materiaalin myötölujuus riippuu vastuksesta, jonka se ominaisuuksillaan antaa dislokaatioliikkeelle. Vastus syntyy materiaalin mikrorakenteessa olevista esteistä, jotka jarruttavat dislokaation liikettä tai mahdollisesti pysäyttävät sen kokonaan. Kun lämpötila on alle puolet materiaalin sulamislämpötilasta ja muodonmuutosnopeus on $\dot{\varepsilon} \leq (1000...5000)\ 1/s$, dislokaatioiden kohtaama vastus muodostuu pääasiassa kahden tyyppisistä esteistä:

a) Laaja-alaisista esteistä, jotka vaikuttavat jännityskentillään metallin rakenteessa useiden tuhansien atomietäisyyksien päähän esteestä.
b) Suppea-alaisista esteistä, joiden vaikutus ulottuu vain muutamien atomietäisyyksien päähän esteestä.

Materiaalia kuormitettaessa dislokaatiot pyrkivät ylittämään tai läpäisemään tiellensä tulevat esteet. Laaja-alaiset esteet dislokaatio pystyy ylittämään vain riittävän suuren kuormittavan jännityksen avulla. Jännitys vaikuttaa myös suppea-alaisten esteiden ylittämiseen, mutta niiden kohdalla myös atomien lämpövärähtely voi avustaa muodonmuutosmekanismeja, jotka mahdollistavat esteiden ylittämisen. Koska avustava "voima" tulee lämpövärähtelystä, kutsutaan tällaisia muodonmuutosmekanismeja termisesti aktivoiduiksi. Niitä muodonmuutosmekanismeja, joihin terminen aktivaatio ei vaikuta, kutsutaan atermisiksi (mekanismit laaja-alaisten esteiden ylittämisessä).

Termisesti aktivoiduissa tapauksissa aktivaation vaikutus esteiden ylityksessä riippuu lämpötilan ohella myös muodonmuutosnopeudesta. Muodonmuutosnopeutta kasvatettaessa todennäköisyys ja "odotusaika" riittävän suurelle satunnaiselle lämpövärähdykselle pienene-

vät ja yksittäisen esteen ylittämiseen vaadittava jännitys kasvaa. Erittäin suurilla muodonmuutosnopeuksilla $\dot{\varepsilon} \geq 5000\ 1/s$ dislokaatioiden liikevastus kasvaa materiaalissa esiintyvien erilaisten vaimennusefektien ansiosta ja lujuus nousee voimakkaasti muodonmuutosnopeutta edelleen kasvatettaessa.

3. MYÖTÖLUJUUDEN JA MUODONMUUTOSNOPEUDEN VÄLISEN YHTEYDEN MALLINTAMINEN

Materiaalitieteellisistä lähtökohdista muodostetuissa malleissa on esteitä ylittävän, termisesti aktivoidun dislokaatioliikkeen määrää kuvattu lämpötilan ja muodonmuutosnopeuden funktiona. Mallit perustuvat Arrheniuksen yhtälöön [2]

$$\Delta U = RT \ln (A/\dot{\varepsilon}), \tag{1}$$

jossa ΔU on muodonmuutokseen käytettävissä oleva aktivaatioenergia, $\dot{\varepsilon}$ on muodonmuutosnopeus, T on lämpötila, R on kaasuvakio ja A on materiaalivakio (ns frekvenssitekijä). Jos yhtälöön (1) sijoitetaan aktivaatioenergian ΔU paikalle jännityksen sisältävä lauseke, jolla kuvataan suppea-alaisen esteen ylittäminen, voidaan jännitys, ainakin periaatteessa, ratkaista muodonmuutosnopeuden, lämpötilan ja materiaaliominaisuuksien funktiona. Näiden mallien pätevyys on rajoittunut vain joihinkin puhtaisiin metalleihin. Mallien ongelmina ovat termisesti aktivoitujen muodonmuutosmekanismien kuvaaminen matemaattisesti, pätevyysalueiden määrittäminen muodonmuutosnopeus- ja lämpötila-asteikolla sekä suurilla muodonmuutosnopeuksilla alkavat muodonmuutosta vastustavat vaimennusefektit.

Kokeellisesti on havaittu useilla metalleilla yhtälön (1) kaltaista riippuvuutta lujuuden ja muodonmuutosnopeuden välillä. Konstruktiometalleille on kehitetty puoliempiirisiä malleja, jotka sisältävät jollakin tavalla termisen aktivaatioenergian ajatuksen ja mahdollisesti myös yhtälön (1) tapaisen muodon. Lisäksi on kehitetty kokonaan empiirisiä malleja. Eräs käytännön suunnitteluun ja lujuusteknisiin tarkasteluihin (mm. ABAQUS-ohjelmisto) soveltuva empiirinen konstitutiivinen malli on Cowper-Symonds yhtälö

$$\dot{\varepsilon}^e_{pl} = D(\sigma^e/\sigma_o - 1)^q, \quad \sigma^e \geq \sigma_o, \tag{2}$$

jossa D ja q ovat materiaalivakioita, σ^e on jännitystilan vertailujännitys, σ_o on materiaalin staattinen myötölujuus ja $\dot{\varepsilon}^e_{pl}$ on efektiivinen, plastinen muodonmuutosnopeus. Yhtälöä voidaan käyttää materiaalin dynaamisen myötölujuuden määrittämiseen kirjoittamalla se muotoon

$$\sigma'_o/\sigma_o = 1 + (\dot{\varepsilon}_{pl}/D)^{1/q}, \tag{3}$$

jossa σ'_o on dynaaminen myötölujuus. Von Mises'n myötöehdolla aksiaalisessa vetokokeessa vertailujännitys on $\sigma'_o = \sigma^e = \sigma_x$ myötöhetkellä ja vastaava efektiivinen venymänopeus $\dot{\varepsilon}_{pl} = \dot{\varepsilon}^e_{pl} \approx \dot{\varepsilon}_x$. Pehmeälle rakenneteräkselle (mild steel) on vakioiden arvoilla q= 5 ja

D=40,4 1/s saavutettu hyvä yhteensopivuus kokeellisesti saatujen dynaamisten myötölujuusarvojen kanssa [5].

4. MUODONMUUTOSNOPEUDEN VAIKUTUS TERÄSTEN LUJUUSARVOIHIN

Tavallisesti konstruktioperusteena käytettävän myötölujuuden (R_{eL} tai $R_{p0,2}$) on pehmeillä rakenneteräksillä todettu kasvavan likimain lineaarisesti logaritmisen muodonmuutosnopeuden funktiona. Suurilla muodonmuutosnopeuksilla myötölujuus kasvaa yleensä voimakkaammin kuin murtolujuus R_m (suhteessa staattisiin arvoihin) ja eniten kasvaa ylempi myötölujuus R_{eH}, jos teräksellä esiintyy terävä myötöraja. Teräksien lujuusarvojen kasvu näyttää ainakin ulkoisesti riippuvan materiaalin staattisesta lujuudesta: kasvu on voimakkainta pehmeillä teräksillä ja vähäisintä lujilla laaduilla. Syynä tähän ei ilmeisestikään ole pelkkä lujuustaso sinänsä, vaan myös terästen koostumukselliset ja mikrorakenteelliset tekijät, mutta niiden yksityiskohtaista vaikutusta ei tunneta kovin tarkasti [6, 7].

Lujuustason vaikutus muodonmuutosnopeudesta aiheutuvaan riippuvuuteen on nähtävissä eri teräslaaduille mitatuista koetuloksista. Jos muodonmuutosnopeutta kasvatetaan staattisen vetokokeen arvosta $\dot{\varepsilon} \approx 10^{-4}$ 1/s arvoon $\dot{\varepsilon} \approx 10^{0}$ 1/s (= 1 1/s), myötölujuuden suhteellinen kasvu ferriittisillä ja martensiittisilla teräksillä on kirjallisuustutkimuksen perusteella [7]:

- Rakenneteräs Fe 37 (vast): n. (35...50) %
- Rakenneteräs Fe 52 (vast): n. 20 %
- Keskihiiliset ferriittis-perliittiset teräkset ja niukkaseosteiset nuorrutusteräkset, joiden staattinen myötölujuus on välillä R_e=(500...700) MPa: n. (15...20) %
- Lujat nuorrutusteräkset ja työkaluteräkset, joiden staattinen myötölujuus on välillä R_e=(1000...1700) MPa: n. (4...8) %.

Lujilla teräksillä lujuusarvojen kasvu on vähäistä, jos muodonmuutosnopeus jää alle $\dot{\varepsilon} < (10...10^2)$ 1/s. Kirjallisuudessa on julkaistu tuloksia, joiden mukaan lujuutta lisäävä vaikutus olisi likimain nolla em. muodonmuutosnopeuksiin saakka ja alkaisi näkyä vasta suuremmilla nopeuksilla. Tältä nopeusalueelta mittaustuloksia on julkaistu melko vähän ja käyttökelpoisia referenssituloksia on kaivattu lisää vielä viimeaikoinakin [6].

5. MITTAUSMENETELMÄT JA -LAITTEET

Keskinopealla muodonmuutosnopeusalueella aineenkoetus toteutetaan yleensä nopeilla hydraulisilla vetokoneilla, vetokokeisiin soveltuvilla iskuvasaroilla tai vauhtipyörävetolaitteilla. Eräitä mittausmenetelmiin ja vetosauvojen geometriaan liittyviä suosituksia ja ohjeita on kuvattu standardiehdotuksessa [3] sekä lähteessä [2]. Vetokokeen dynaamisuudesta johtuen perinteiset kiinteisiin voima-antureihin, leukojen liikkeeseen tai (sähkö-) mekaaniseen venymäanturiin (ekstensiometriin) perustuvat mittaukset eivät anna luotettavia tuloksia.

5.1 Koejärjestelyt

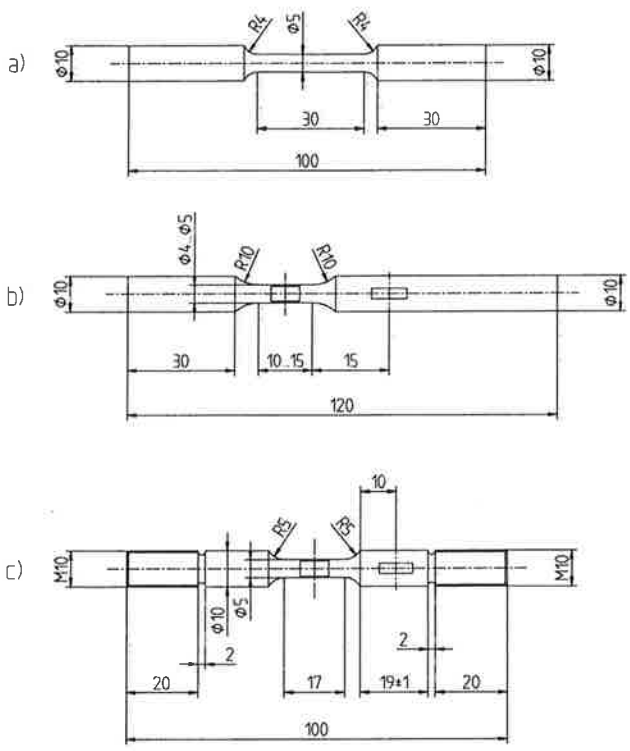
Ase- ja ammusteräksille suoritetuissa dynaamisissa vetokokeissa vetosauvaan kohdistuva voima ja vartalon venymä mitattiin sauvoihin liimatuilla venymäliuskoilla. Voimaa mittaava(-t) liuska(-t) sijoitettiin sauvan paksulle osalle ja venymää mittaava liuska vartaloon. Ensimmäisissä kokeissa [7] voimaa mitattiin vain yhdellä liuskalla ja voimamittauksessa esiintyi suurta hajontaa, jonka arvioitiin johtuvan taivutuksesta. Myöhemmissä kokeissa voima mitattiin taivutuskompensoidulla kahden venymäliuskan mittauskytkennällä ja sauvan vartalon venymä yhdellä liuskalla suoraan koesauvasta. Kuvassa 1 on esimerkkejä erään koesarjan vetosauvoista. Muodonmuutosnopeutta säädettiin vetoleukojen liikenopeutta ja/tai sauvan geometriaa muuttamalla. Vertailuarvoina käytetyt staattiset lujuusarvot määritettiin standardin [1] mukaisilla vetokokeilla.

Voiman ja venymän mittauksissa kokeiltiin liuskojen ja liimojen yhdistelminä:
- Voima: Liuskat Kyowa KFG-3 ja KFG-5, liima Kyowa CC-33A pikaliima.
- Venymä: Liuskat Kyowa KFG-2N ja KFG-3 (ilmoitettu mittaavan venymää vähintään 5% saakka), liimana Kyowa CC-33A. Liuskat irtosivat sauvoista $\varepsilon \leq 3$ % venymillä.
- Venymä: Liuskat Kyowa KFE-2-120-C1 (ilmoitettu mittaavan venymää vähintään 8 % saakka), liimoina lämmöllä (~80°C) kovettuvat Kyowa EP 18 sekä Micro-Measurements A-12 liimat. Liuskat mittasivat venymää luotettavasti (5...8) % saakka, joissakin kokeissa jopa (10...12) % asti.

Mittavälineet:
- Tasajännitesiltavahvistin Kyowa CDV-230C, taajuuskaista (0...200) kHz
- Signaalitallentimet Data Precision D6000 tai D6500 pistoyksiköillä mod 611 tai mod 650 varustettuna, taajuuskaista vaihteli vetonopeudesta riippuen alueilla (0...20...25...40...100 ...200) kHz
- Muutamissa vetokokeissa signaalitallentimen rinnalla käytettiin varmistuksena digitaalinauhuria TEAC RD-145 T, taajuuskaista 10 kHz
- Fluke 77 II yleismittari voimamittasillan staattisessa kalibroinnissa.

Aineenkoetuskoneet:
- Instron 1186 mekaaninen aineenkoetuskone, staattiset vetokokeet, PvTeknTL
- MTS 100 kN hydraulinen aineenkoetuskone, staattiset ja dynaamiset vetokokeet, TTKK
- Matertest HSMT 22 kN hydraulinen nopeavetokone, dynaamiset vetokokeet, PvTeknTL.

Instron ja MTS laitteet ovat tavallisia aineenkoetuskoneita, joissa koesauva kiinnitetään ruuvi- tai hydraulileuoilla. MTS-vetokoneella on mahdollista saavuttaa v_{max}=150 mm/s leukojen liikenopeus, jolloin sauvaan kohdistuva muodonmuutos on keskinopealla alueella. Matertest HSMT 22 kN hydraulinen nopeavetokone on tarkoitettu dynaamiseen aineenkoetukseen, jonka vetoliikkeen nopeus suurimmillaan on v_{max}= 1 000 mm/s = 1 m/s. Matertest-vetokokeessa koesauva kiihdytetään äkillisesti suureen vetonopeuteen törmäyttämällä massiivinen vedin vetosauvan päähän. Vedin on kiinnitetty hydraulisylinterin männän varteen ja vetokoe jatkuu törmäyksen jälkeen männän vetämänä. Voima on mahdollista mitata voima-anturilla, johon koesauva on kiinnitetty yläpäästään. Leukojen liikettä mitataan sähkö-



Kuva 1. *Erään koesarjan vetokoesauvojen mitat ja venymäliuskojen sijainti.*
a) Staattiset vetokokeet, Instron 1186 aineenkoetuskone, ei venymäliuskoja.
b) Staattiset ja dynaamiset vetokokeet, MTS 100 kN.
c) Dynaamiset vetokokeet, Matertest HSMT 22 kN.

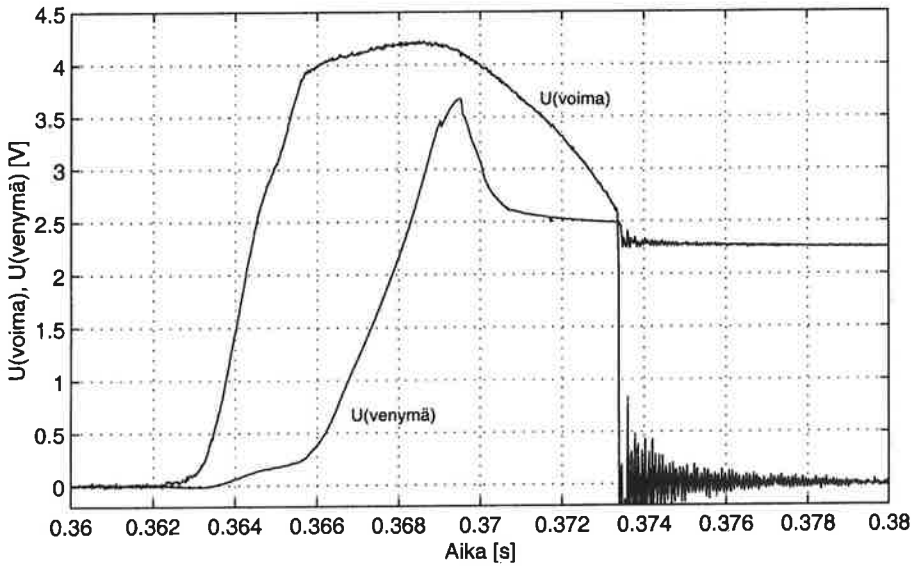
mekaanisella siirtymäanturilla. Nopean vetokokeen dynaamisesta luonteesta (äkillinen isku, runkorakenteiden ja vetomekanismin värähtelyt) johtuen vetokoneen oma voimamittaustulos todettiin kuitenkin epäluotettavaksi suurilla vetonopeuksilla. Aineenkoetuskokeiden vetonopeuksien asetusarvot vaihtelivat alueella v=(0,0045...300) mm/s. Alueen alaraja edustaa staattisen kokeen kimmoisen alueen vetonopeutta ja yläraja dynaamisen Matertest-nopeavetokokeen suurinta käytettyä vetonopeutta.

Voimaliuskojen kytkennät kalibroitiin ennen kokeita vertaamalla mittasillan antamaa tulosta aineenkoetuskoneiden kalibroitujen voima-antureiden näyttämiin. Kalibrointivoimana käytettiin kaikissa koesauvoissa kuormituksia, jotka jännityksinä vastasivat (30...40) % koemateriaalien staattisten vetokokeiden myötölujuuksista (R_{eL} tai $R_{p0,2}$).

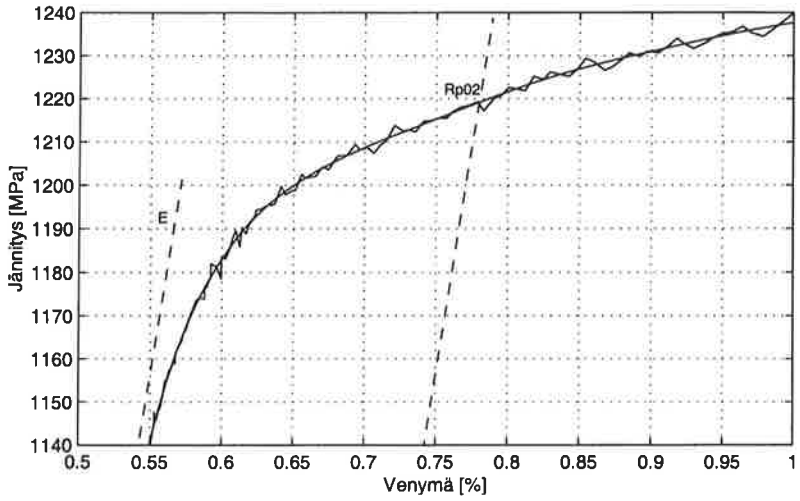
5.2 Mittaustulosten jälkikäsittely

Vetokokeiden venymäliuskamittausten tulokset käsiteltiin Matlab-ohjelmistolla:
- Signaalien nollatasot korjattiin alkuosien keskiarvolla, esimerkki kuvassa 2.
- Mitatut jännitteet muutettiin nimellisiksi jännitykseksi ja venymiksi (insinöörivenymiksi).
- Mittaustulosten epäolennaiset loppuosat leikattiin pois.
- Mittaustulosten hyväksytyt osuudet suodatettiin laskennallisesti alipäästösuodattimella, jonka katkaisutaajuudet olivat f_{cut}= {0,2 235 2000} Hz staattisissa ja dynaamisissa kokeissa. Katkaisutaajuudet valittiin "silmämääräisesti" kokeilemalla suodatuksen vaikutusta mittaustuloksista piirrettyihin jännitys-venymäkäyriin, kuva 3.

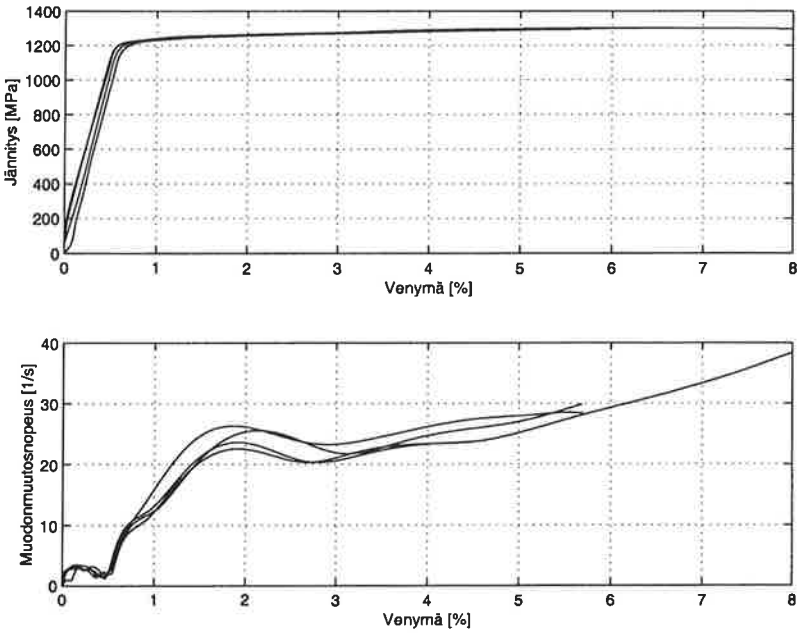
Alipäästösuodatetuista mittaustuloksista piirrettiin vetokokeiden jännitys-venymäkuvaajat, laskettiin kimmomoduulit E ja määritettiin materiaalista riippuen joko venymisrajat $R_{p0,2}$ tai ylemmät ja alemmat myötörajat R_{eH}, R_{eL} sekä murtolujuudet R_m kokeittain. Venymisraja $R_{p0,2}$ määritettiin kuvan 3 tapaan piirtämällä kimmoisen osan suuntainen suora kunkin kokeen kimmomoduulia apuna käyttäen 0,2 % pysyvän venymän etäisyydelle alkuperäisen kuvaajan suoralta osalta ja hakemalla näiden leikkauspiste. Muodonmuutosnopeus määritettiin peräkkäisten venymäpisteiden perusteella $\dot{\varepsilon}(t) = d\varepsilon/dt = (\varepsilon_{i+1} - \varepsilon_i)/(t_{i+1} - t_i)$. Esimerkkejä jännitys-venymä- ja muodonmuutosnopeus-venymäpiirroksista on esitetty kuvassa 4.



Kuva 2. *Esimerkki Matertest-nopeavetokokeen alkuperäisestä mittausdatasta; Voima-aika- ja venymä-aika-käyrät. Tulokset jännitteinä. Vetonopeus 300 mm/s.*



Kuva 3. *Alipäästösuodatuksen vaikutus MTS-kokeen jännitys-venymäpiirrokseen ja venymisrajan $R_{p0,2}$ määritys myötöalueelta. Vetonopeus 75 mm/s.*



Kuva 4. *Esimerkki Matertest-nopeavetokokeen tuloksista. Vetonopeus 300 mm/s. Yläkuvassa alipäästösuodatettu jännitys-venymäpiirros neljästä kokeesta. Alakuvassa muodonmuutosnopeus-venymäpiirros samoista kokeista.*

Muodonmuutosnopeus-venymä piirroksista määritettiin kullekin kokeelle kimmoisten ja plastisten muodonmuutosnopeuksien arvot. Koska muodonmuutosnopeuden pitäminen vakiona on dynaamisissa vetokokeissa lähes mahdotonta (vrt kuva 4), koekohtaiset arvot laskettiin useilla eri tavoilla:

1.) *Kimmoiset muodonmuutosnopeudet*

1a) Keskimääräinen kimmoinen muodonmuutosnopeus $\dot{\varepsilon}_{elke}$

- jännitys-venymäpiirroksen kimmoisalta väliltä laskettu muodonmuutosnopeuksien keskiarvo, kuvassa 5 alueen (200...1000) MPa muodonmuutosnopeuksien keskiarvo.

1b) Kimmoisen alueen loppunopeus $\dot{\varepsilon}_{ellop}$

- jännitys-venymäpiirroksen kimmoisen alueen (suoran) loppuosaa (~100 MPa:n jännitysmuutosta vastaavalta alueelta) vastaavien muodonmuutosnopeuspisteiden keskiarvo, kuvassa 5 alueen (950...1050) MPa muodonmuutosnopeuksien keskiarvo.

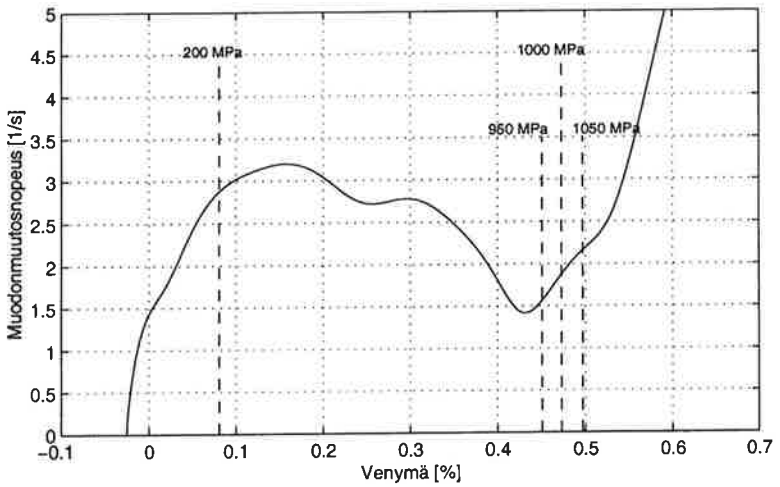
2.) *Venymisrajan $R_{p0,2}$ muodonmuutosnopeus* $\dot{\varepsilon}_{0,2}$

- jännitys-venymäkäyrän 0,2-rajaa vastaava muodonmuutosnopeuden arvo.

3.) *Plastinen muodonmuutosnopeus* $\dot{\varepsilon}_{pl}$

- muodonmuutosnopeuden keskiarvo venymisrajan (0,2-rajan) jälkeisistä pisteistä suurimpaan hyväksyttyyn venymäarvoon asti.

Tutkimuksessa keskityttiin myötölujuuden kokeelliseen määrittämiseen kimmoisen muodonmuutosnopeuden funktiona, koska myötölujuuden arvo ko rasitustilassa esiintyvällä muodonmuutosnopeudella määrää sekä ammuksen säteettäisten siirtymien äkillisen kasvamisen (putkikosketus) että aseen putken vaurioitumisen (pysyvä muodonmuutos).



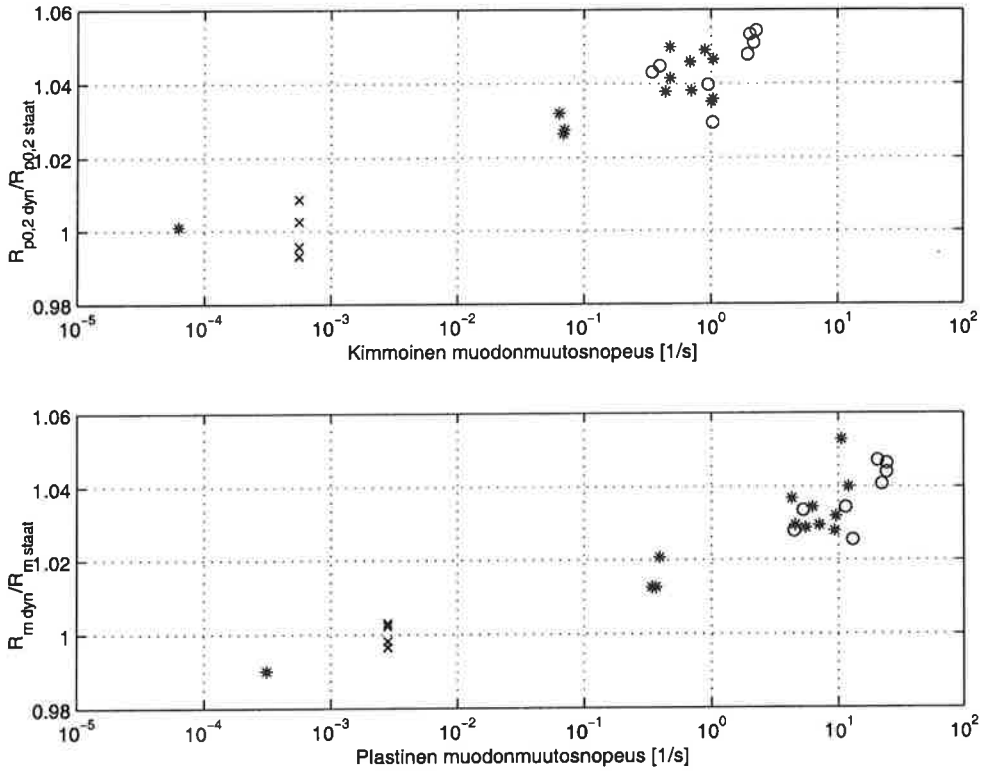
Kuva 5. *Kimmoisten muodonmuutosnopeuksien määritys Matertest-vetokokeen muodonmuutosnopeus-venymäkäyrältä. Kuvaan merkitty jännitys-venymäkäyrän jännitystasojen σ_i={200 950 1000 1050}MPa kohdat. Vetonopeus 300 mm/s.*

6. ESIMERKKIKOKEIDEN TULOKSET

Kuvissa 6 ja 7 on esitetty tuloksia kahdella eri koemateriaalilla tehdyistä aineenkoetuskokeista. Tulokset on esitetty myötö- ja murtolujuuksien suhteellisina arvoina muodonmuutosnopeuksien funktiona. Koemateriaaleina ovat luja nuorrutettu aseteräs (kuva 6, 25 vetokoetta) ja niukasti seostettu nuorrutettu kranaattiteräs (kuva 7, 23 koetta). Mittapisteet kuvissa ovat:

(x) staattinen vetokoe, "Instron 1186" vetokone, PvTeknTL.
(*) staattinen ja dynaaminen vetokoe, "MTS 100 kN", TTKK/Materiaaliopin lait.
(o) dynaaminen vetokoe, "Matertest 22 kN", PvTeknTL.

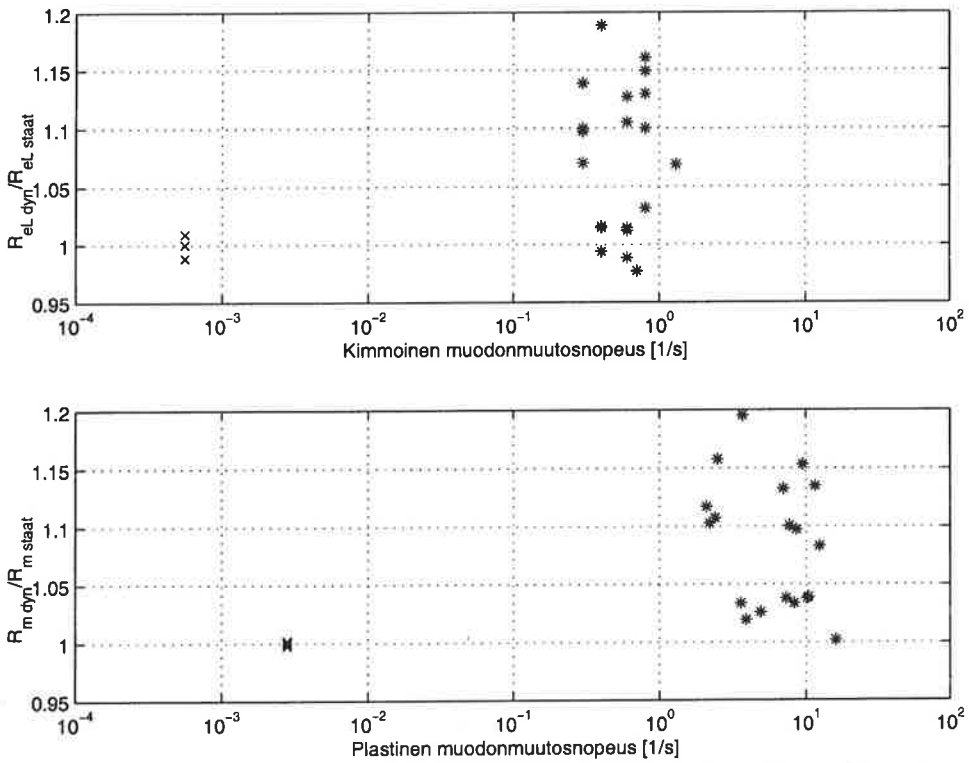
Vertailuarvoina käytetyt staattiset lujuusarvot $R_{xx\ staat}$ määritettiin staattisten vetokokeiden keskiarvoina tai yksittäisinä vetokokeen tuloksina materiaalista ja aihioista riippuen.



Kuva 6. *Aseteräksen suhteellinen venymisraja $R_{p0,2\ dyn}/R_{p0,2\ staat}$ kimmoisen muodonmuutosnopeuden $\dot{\varepsilon}_{ellop}$ funktiona ja suhteellinen murtolujuus $R_{m\ dyn}/R_{m\ staat}$ plastisen muodonmuutosnopeuden $\dot{\varepsilon}_{pl}$ funktiona. Staattinen venymisraja $R_{p0,2\ staat}$ ja murtolujuus $R_{m\ staat}$ on neljän staattisen kokeen (x) keskiarvo.*

Mittaustulosten perusteella lujan aseteräksen dynaaminen venymisraja $R_{p0,2\ dyn}$ kasvoi keskimäärin ~5 % staattiseen venymisrajaan verrattuna, kun kimmoinen muodonmuutosnopeus oli $\dot{\varepsilon}_{ellop} \approx 2\ 1/s$ (loppunopeus). Vastaavasti dynaamisen murtolujuuden $R_{m\ dyn}$ suhteellinen kasvu oli ~4,5 % plastisella muodonmuutosnopeudella $\dot{\varepsilon}_{pl} \approx 22\ 1/s$. Koesarjan venymisrajojen hajonta oli kohtuullista ja verrattavissa staattisten kokeiden sekä kirjallisuudessa julkaistujen dynaamisten kokeiden mittaustuloksiin. Venymisrajan määrityksen suhteelliseksi epävarmuudeksi arvioitiin koelaitteistosta riippuen (1,5...3,5) %.

Kranaattiteräksen myötö- ja murtolujuuksien (R_{eL}, R_{eH}, R_m) kasvu oli keskimäärin (7...9) %, kun kimmoinen muodonmuutosnopeus oli alueella $\dot{\varepsilon} \approx$ (0,8...1,1) 1/s. Pehmeällä aihiolla, jonka staattinen myötölujuus oli vaatimusrajalla, lujuuden kasvu oli keskimääräistä suurempaa ja lujalla aihiolla ($R_{eL} \approx 1{,}2 * R_{eL\ vaad}$) lujuusarvot kasvoivat vähemmän. Kranaattiteräksellä tehdyissä kokeissa tulosten hajonta oli likimain yhtä suurta kuin havaittu keskimääräinen lujuuden kasvu, joten tilastollisen luotettavuuden saavuttamiseksi kokeita on jatkettava. Hajonnan suuruuteen kranaattiteräksellä vaikuttivat koesauvojen valmistaminen lujuustasoltaan erilaisista aihioista sekä voiman mittaus vain yhdellä venymäliuskalla, jolloin koesauvan taivutuksen vaikutusta ei kyetty eliminoimaan.



Kuva 7. *Kranaattiteräksen suhteellinen alempi myötölujuus $R_{eL\ dyn}/R_{eL\ staat}$ kimmoisen muodonmuutosnopeuden funktiona ja suhteellinen murtolujuus $R_{m\ dyn}/R_{m\ staat}$ plastisen muodonmuutosnopeuden funktiona. Instron (x) ja MTS (*) kokeet.*

7. YHTEENVETO

Aseiden ja ammuksien sisäballistisen vaiheen lujuusteknisessä mitoituksessa voidaan hyödyntää laukausrasituksista aiheutuvan muodonmuutosnopeuden tuoma teräksen lujuuden lisäys, jos lujuusarvojen kasvu kyetään todentamaan luotettavasti keskinopeilla dynaamisilla aineenkoetuskokeilla. Suoritetuissa vetokokeissa saavutettiin aseteräkselle ~5 % ja kranaattiteräkselle ~7 % keskimääräinen myötölujuuden lisäys, kun kimmoinen muodonmuutosnopeus oli luokkaa ~10° 1/s, jolla laukausrasituksista aiheutuvien muodonmuutosten arvioidaan syntyvän sekä aseen putkeen että ammuskuoreen. Kirjallisuustietojen perusteella saavutetut lujuusarvot vastaavat saman tyyppisille teräksille julkaistujen dynaamisten aineenkoetuskokeiden tuloksia.

KIITOKSET

Tutkimus toteutettiin Puolustusvoimien Materiaalilaitoksen Esikunnan Ase- ja Ampumatarvikeosastojen tukemana, joille osoitamme kiitoksemme mielenkiintoisesta tutkimusaiheesta, koemateriaalien luovuttamisesta sekä kirjoitelmassa esitettyjen tutkimustulosten julkaisuluvasta. Teknikko Pauli Lintulaa kiitämme Matertest 22 kN nopeavetokoneen käyttöönottoon liittyneistä työkalusuunnittelusta ja -hankinnoista, käyttöohjeiden laatimisesta sekä alustavien että varsinaisten nopeavetokokeiden teknisestä suorituksesta.

VIITTEET

1. SFS-EN 100002-1. *Metallien vetokoe. Osa 1: Menetelmä.* Standardi, Suomen Standardoimisliitto SFS. Helsinki, 21.12.1990. 33 s.

2. Meyer, L.W. *Werkstoffverhalten hochfester Stähle unter einsinnig dynamischer Belastung.* Dissertation. Dortmund, 1982.

3. *Draft Standard for Dynamic Tensile Tests.* Stuttgart, 28.11.1991. 13 p.

4. Dejus, R. & Nilsson, U. & Torstensson, H. *Quality in Mechanical Testing of Metallic Materials.* Nordtest, NT Tech Report 238. Espoo, 1994. 52 p.

5. Jones, N. *Structural Impact.* Cambridge Univ Press. Cambridge, 1989. 575 p.

6. Itabashi, M. & Kawata, K. *Carbon content Effect on high-strain-rate tensile properties for carbon steels.* In: International Journal of Impact Engineering, Pergamon Press. vol 24. 2000. pp. 117...131.

7. Erkkilä, T. & Salonen, L. *Muodonmuutosnopeuden vaikutus teräksen myötölujuuteen, aineenkoetuskokeet vuonna 1993.* PvTKfys-os:n tutkimustodistus, Ylöjärvi, 1996. 19 s + 1 liites. (Luottamuksellinen, julkaisematon).

VISCOPLASTIC MODELS OF ALUMINIUM ALLOY AA 6063-T6 AT ELEVATED TEMPERATURES

J. MYLLYMÄKI[1]
Laboratory of Structural Mechanics, Helsinki University of Technology
FIN-02150 Espoo, Finland

ABSTRACT

Mechanical properties of aluminium alloy AA 6063-T6 in the temperature range of 20 °C to 400 °C has been investigated. The effect of the strain rate and the heating rate to the behaviour of the alloy at high temperature tests was studied. Additive and multiplicative strain rate dependent viscoplastic constitutive laws have been used in the analysis of the behaviour of the aluminium alloy AA 6063-T6 at high temperatures. In the identification of the parameters of the viscosity-hardening laws an inverse solution technique, regularized output least squares method (RLS), has been used. In the inverse solution stress controlled explicit time integration was applied. Applying the viscoplastic models to tests, which were not used in the parameter identification, validates the applicability of the parameters. In the validation computations strain controlled implicit time integration methods were used. It is shown that both additive and multiplicative viscoplasticity laws satisfactorily predict behaviour of aluminium alloy AA6063-T6 at elevated temperatures both in transient temperature creep tests and in steady state tests of different strain rates.

INTRODUCTION

The process of deterioration of metal-supported building elements in fire can be predicted with reasonable accuracy, if the plastic behaviour of metals at elevated temperatures is known. Substantial amount of information is already available concerning various properties of steels at elevated temperatures [1-4]. Also a lot of investigations have been conducted on the properties of aluminium at elevated temperatures [5-10]. Unfortunately most of the information is studied in the metallurgical point of view and the results are in many cases not applicable to the commercial aluminium alloys. Most of the tests of available knowledge of the properties of aluminium alloys have been obtained using classical steady-state tensile tests or creep tests [9-11]. Although one of the first studies on

[1] Present address: VTT BUILDING TECHNOLOGY, Fire Technology, P.O.Box 1803, FIN-02044 VTT
Email: Jukka.Myllymaki@vtt.fi

metals at transient state has been conducted with alumium alloys [8] there is not much information of the creep behaviour of aluminium alloys at transient-state temperature.

To provide information to design engineers, primarily to those concerned with various aspects of fire endurance of aluminium structures in fire, a research program was launched at Laboratory of Structural Mechanics of Helsinki University of Technology supported Technology Development Centre under grant 4096/95.

This paper is presentation of the test results and theoretical work of the project. It is shown that the strain rate has a considerable effect on the yield properties of the alloy at higher temperatures. The transient test results show temperature rate dependence. These phenomena can be explained only by using strain rate (creep) dependent plastic models [2]. The creep effects of steel and aluminium have been treated in the fire safety science [2, 5] using elasto-plastic models with Dorn-Harmathy creep model [1]. Traditionally the parameters of the strain rate (creep) dependent are identified using tensile, relaxation and creep tests at constant temperature using a log-log fitting [6, 15].

In this paper the plastic and creep behaviour of the aluminium alloy has been modeled using additive and Norton-Bailey type multiplicative viscoplasticity laws, which are suitable for several metals [15] also for aluminium [6, 7]. Identification of the parameters of the models is based both on steady-state and transient test results. This requires advanced mathematical tools. Here technique called regularized output least squares method (RLS) in the theory of inverse problems [13] has been applied. Regularization by mesh coarsing and Tikhonov-regularization [16] was used in order to get more stabilized inverse solutions.

EXPERIMENTS

Test material used was tempered aluminium alloy AA 6063-T6 manufactured by Nordic Aluminium. Test specimens (Fig. 1) were cut out from as received 20-mm-wide and 4-mm-thick aluminium alloy AA 6063-T6 sheet longitudinally to wrought direction. Chemical composition (%) of the test material measured by spectral analysis was as follows:

Mg	Si	Mn	Cu	Fe	Zn	Ti	V	Pb	Cr	Zr	Ti+Zr
0.48	0.44	0.0043	0.0038	0.22	0.012	0.022	0.009	0.009	0.0031	0.003	0.025

Hardening tests at constant temperature

It is important to distinguish between the two main methods used in this study. In the steady state hardening test the unloaded specimen is brought into thermal equilibrium at a certain temperature and the specimen is subjected to a deformation at constant speed. In the steady state hardening tests the oven was controlled so that temperature rate of the air of

the oven was first 20 or 30 K/min and after that the temperature was maintained constant. After the specimen temperature had reached a constant value, the tension test was started. The maximum heating time before the test was about 60 min so all the results also transient ones represent results of less than 1 h holding time.

Steady-state hardening tests were carried out with two equal tests at strain rate $7x10^{-5}$ s^{-1} (crosshead speed 0,18 mm/min) and one test at strain rate 7×10^{-4} s^{-1} (crosshead speed 1,8 mm/min) at temperatures 50 °C, 100 °C, 150 °C, 200 °C, 225 °C, 250 °C and 300 °C. The strain was recorded as a function of time and the strain rate could be calculated from that data. The strain rate varied being lower at elastic part and higher after yielding. Given values are the strain rates just before and after yield. Three tensile tests were carried out at room temperature to determine the mechanical properties of aluminium alloy AA 6063-T6 at room temperature.

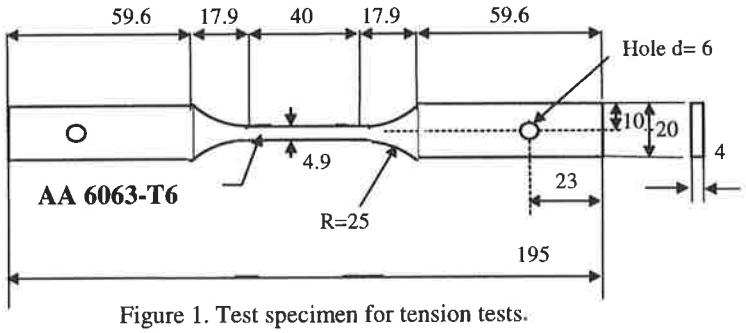


Figure 1. Test specimen for tension tests.

Creep tests at transient temperature

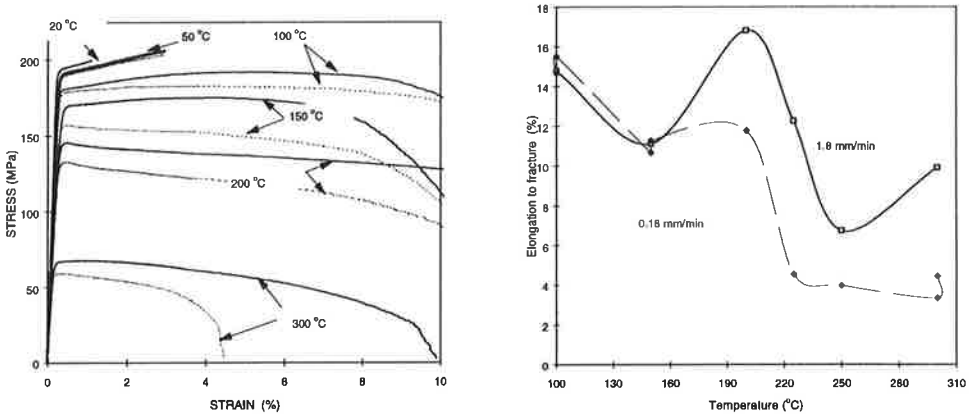
It is common for loaded metal component to be subject to a change in temperature and it is important to know how the resulting deformation of the material will develop. Therefore tensile creep tests under transient heating conditions have been devised [1, 3, 4] in which the load on the tensile specimen is maintained constant while its temperature is increased at a given rate.

In the transient temperature tests the oven was controlled so that the air temperature rate in the oven was 5, 10 and 30 K/min. Transient state tensile tests were carried out with two equal tests at each stress level of 3, 20, 40, 60, 80, 100, 120, 140, 160, 180 and 190 MPa. Thermal strain of the aluminium alloy was determined with three tests at load level of 3 MPa.

TEST RESULTS

The stress-strain curves of AA 6063-T6 alloy at strain rates 7×10^{-5} and 7×10^{-4} are shown in Fig. 2a. Following feature is apparent; at the same strain, the higher the strain rate is, the higher will be the stress. The effect of temperature on the ductility of quenched and aged specimens of AA 6063-T6 alloy and tensile tested at strain rates 7×10^{-5} and 7×10^{-4} is shown in Fig. 2b. For the same strain rate, the elongation of the alloy increases with test temperature to a maximum and then decreases reaching a minimum value followed by subsequent increase at higher temperatures. At higher strain rate the elongation to fracture is also higher.

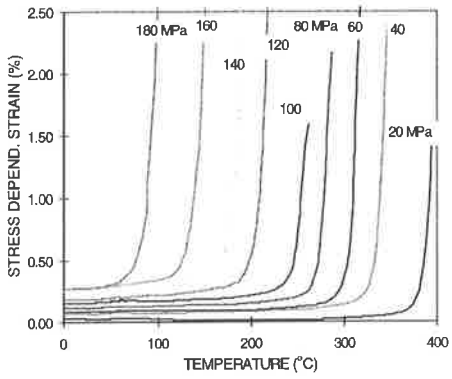
In the tests under transient heating conditions a series of stress dependent strain ($\varepsilon_\sigma = \varepsilon_e + \varepsilon_p$) curves was obtained for each heating rate by excluding the thermal strain. Typical curves, using a heating rate of 10 K/min are shown up to 2.5 % strain in Fig. 3 a . The principal features comprise a small initial elastic strain, when the load was applied at 20 °C, followed by a gradual strain increase as the temperature was increased.



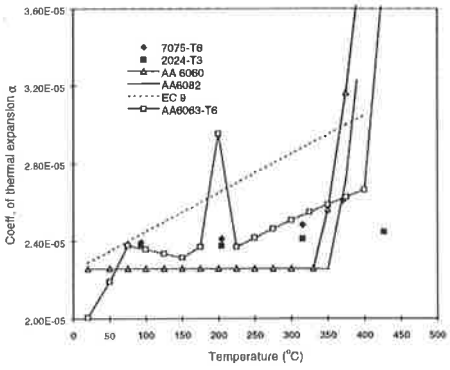
a) b)

Figure 2 a)The hardening (stress-strain) curves of aluminium alloy AA 6063-T6 at strain rates 7×10^{-5} and 7×10^{-4}. b) Effect of temperature on the elongation to fracture, strain rates 7×10^{-5} and 7×10^{-4} (crosshead speed 1.8 and 0.18 mm/min).

Thermal strain was obtained from the tests conducted at almost zero loading (3 MPa). Coefficient of thermal expansion α was calculated using regularized output least squares method []. The calculation results are shown in Fig 3 b and compared to the values of other experimental data.



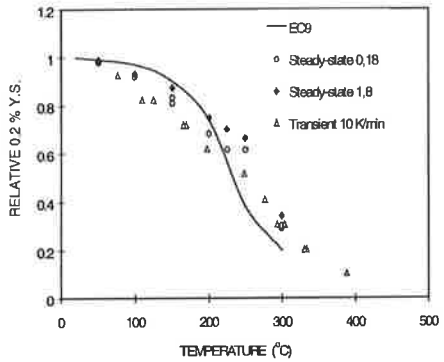
a)



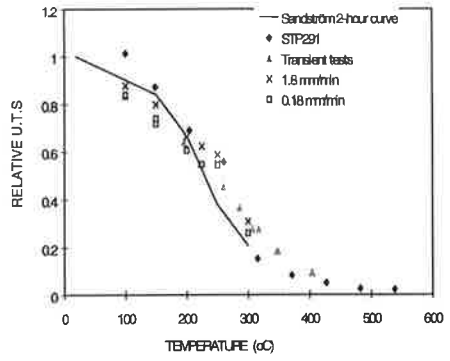
b)

Figure 3.a) Measured stress dependent strain (thermal strain excluded) of AA6063-T6 and b) coefficient of thermal expansion α derived from transient tests of 10 K/min . Thermal expansion curves of AA 6063 and AA6082 from Järvstråt [9] and 7075-T6 and 2024-T6 from Heimerl [8] and Eurocode 9 curve [12].

The variation of 0.2 % yield strength is shown in Figs.4a. The dependence of the ultimate tensile strength (U.T.S) on the test temperature is shown in Fig 4 b. In general, the U.T.S and yield strength decrease slowly from room temperature up to 150 °C and decrease rapidly at higher temperatures.



a)



b)

Figure 4 a) The dependence of the 0.2 % yield strength and b) the UTS on temperature, steady-state hardening test results at strain rate 7×10^{-5} and 7×10^{-4} (crosshead speed 1.8 and 0.18 mm/min) and transient creep test results at 10 K/min compared to Eurocode 9 curve [12], ASTM STP 291 data [11] and Sandström 2-hour curve [10] for alloy AA 6063-T6.

VISCOPLASTIC CONSTITUTIVE EQUATIONS

In the following formulations of unified viscoplastic constitutive equations are shortly explained. The viscoplastic models considered are the additive and multiplicative strain hardening models [15]. The models are treated in a simple one-dimensional form; only with isotropic hardening. Kinematic (directional) hardening and damage are not taken into account in the equations. Only the case of small strains (less than about 5 %) is discussed here. The total strain rate is assumed to be decomposable into elastic, plastic and thermal components

$$\dot{\varepsilon}_{ij} = \dot{\varepsilon}^{e}_{ij} + \dot{\varepsilon}^{p}_{ij} + \alpha \dot{T} \delta_{ij} \tag{1}$$

where the elastic component is given by the time derivative of Hooke´s law

According to Chaboche and Lemaitre [15] both in the additive and multiplicative viscoplastic laws the plastic strain rate expressed by equation

$$\dot{\varepsilon}^{p}_{ij} = \frac{3}{2}\ \dot{p}\ \frac{s_{ij}}{\sigma_{eq}} \tag{2}$$

where $\dot{e}^{p}_{ij}$ is the deviatoric plastic strain rate tensor and s_{ij} is deviatoric stress tensor and $\sigma_{eq} = \sqrt{\frac{3}{2} s_{ij} s_{ij}}$ is von Mises equivalent stress.

Multiplicative viscoplastic law

The evolution equation for the accumulated plastic strain is following in the case of multiplicative model

$$\dot{p} = \sqrt{\frac{2}{3} \dot{\varepsilon}^{p}_{ij} \dot{\varepsilon}^{p}_{ij}} = \left(\frac{\sigma_{eq}}{K(T)\ p^{1/M(T)}}\right)^{N(T)} \tag{3}$$

The uniaxial expression of the multiplicative model (3) viscosity-hardening law is

$$\dot{\varepsilon}_{p,11} = \left(\frac{\sigma}{K(T)\ \varepsilon_{p,11}^{1/M(T)}}\right)^{N(T)} \tag{4}$$

In uniaxial case we obtain following equations

$$\dot{\varepsilon} = \dot{\varepsilon}_e + \dot{\varepsilon}_p + \alpha(T)\dot{T} \qquad (5)$$
$$\sigma_{11} = E(T)\varepsilon_e$$
$$\sigma_{11} = K(T)\varepsilon_p^{1/M(T)}\dot{\varepsilon}_p^{1/N(T)}$$

where N is the viscosity exponent, M is the hardening exponent, K is the coefficient of resistance and E is the modulus of elasticity, ε_e is the linear elastic, ε_p the viscoplastic, ε_T the thermal strain and $\dot{\varepsilon}_p$ is the rate of the viscoplastic strain.

Additive viscoplastic law

The evolution equation for the plastic strain is following in the case of additive model

$$\dot{p} = \sqrt{\frac{2}{3}\dot{\varepsilon}_{ij}^p\dot{\varepsilon}_{ij}^p} = \left\langle \frac{\sigma_{eq} - R(T) - k(T)}{K_a(T)} \right\rangle^{N_a(T)} \qquad (6)$$

The uniaxial expression of the additive viscosity-hardening law is

$$\dot{\varepsilon}_{p,11} = \left\langle \frac{\sigma - R(\varepsilon_p, T) - k(T)}{K_a(T)} \right\rangle^{N_a(T)} \qquad (7)$$

The relation between R and ε_p must be chosen so as to approximate the hardening correctly. A simple relation [15] uses a linear term and an exponential term

$$R(\varepsilon_p, T) = Q_1(T)\ \varepsilon_p + Q_2(T)\left[1 - \exp(-b(T)\varepsilon_p)\right] \qquad (8)$$

The equation (7) can be inverted to give

$$\dot{\varepsilon} = \dot{\varepsilon}_e + \dot{\varepsilon}_p + \alpha(T)\dot{T} \qquad (9)$$
$$\sigma_{11} = E(T)\varepsilon_e$$
$$\sigma_{11} = k + R(T, \varepsilon_p) + K_a(T)\ \dot{\varepsilon}_p^{1/N_a(T)}$$

The coefficients N_a, K_a, k, Q_1, Q_2 and b may depend on temperature. This type of law is rarely used because of the identifying problems. Instead, the multiplicative product form, which is easier to use and identify, is more generally used. A simple form of the additive model used here is following where no strain hardening is assumed (R=0 and k=0)

$$\sigma_{11} = K_a(T)\ \dot{\varepsilon}_p^{\,1/N_a(T)} \tag{10}$$

PARAMETER IDENTIFICATION OF THE VISCOPLASTICITY LAWS

Consider a coefficient determination problem, i.e. the problem of determining a non-constant coefficients $\vec{a}(\mathbf{T}) = \left(K(T) \quad M(T) \quad N(T)\right)$ in an initial value problem (5) or (9) based on the existing measured data about the solution of uniaxial strain ε. The non-linear inverse problem was solved using the *regularized output least squares method* (RLS). The distributed unknown parameters $K(T), M(T), N(T), E(T)$ were discretized into a certain number of sub-intervals $[T_i, T_{i+1}]$ of arbitrary length $T_{i+1} - T_i$ using piecewise linear basis functions. The unknown parameters were seeked to minimize the functional

$$\min_{\vec{a}\in D} \left\| \varepsilon_{11}(\vec{a};\vec{x};t) - \varepsilon_{11,data}(\vec{x},t) \right\|^2 + \alpha \left\| \mathbf{L}\ \vec{a} \right\| , \tag{11}$$

where ε_{11} and $\varepsilon_{11,data}$ are the vectors of computed and measured uniaxial strain and $\alpha > 0$ is a regularization parameter depending on the noise level of the data. Explicit Euler scheme was used in the time integration of Equations (5) and (8). Either Newton or Conjugate Gradient methods were used in the minimization. The Euclidean norms were taken over each test i used in the identification at each collocation time step j $\|f(t)\|^2 = \sum_i \sum_j |f(t_j)|$.

As a regularization operator $\mathbf{L}$ a discrete version of the Laplacian with respect to the temperature was used. At first, the parameters were solved at constant temperatures using the hardening test results at two deformation rates. Part of the transient results at oven temperature rate 10 K/min were used to calculate parameters at temperature range 20 °C-400 °C.

In Fig. 5 the variables K_a, N_a of the additive model and the values of Young's modulus E are presented. Figures 6 shows the parameters of the multiplicative model; coefficient of resistance K and the viscosity and hardening exponents N, M . In both models the parameters are generally temperature and material dependent. It has been shown that M approximately varies between 3 and 100 and N has a value approximately between 3 and 50 [15] for several materials. These ranges apply also for some aluminium alloys [6, 7].

Example of the inverse solution show that the use of creep dependent models as the viscoplastic constitutive model illustrated above makes it possible to couple constant temperature test results obtained from hardening or relaxation tests with transient temperature results when the parameters of the constitutive law are identified.

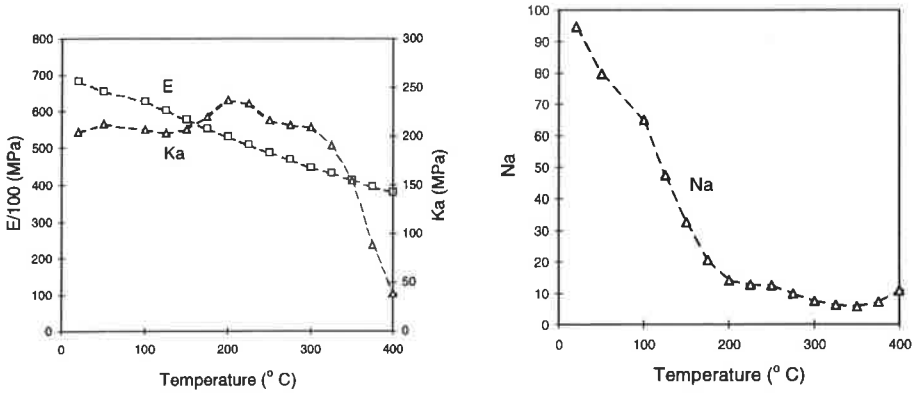


Figure 5. The parameters E, K_a, N_a of the additive viscosity law for aluminium alloy AA6063-T6.

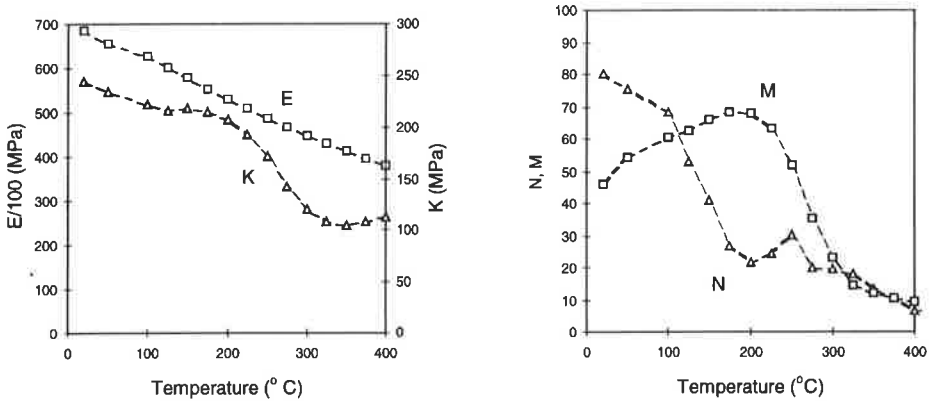


Figure 6. The parameters of the multiplicative viscosity law for aluminium alloy AA6063-T6; E, K, M and N as a result of the RLS inverse solution.

NUMERICAL EXAMPLES OF HARDENING, TRANSIENT AND CYCLIC TESTS

In Figures 7-9 both the experimental results and calculated ones are illustrated. It can be seen that the multiplicative and additive viscoplasticity laws satisfactorily predict behaviour of aluminium alloy AA6063-T6 in transient temperature tests of different temperature rates (Fig. 8) and in constant temperature hardening tests of different strain rates (Fig. 7). It must be remembered that the test results at temperature rates 5 K/min and 30 K/min has not been used in the model identification and the model predicts also them (Fig. 8 a).

REFERENCES

1. Harmathy T.Z., Stanzak W.W. 1970. Elevated-Temperature Tensile and Creep Properties of Some Structural and Prestressing Steels, ASTM STP 464. p.186-.
2. Anderberg Y. 1988. Modelling Steel Behaviour, Fire Safety Journal 13 (1988) 17-26.
3. Kirby B.R. and. Preston R. R. 1988, High Temperature Properties of Hot-rolled, Structural Steels for Use in Fire Engineering Studies. Fire Safety Journal 13 (1988) 27-37.
4. Outinen, J. Kesti, J. and Mäkeläinen, P. Fire Design Model for Structural Steel S355 Based Upon Transient State Tensile Test Results. J. Construct. Steel Res. Vol. 42, No 3, pp. 161-169, 1997.
5. Forsen, N. Fire resistance In: Mazzolani F.M. Aluminium Alloy Structures. E & FN Spon. Second edition 1995. pp.657-681.
6. Hammad, A.M. and Ramadan, K.K. Mechanical properties of Al-Mg Alloys at Elevated Temperatures. Z. Metallkde. Bd. 80 (1989) H. 6. pp. 431-438.
7. Zhou, M. and Clode, M.P. Modelling of high temperature viscoplastic flow of aluminium alloys by hot torsion testing. Materials Science and Technology 1997 Vol. 13. Pp. 818-824.
8. Heimerl, G., Inge, J. 1955. Tensile properties of 7075-T6 and 2024-T3 aluminium alloy sheet heated at uniform temperature rates under constant load. NACA TN 3462, Washington 1955.
9. Järvstråt, N. 1994. Calculation of thermally induced stress and distortion in extruded aluminium and short fibre composites. Lindköping Studies in Science and Technology. Dissertations No. 349. Lindköping University. Lindköping 1994.
10. Engström, H., Sandström, R. Evaluation of high temperature strength values of aluminium alloys. ALUMINIUM 69 (1993) 11 pp.1007-1013.
11. Voorhees, H.R., Freeman, J.W. Report on the Elevated-temperature properties of Aluminium and Magnesium alloys. ASTM, STP No. 291 (1960).
12. EUROCODE 9. 1995. Design of aluminium alloy structures Part 1.2 May 1995. Structural fire design. Draft prENV 1999 Part 1.2. CEN/TC 250/SC 9.
13. Groetsch C.W. 1993. Inverse Problems in the Mathematical Sciences, Wieweg Mathematics for Scientists and Engineers, Vieweg, 1993 p.151.
14. Hansen, P.C. 1990. Analysis of discrete ill-posed problems by means of the L-curve. Technical Report MCS-p157-0690. Mathematics and Computer Science Division, Argonne National Laboratory.
15. Lemaitre J and Chaboche J.-L. 1994. Mechanics of solid materials. Cambridge University Press. 1994. pp. 253-345.
16. Tikhonov A.N. and Arsenin V.Y., Solutions of Ill-Posed Problems. Wiley, New York, 1977.

SiC-POHJAISTEN KUUMASUODATTIMIEN LUJUUDESTA

P. Pastila, A.-P. Nikkilä ja T. Mäntylä
Materiaaliopin laitos
Tampereen teknillinen korkeakoulu
PL 598
33101 Tampere

TIIVISTELMÄ

SiC-pohjaisia kuumasuodattimia käytetään uusia polttotekniikoita, kuten paineistettua polttoa ja kaasutusta, hyödyntävissä kombivoimalaitoksissa. Käyttöympäristön korkea lämpötila ja vesihöyry aiheuttavat suodatinmateriaalissa muutoksia, jotka vaikuttavat suodattimen lujuuteen. Käytön aikana suodattimen täytyy kestää puhdistusmenetelmän paineiskuja ja lämpöshokkeja, lämpöjännityksiä sekä matalataajuuksisia värähtelyjä. Tässä esityksessä käsitellään suodattimen mikrorakenteen ja sen käytön aikana tapahtuvan muuttumisen vaikutusta lujuuteen sekä suodattimen vaurioitumismekanismeja. Lujuuden tilastollinen luonne ja hidas särönkasvu ovat tyypillisiä keraamimateriaaleille.

1. JOHDANTO

SiC-pohjaisia kuumakaasusuodattimia käytetään uusia polttotekniikoita hyödyntävissä kombivoimalaitoksissa. Uusilla polttotekniikoilla, kuten paineistettu poltto ja kaasutus, voidaan nostaa fossiilisia polttoaineita käyttävien voimaloiden hyötysuhdetta. Paineistetussa poltossa syntyvä kuumakaasu sisältää kiintoainespartikkeleita ja tuhkaa sekä korrodoivia alkalihöyryjä ja rikkiyhdisteitä [1,2]. Nämä aiheuttavat turbiinin siipien korroosiota ja eroosiota. Jotta päästäisiin taloudellisesti kannattavaan kaasuturbiinin käyttöikään, yli 10 μm:n partikkeleiden ja tuhkan pääsy kaasuturbiinille on estettävä. Tähän on kehitetty korkeassa lämpötilassa ja paineessa toimivia kuumasuodatusmenetelmiä, joista lupaavimmalta vaikuttavat keraamiset kynttiläsuodattimet. Kyntiiläsuodatuksessa kaasua ei tarvitse jäähdyttää puhdistusta varten yhtä paljon kuin nykyisin käytettävissä kaasun pesutekniikoissa. Kuumakaasun suodatuksella voidaan myös yksinkertaistaa voimalaitosprosessia ja pienentää ympäristölle haitallisia päästöjä.

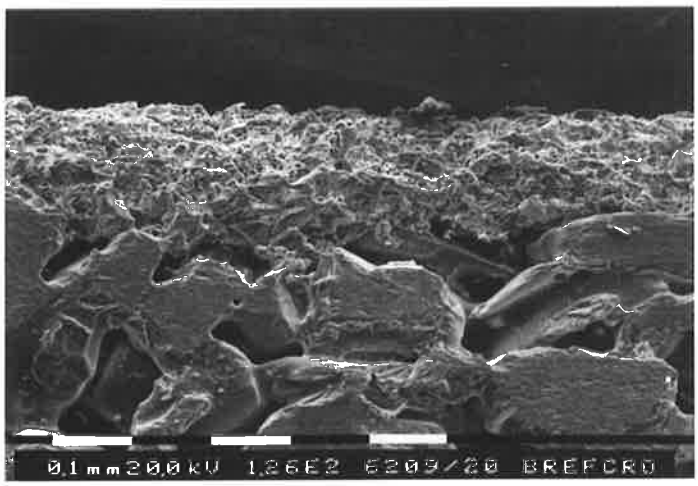
Piikarbidi (SiC) on suosittu suodatinmateriaali sen hyvän korkean lämpötilan lujuuden ja lämmönjohtavuuden vuoksi. Hyvä lämmönjohtavuus parantaa oleellisesti materiaalin lämpöshokinkestävyyttä, mikä puolestaan on oksidipohjaisten suodatinmateriaalien suurin heikkous. Piikarbidin heikkous on sen korroosio hapen ja vesihöyryn vaikutuksesta korkeissa lämpötiloissa. Kuumasuodattimien laajamittaisen kaupallisen käytön esteenä on yhä suodattimien pitkäaikaiskestävyyden ja luotettavuuden varmistaminen. Ongelmana on suodattimien lujuuden muuttuminen käytön aikana korkean käyttölämpötilan,

suodatinmateriaalin korroosion ja puhdistukseen liittyvän termisen syklauksen yhteisvaikutuksesta.

2. SiC-POHJAISET KUUMASUODATTIMET

SiC-pohjaiset kynttiläsuodattimet koostuvat suurihuokoisesta runkomateriaalista ja ohuesta noin 100...200 µm:n paksuisesta suodattavasta membraanipinnasta, kuva 1. Suodattimen huokoisuus on n. 40 %. Runkomateriaali koostuu halkaisijaltaan n. 100...200 µm SiC-rakeista, joita ympäröi ja yhdistää yleensä savipohjainen, alkalipitoinen alumiinisilikaatti-sideaine. Sideaineen koostumus, määrä ja mikrorakenne määräävät suodattimen mekaaniset ominaisuudet.

Kynttiläsuodattimet ovat 1,5 metriä pitkiä, toisesta päästään umpinaisia putkia. Suodatus tapahtuu paineenalaisen kaasun virratessa putken ulkopuolelta sisäpuolelle ja edelleen avonaisesta päästä puhtaan kaasun keräilyputkistoon. Suodatettu kiintoaines kerääntyy putken ulkopinnalle pölykerrokseksi, joka nostaa oleellisesti suodatustehoa. Pölykerroksen ja putken seinämän läpi tapahtuvan painehäviön kasvaessa kriittiseen arvoon suodatin puhdistetaan vastakkaiseen virtaussuuntaan annettavalla kaasupulssilla. Ylipaineinen, huomattavasti suodatettavaa kaasua kylmempi kaasu aiheuttaa lämpöshokin ja lämpöjännityksiä varsinkin putken kiinnityskohdan läheisyyteen ja suodattimen sisäpinnoille. Muulloin suodatettavan kaasun paine aiheuttaa puristusjännityksen suodattimeen. Kynttiläsuodattimet ovat jäykkyytensä ja putkimaisen konstruktionsa vuoksi alttiita värähtelyille, joita aiheuttavat puhdistuspulssit sekä mahdolliset suodatinryhmien liikahdukset.



Kuva 1. Elektronimikroskooppikuva SiC-pohjaisen kuumasuodattimen poikkileikkauksesta. Ylhäällä on suodattimen ulkopinnan membraani ja sen alapuolella suurihuokoista runkoa.

3. KERAAMIMATERIAALIEN LUJUUDESTA

Keraamimateriaalit ovat hauraita, koska plastiseen muodonmuutokseen johtava dislokaatioliike on keraamien atomisidosten luonteen ja monimutkaisten kiderakenteiden vuoksi erittäin vaikeaa. Hauraan materiaalin murtumiskäyttäytymistä voidaan kuvata ns. Griffithin yhtälöllä, jonka mukaan murtumiseen vaadittava jännitystaso on kääntäen verrannollinen vikakoon neliöjuureen [3]:

$$\sigma_f = \sqrt{2E\gamma_o / \pi a}, \qquad (1)$$

missä σ_f = vaadittava jännitys
E = kimmomoduli
γ_o = uuden pinnan muodostamiseen vaadittava pintaenergia
a = alkusärön tai murtuman synnyttävän virheen pituuden puolikas

Keraamikappaleissa on aina valmistuksen jäljiltä alkusäröksi sopivia vikoja, kuten raerajoja tai huokoisuutta, ja kappaleen lujuuden määrää suurin sopivasti kuormitukseen nähden orientoitunut vika. Tämän vuoksi ulkoisesti identtisten koekappaleden lujuudelle saadaan mitattaessa usein suuri hajonta.

Keraamien lujuuskoetuloksia käsitellään yleensä Weibull-statistiikan avulla. Se perustuu heikoimman lenkin teoriaan (weakest link argument) ja oletukselle, että materiaalin vikajakauma on homogeeninen [3]. Weibull yhtälö antaa todennäköisyyden P_s, jolla kappale kestää tietyn jännityksen:

$$P_s = \exp\left[-\int_V \left(\frac{\sigma - \sigma_u}{\sigma_0} \right)^m dV \right] \qquad (2)$$

Yhtälöstä 2 käytetään yleensä ns. kaksiparametrimuotoa, jossa σ_u eli jännitys, jonka kappale varmasti kestää, oletetaan nollaksi. Materiaalin parametri σ_0 on jännitys, jolla materiaali murtuuu 63,2 %:n todennäköisyydellä. Weibull moduli m kuvaa lujuushajonnan suuruutta; mitä suurempi m, sitä pienempi lujuushajonta. On huomattava, että heikoimman lenkin teorian mukaisesti lujuus riippuu vetojännityksen σ kuormittaman materiaalin tilavuudesta. Tämän vuoksi tietylle materiaalille saadaan erilainen lujuus eri kokoisilla koekappaleilla ja eri menetelmillä, kuten esimerkiksi vetokoe, 4-pistetaivutus tai sisäpuolisesti kuormitettu rengaskoe eli IHP-koe (Internal Hydraulic Pressure).

Lineaaris-elastinen murtumismekaniikka soveltuu hyvin hauraille keraameille. Sen mukaan murtuma etenee katastrofaalisesti, kun jännitysintensiteettikertoimen K_I arvo saavuttaa kriittisen arvon K_{IC}, joka vastaa materiaalin murtositkeyttä. Jännitysintensiteettikertoimen arvo riippuu jännitystasosta, kuormituksen tyypistä, särön pituudesta ja kuormitettavan kappaleen sekä särön geometriasta. Murtositkeys on materiaalikohtainen ominaisuus, joka voi riippua särön pituudesta.

Keraameille, varsinkin piidioksidille sekä silikaateille, on tyypillistä hidas murtuman kasvu, joka tarkoittaa murtuman etenemistä myös pienemmillä arvoilla kuin K_{IC}. Tällöin murtuman etenemisnopeus on verraten pieni, jopa alle 10^{-10} m/s, ja murtuman etenemisnopeus riippuu jännitysintensiteettikertoimen arvosta. Hidas murtuman kasvu on seurausta kemiallisten reaktioiden aiheuttamista atomien sidosvoimien pienentymisistä särön kärjessä. Myös mikrosäröjen muodostuminen ja yhdistyminen särön kärjen edessä johtaa hitaaseen murtuman kasvuun varsinkin korkeissa lämpötiloissa. Murtuman etenemisnopeuden v riippuvuus jännitysintensiteettikertoimesta on materiaalikohtainen ominaisuus, jolle on löydetty empiirinen yhteys [3]:

$$v = AK_I^n \tag{3}$$

Hitaalla murtumankasvulla on merkittävä vaikutus materiaalin lujuuteen ja pitkäaikaiskestävyyteen. Viivästyneellä murtumisella tarkoitetaan materiaalin murtumista vakiojännityksen alaisena hitaan murtuman kasvun seurauksena. Ilmiötä on keraameilla tutkittu jo hyvin pitkään nimenomaan laseilla, kun kyseessä on korrodoiva ympäristö, kuten esimerkiksi vesi. Jo huoneenlämpötilassa ilmankosteus aiheuttaa helposti todettavaa hidasta murtuman kasvua keraameilla. Myös keraamien väsyminen vaihtelevan jännityksen alaisena on normaalisti seurausta murtuman hitaasta kasvusta, joka jatkuu, kunnes saavutetaan kriittinen vikakoko. Tällöin murtuma etenee nopeasti läpi koko rakenteen. Terminen väsyminen on materiaalin väsymistä lämpötilanvaihteluiden aiheuttamien jännitysten vuoksi.

Hidasta murtuman kasvua ja sen seurauksena tapahtuvaa viivästynyttä murtumista (keraamien jännityskorroosiota) on tutkittu monilla tärkeistä keraamimateriaaleista. Murtumiskäyttäytymisen on useissa tapauksissa todettu noudattavan seuraavaa riippuvuutta:

$$\sigma^n t_B = C \tag{4}$$

missä σ = maksimijännitystila väsytyksessä
t_B = murtumiseen tarvittava aika tai syklien lukumäärä N
n = murtuman etenemistä kuvaava parametri (väsymisparametri)
C = vakio

Suuri väsymisparametrin arvo (sama parametri yhtälöissä 3 ja 4) tarkoittaa hyvää kestävyyttä hidasta murtuman kasvua vastaan ja siten hyvää väsymiskestävyyttä. Eri materiaalit käyttäytyvät eri tavoin ja merkittäviä eroja löytyy myös saman materiaalityypin sisällä. Tärkeä kestävyyteen vaikuttava tekijä on raerajafaasit. Esimerkiksi piinitrideityypit, joiden raerajafaasi on kiteytynyt valmistuksen yhteydessä, ovat hyvin kestäviä hidasta murtuman kasvua vastaan.

Pitkäaikaiskestävyyden arviointia varten voidaan määrittää SPT (Strength-Possibility-Time)-diagrammeja. Diagrammi voidaan laatia yhdistämällä yhtälöt 2 ja 4, kun tunnetaan materiaalin Weibullin moduli ja murtuman etenemistä kuvaavan parametrin n arvo [3]:

$$\ln \ln (1/(1-F)) = (m/n)(\ln t) + m(\ln \sigma) + C \qquad (5)$$

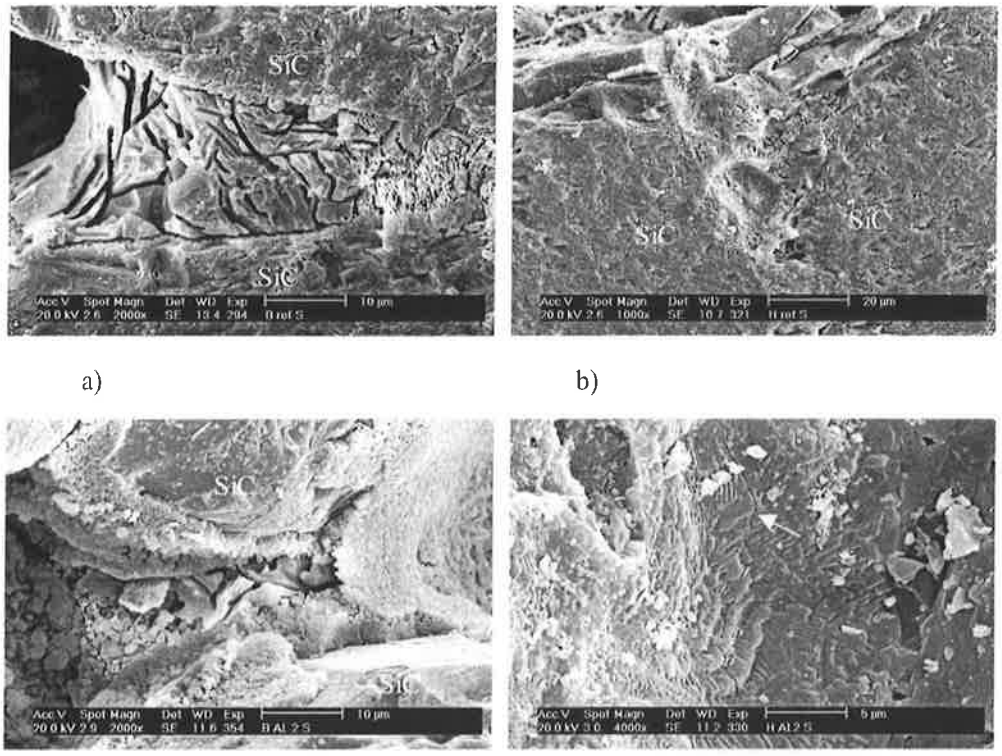
missä F = kappaleen murtumistodennäköisuus
σ = murtumistodennäköisyyttä vastaava lujuus
m = Weibullin moduli
t = murtumiseen kuluva aika tai väsytyskierrosten lukumäärä
C = vakio

Yhtälöstä 5 voidaan todeta, että mitä suurempi parametri n on, sitä vähemmän aika vaikuttaa murtumistodennäköisyyteen.

4. MIKRORAKENNEMUUTOSTEN VAIKUTUS KUUMASUODATTIMIEN LUJUUTEEN

Kuumasuodattimien käyttöympäristö on varsin aggressiivinen. Korkean lämpötilan (500...900°C) ja paineen (10...25 bar) lisäksi suodattimille tuleva polttokaasu sisältää mm. happea, vesihöyryä, rikkiyhdisteitä ja alkalimetallihöyryjä [1]. Suodattimien huokoinen rakenne mahdollistaa korrodoivan väliaineen helpon pääsyn kosketuksiin materiaalin kanssa. Koevoimalaitoksissa tehdyt suodattimien käyttökokeet ja TTKK:n Materiaaliopin laitoksella tehdyt laboratorioaltistukset osoittavat SiC-pohjaisen suodatinmateriaalin mikrorakenteen muuttuvan merkittävästi käytön aikana [2, 4, 5]. Käyttöympäristön happi ja vesihöyry voivat reagoivat SiC:n kanssa ja hapettaa sen lasimaiseksi piidioksidiksi SiO_2. Näin muodostuva tiivis piidioksidilasikerros yleensä hidastaa merkittävästi SiC:n hapettumista, mutta vesihöyry pystyy diffuntoitumaan lasimaisen kerroksen läpi ja ylläpitämään SiC:n hapettumista. Lisäksi korkea lämpötila ja vesihöyry aiheuttavat lasimaisen SiO_2:n kiteytymistä kristobaliitiksi. Lasimaisen kerroksen kiteytyminen johtaa SiC:a ympäröivään kerrokseen halkeiluun, koska kristobaliitin tiheys on suurempi kuin lasimaisen piidioksidin. Suojaavan lasikerroksen halkeilu mahdollistaa SiC:n nopeamman reagoinnin ympäristön kanssa. Myös alkalit helpottavat pieninä määrinä lasimaisen piidioksidin kiteytymistä kristobaliitiksi [6].

Laboratoriokokeissa näytteitä viidestä SiC-pohjaisesta kuumasuodattimesta altistettiin 1000 tuntia 870°C lämpötilassa erilaisille vesihöyryn, ilman, hiilidioksidin, rikkidioksidin ja natriumhöyryn seoksille. Kuvassa 2 on kahden tutkitun suodattimen sideaineen tyypillistä mikrorakennetta ennen altistusta ja altistuksen jälkeen. Suodattimen B sideaine on suurimmaksi osaksi amorfista alumiinisilikaattia, jonka alkalipitoisuus on melko suuri. Suuri alkalipitoisuus laskee amorfisen faasin pehmenemislämpötilaa ja pienentää sen viskositeettia. Edelleen korkea alkalipitoisuus, kuten natrium, pyrkii pitämään lasimaisen alumiinisilikaatin amorfisena. Suodattimen H sideaine taas on lähes kokonaan kiteytynyttä alkalialumiinisilikaattia, ja sen alkalipitoisuus on hyvin pieni.

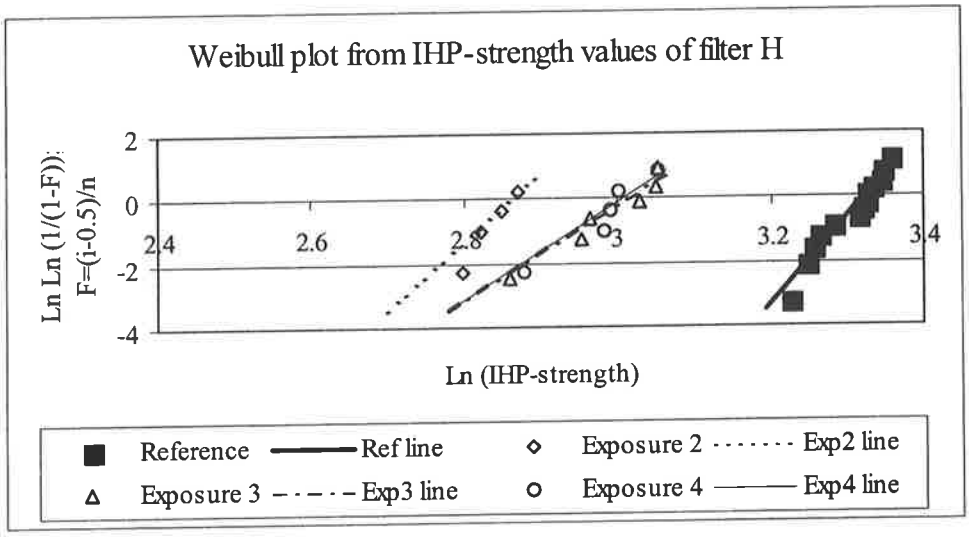


Kuva 2. SiC-pohjaisten kuumasuodattimien sideaineen HF-syövytettyä mikrorakennetta. a) Suodattimen B sideaine on ennen altistusta lähes kokonaan amorfista alkalialumiinisilikaattia. b) Suodattimen H sideaine ennen altistusta on lähes kokonaan kiteistä alkalialumiinisilikaattia. c) Suodattimen B sideainetta altistuksen jälkeen. Sideaineen ulkopinnat kiteytyivät ja SiC:n hapettumisen seurauksena sen pinnalle kasvoi kiteinen kerros. d) Halkeama suodattimen H sideaineen pinnalla altistuksen jälkeen.

Altistuksessa suodattimen A sideaine on kiteytyi ulkopinnoiltaan pääasiassa kristobaliitiksi ja alkalialumiinisilikaateiksi. Piikarbidirakeet hapettuivat altistuksessa piidioksidiksi, joka kiteytyi kristobaliitiksi SiC:n pinnalle. Myös suodattimen B sideaineessa tapahtui altistuksessa kiteytymistä ja sen seurauksena sideaineen ulkopinnoille syntyi säröjä.

Tutkittujen suodattimien huoneenlämpötilassa mitattu lujuus laski noin 30% altistuksien vuoksi. Suoritetut altistukset ovat siten joko kasvattaneet suodattimen mikrorakenteessa olevia suurimpia vikoja tai mahdollisesti pienentäneet materiaalin murtopintaenergiaa. Muutokset ovat tapahtuneet sideaineessa ja sideaineen ja SiC-rakeen rajapinnalla. Vikakoon kasvu johtuu altistuksessa tapahtuneesta lasimaisen lähes puhtaan piidioksidin kiteytymisestä ja siihen liittyvistä tilavuuden muutoksista, jotka aiheuttavat säröjä materiaaliin. Kristobaliitilla esiintyy faasimuutos lämpötilavälillä 200...270°C, johon

liittyy 3%:n tilavuuden pieneneminen jäähdytettäessä [7]. Sen lisäksi kristobaliitin lämpölaajenemiskerroin on suurempi kuin SiC:n tai amorfisen piidioksidin [8]. Pintaenergian muuttuminen liittyy sideaineen kemiallisen koostumuksen muutoksiin ja uusien raerajojen syntymiseen. Nämä tekijät yhdessä selittävät suodattimien lujuuden laskua. Myös hajonta lujuusarvoissa kasvoi altistuksen vuoksi, mikä olisi jo ilman lujuuden merkittävää laskuakin pitkäaikaiskäytön kannalta huono asia. Tämä näkyy kuvan 3 Weibull-kuvaajassa mittausdataan sovitetun suoran kulmakertoimen eli Weibull-modulin m arvon pienenemisenä. Suodatinnäytteiden lujuus altistusten jälkeenkin oli kuitenkin riittävä, jotta ne kestäisivät kuumasuodatuksen rasitukset. Lyhytaikaiset altistuskokeet ilman kuormitusta eivät kuitenkaan anna tietoa materiaalin pitkäaikaiskestävyydestä hidasta murtumankasvua vastaan.



Kuva 3. Weibull-kuvaaja suodattimen H IHP-menetelmällä ennen altistusta ja altistusten jälkeen tehtyjen lujuuskokeiden tuloksista. Mittausdataan maximum likelihood-menetelmällä sovitetun kuvaajan kulmakerroin antaa Weibull-modulin. Kuvaajan ja x-akselin leikkauskohta antaa jännitystason ($\ln\sigma$), jolla koekappaleen murtumistodennäköisyys on 63,2%.

Hitaan murtumankasvun nopeuden määritystä vaikeuttaa suodattimien mikrorakenteen jatkuva muuttuminen ja samalla esimerkiksi murtositkeyden muuttuminen. Hidas särönkasvu on mahdollista sekä mikromurtumien muodostumisen ja yhdistymisen mekanismilla että kemiallisen reaktion seurauksena. Mikromurtumien syntymisen ja kasvamisen kannalta SiC:n hapettumiseen ja amorfisten faasien kiteytymiseen liittyvät paikalliset jännityskeskittymät ovat merkittäviä. Erot lämpölaajenemiskertoimissa voivat aiheuttaa vetojännityksiä suodattimen puhdistussyklien yhteydessä. Käyttöympäristön vesihöyry voi reagoida silikaatisideaineen kanssa aiheuttaen hidasta särönkasvua. Näin SiC-pohjaisten kuumasuodattimien vaurioituminen on monimutkainen prosessi, jonka

arviointi edellyttää käytön aikaisten kuormitusten tarkkaa analysointia ja muuttuvien materiaaliominaisuuksien tuntemista.

5. LÄHTEET

1. J. E. Oakey ja I. R. Fantom, Hot gas cleaning - materials and performance. Materials at high temperature, 1997, **14**, 337-345.

2. M. A. Alvin, Performance and stability of porous ceramic candle filters during PFBC operation. Materials at high temperatures, 1997, **14**, 355-364.

3. J. B. Wachtman, Mechanical properties of ceramics. Jhon Wiley &Sons, 1996.

4. J. E. Oakey, T. Lowe, R. Morrel, W. P. Byrne, R. Brown ja J. Stinger, Grimethorpe filter element performance - the final analysis. Materials at high temperatures, 1997, **14**, 371-381.

5. A.-P. Nikkilä, V. Helanti, P. Pastila ja T. Mäntylä, Keraamisten kuumasuodattimien pitkäaikaiskestävyys ja käyttöluotettavuus. Raportti 15/ 1997, TTKK Materiaalioppi.

6. W. A. Weyl ja E. C. Marboe, The constitution of glasses, A dynamic interpretation, Voulume II. Jhon Wiley&Sons, 1967.

7. R. W. Chan, P. Haasen ja E. J. Kramer, Materials Science and Technology, Volume 11, Structure and properties of ceramics. VCH, 1994.

8. N. S. Jacobson, Corrosion of silicon-based ceramics in combustion environments. Journal of the American Ceramic Society, 1993 **76**, 3-28.

THERMOMECHANICAL FORMULATION FOR THE GRADIENT THEORY

K. SANTAOJA
Laboratory for Mechanics of Materials
Helsinki University of Technology
P.O.Box 4100
FIN-02015 HUT, FINLAND

ABSTRACT

This work demonstrates how gradients of internal state variables can be used in a set of internal variables when the thermomechanics of internal state variables is utilised. This is done here by introducing a thermal internal variable called specific dissipative entropy. To demonstrate the derived new concept, a material model showing Hookean deformation and primary and secondary creep is evaluated. Damage affects the Hookean response of the material. The damage evolution equation contains a term in which the Laplacian term for damage is present. This term is introduced to avoid localisation of damage and mesh-dependence of a finite element solution. The potential for utilising an ABAQUS coupled thermal-stress solution procedure to solve material models with the Laplacian term in the damage evolution equation is discussed.

1. INTRODUCTION

Strain softening has been observed in many materials such as concrete, rocks, ice, metallic materials at elevated temperatures etc. In order to model strain softening, usually a new variable called damage and denoted by D is introduced. Damage D does not model any specific deformation mechanism, but is a general concept for strain softening. In the high-temperature materials of power plant components, where the service temperature is around 550 °C, damage D may be nucleation, growth and coalescence of voids on grain boundaries or carbide coarsening. In ice, by comparison, damage is related to microcracking.

The standard form of continuum mechanics is a local theory. This means that the functions are dependent only on the values of the variables at the same material points. For Hooke's law this is as follows: stress σ depends only on the value of the strain ε at that specific point, i.e. $\sigma = E\,\varepsilon$. The problem with local theories is that when strain softening is modelled they lead to stability problems and to mesh-dependency in finite element analysis. Non-local theories solve these problems. In non-local theories (some) functions are not

only dependent on the values of the variables at that particular point but also on the values of the variables in the neighbourhood of the investigated point. As mentioned by Bazant (1994, p. 593)[1], finite element analysis of distributed strain-softening damage, including its final localization into a sharp fracture, requires the use of some type of non-local continuum.

Non-local continuum theories have a material length scale. According to Ramaswamy and Aravas (1998, pp. 11 and 12)[2] there is a variety of methods which have introduced a length scale. In micropolar continuum theories the rotational degrees of freedom are added to the conventional translational degrees of freedom. In integral-type constitutive models the evolution of certain internal variables is expressed by means of integral equations. In the gradient-type of plasticity models, for example, the Laplacian value of the effective plastic strain ε^v can be a vital part of the yield function. A counterpart gradient-type of a constitutive equation for the above Hooke's law might be $\sigma = E\,\varepsilon + k\nabla^2\varepsilon$ where the notation ∇^2 stands for the Laplacian operator.

This work evaluates material models within the framework of thermomechanics with internal variables. It has an ingredient beyond the standard formulation: The gradient theory. Thermomechanics is a theory that combines thermodynamics and continuum mechanics. The formulation of thermomechanics with internal variables introduces variables beyond controllable ones present in the basic laws of thermomechanics. As already mentioned, these variables are called internal variables. Usually, internal variables describe the mechanical part of the response of a material, i.e. deformation. Internal variables are scalars or tensors of any order. Sometimes the material time derivatives of internal variables belong to the set of internal variables. In the gradient theory, however, also the gradients of internal variable(s) are a vital part of the set of internal variables. This work demonstrates how gradients can be implemented into the thermomechanics of internal variables. This is done by introducing a thermal internal variable called the specific dissipative entropy.

At the end of the paper, the implementation of a material model with a gradient term into the ABAQUS finite element method computer program is discussed. The present method utilises the coupled thermal stress analysis present in the gallery of analysis procedures of ABAQUS.

2. THERMOMECHANICS WITH INTERNAL VARIABLES

The present chapter evaluates the problems which originates when the classical approach to thermomechanics with internal variables is used with a gradient term in the list of internal variables. This chapter follows the work by Santaoja (2000, Chapt. 3)[3] although that work used a Laplacian of an internal variable in the list of state variables.

Controllable state variables for thermomechanical processes in deformable solids are:

$\boldsymbol{\varepsilon}$ The strain tensor. This is a second-order tensor describing both mechanical and thermal deformation.

s Specific entropy.

The internal state variables and their form are determined by the material model under consideration. For the sake of simplicity this work studies the Maxwell type of material models, which include the inelastic strain tensor $\boldsymbol{\varepsilon}^{\mathrm{i}}$ in the set of internal state variables and also assume that the difference $\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^{\mathrm{i}}$ describes the state of a material point. The present work studies a special case of gradient theory. Traditionally, when strain softening has been studied with the theory of plasticity the higher-order gradients of inelastic strain tensor $\boldsymbol{\varepsilon}^{\mathrm{i}}$ have been used in the yield functions [e.g. Aifantis (1992)][4]. This has been done to avoid localisation [e.g. Kukudžanov et al. (1995, pp. 2-6)][5]. Here a different approach is used, where the gradient operator $\vec{\nabla}$ is acting on the internal state variable. This work follows the popular notation and assumes that this particular variable is a scalar-valued quantity denoted by D. Its gradient is $\vec{\nabla}D$. The quantity D is only an example and can therefore be replaced by a vector or a tensor of any order.

The state $\mathscr{E}$ of the material is dependent on the above-mentioned independent state variables. Therefore they are called state variables. Functions whose arguments are state variables are called state functions. According to the caloric equation of state, the state $\mathscr{E}$ is expressed by the specific internal energy $u(\vec{x},t)$ through the caloric equation of state. Thus the state $\mathscr{E}$ is expressed by

$$u = u(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^{\mathrm{i}}, D, \vec{\nabla}D, \boldsymbol{\alpha}, s), \tag{1}$$

where $\boldsymbol{\alpha}$ represents the set of other mechanical internal state variables. The set $\boldsymbol{\alpha}$ can contain scalars, vectors or tensors of any order. Instead of the specific internal energy u, the state of solids is usually described by the specific Helmholtz free energy ψ which is a Legendre partial transformation of the specific internal energy u. This is done because writing a material model using the specific internal energy u is very difficult, given that the specific entropy s is an argument of the specific internal energy u. It is very difficult to construct a constitutive model as a function of the specific entropy s. In the formulation of the specific Helmholtz free energy ψ the specific entropy s is replaced by the absolute temperature T. Since the absolute temperature T is a well-known quantity for a human being, writing a material model using the specific Helmholtz free energy ψ is much simpler than doing so with the specific internal energy u.

State functions are obtained as partial derivatives of the specific internal energy u with respect to the state variables. Due to the introduction of the specific Helmholtz free energy ψ state equations take the forms

$$\boldsymbol{\sigma} := \rho \frac{\partial \psi(\ldots)}{\partial(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^{\mathrm{i}})} \qquad \textit{and} \qquad v := -\rho \frac{\partial \psi(\ldots)}{\partial D} \tag{2}$$

and further

$$\vec{\zeta} := -\rho \frac{\partial \psi(\ldots)}{\partial \vec{\nabla} D} \qquad \textit{and} \qquad \boldsymbol{\beta} := -\rho \frac{\partial \psi(\ldots)}{\partial \boldsymbol{\alpha}} \tag{3}$$

and finally

$$s := -\frac{\partial \psi(\ldots)}{\partial T} \qquad \textit{where} \qquad \psi = \psi(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^{\mathrm{i}}, D, \vec{\nabla} D, \boldsymbol{\alpha}, T). \tag{4}$$

In State Equations (2) and (3) the notation $\boldsymbol{\sigma}$ stands for the stress tensor, v, $\vec{\zeta}$ and $\boldsymbol{\beta}$ are internal forces and ρ is the density.

The principle of conservation of energy, also referred to as the first law of thermodynamics, can be stated as follows: The time rate of change of the sum total of the kinetic energy K and the internal energy U in the body is equal to the sum of the rates of work done by the surface and body loads in producing the deformation (or flow) together with heat energy that may leave or enter the body at a certain rate. Thus the following is obtained:

$$\frac{\mathrm{D}}{\mathrm{D}t}(K + U) = P^{\mathrm{ext}} + Q. \tag{5}$$

In Basic Law (5) P^{ext} is the power input of the external forces and Q is the heat input rate. The local form for the first law of thermodynamics is called the energy equation (in the non-polar case) or the equation of balance of energy. It has the following form:

$$\rho \dot{u} = \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} + \rho r - \vec{\nabla} \cdot \vec{q}, \tag{6}$$

where r is the heat source per unit mass and where $\vec{q}$ is the heat flux vector. The second law of thermodynamics can be written in the form

$$\dot{S} \geq -\oint_{\partial V} \frac{\vec{n} \cdot \vec{q}}{T} \, \mathrm{d}A + \int_V \rho \frac{r}{T} \, \mathrm{d}V, \tag{7}$$

where $\dot{S}$ is the entropy rate and $\vec{n}$ is the outward unit vector for volume V, the surface of which is denoted by ∂V. The local form of the second law of thermodynamics takes the form

$$\rho T \dot{s} + \vec{\nabla} \cdot \vec{q} - \frac{\vec{\nabla} T}{T} \cdot \vec{q} - \rho r \geq 0. \tag{8}$$

The internal energy U and the entropy S are defined by

$$U := \int_V \rho u \, \mathrm{d}V \qquad \textit{and} \qquad S := \int_V \rho s \, \mathrm{d}V. \tag{9}$$

Combination of the local forms of the first and second law of thermodynamics, i.e. Expressions (6) and (8), is called the Clausius-Duhem inequality. For the present set of state variables [see Expression (4b)] it takes the following form:

$$\boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}^{\mathrm{i}} + v \dot{D} + \vec{\zeta} \cdot \vec{\nabla} \dot{D} + \boldsymbol{\beta} : \dot{\boldsymbol{\alpha}} - \frac{\vec{\nabla} T}{T} \cdot \vec{q} \geq 0. \tag{10}$$

Based on the Clausius-Duhem Inequality (10) the specific entropy production rate $\dot{s}^{i}$ is introduced. It is defined by

$$\rho T \dot{s}^{i} := \boldsymbol{\sigma}:\dot{\boldsymbol{\varepsilon}}^{i} + v\dot{D} + \vec{\zeta}\cdot\vec{\nabla}\dot{D} + \boldsymbol{\beta}:\dot{\boldsymbol{\alpha}} - \frac{\vec{\nabla}T}{T}\cdot\vec{q} \quad (\geq 0). \tag{11}$$

The principle of maximal rate of entropy production by Ziegler (1963, Chapt. 4)[6] is used to obtain the normality rule. The mechanical part of the problem [i.e. Expression (11)] is considered. The principle of maximal rate of entropy production is written in the following mathematical form:

maximise with respect to the fluxes $(\dot{\boldsymbol{\varepsilon}}^{i}, \dot{D}, \vec{\nabla}\dot{D}, \dot{\boldsymbol{\alpha}})$

$$\dot{s}^{i}_{loc} = \frac{1}{\rho T}\left(\boldsymbol{\sigma}:\dot{\boldsymbol{\varepsilon}}^{i} + v\dot{D} + \vec{\zeta}\cdot\vec{\nabla}\dot{D} + \boldsymbol{\beta}:\dot{\boldsymbol{\alpha}}\right) \tag{12}$$

subject to:

$$\tau_{loc} = \frac{1}{T}\varphi_{loc} - \frac{1}{\rho T}\left(\boldsymbol{\sigma}:\dot{\boldsymbol{\varepsilon}}^{i} + v\dot{D} + \vec{\zeta}\cdot\vec{\nabla}\dot{D} + \boldsymbol{\beta}:\dot{\boldsymbol{\alpha}}\right) = 0, \tag{13}$$

where $\tau_{loc} = 0$ is a constraint and where

$$\varphi_{loc} = \varphi_{loc}(\dot{\boldsymbol{\varepsilon}}^{i}, \dot{D}, \vec{\nabla}\dot{D}, \dot{\boldsymbol{\alpha}}; \boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^{i}, D, \vec{\nabla}D, \boldsymbol{\alpha}, T) \tag{14}$$

is the specific dissipation function for mechanical behaviour.

The specific dissipation function φ_{loc} (the subscript loc refers to mechanical behaviour) is defined by

$$\varphi_{loc} = T\dot{s}^{i}_{loc} \quad \Rightarrow \quad \frac{1}{T}\varphi_{loc} - \dot{s}^{i}_{loc} = 0 \quad \textit{for an actual process}. \tag{15}$$

Instead of the specific dissipation function φ_{loc} this work utilises the specific complementary dissipation function φ^{c}_{loc}. It is a Legendre partial transformation of the specific dissipation function φ_{loc} and it is expressed as

$$\varphi^{c}_{loc} = \varphi^{c}_{loc}(\boldsymbol{\sigma}, v, \vec{\zeta}, \boldsymbol{\beta}; \boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^{i}, D, \vec{\nabla}D, \boldsymbol{\alpha}, T) \tag{16}$$

Now the mechanical part of the normality rule takes the form

$$\dot{\boldsymbol{\varepsilon}}^{i} = \rho\frac{\partial\varphi^{c}_{loc}(\ldots)}{\partial\boldsymbol{\sigma}} \quad \textit{and} \quad \dot{D} = \rho\frac{\partial\varphi^{c}_{loc}(\ldots)}{\partial v} \tag{17}$$

and

$$\vec{\nabla}\dot{D} = \rho\frac{\partial\varphi^{c}_{loc}(\ldots)}{\partial\vec{\zeta}} \quad \textit{and} \quad \dot{\boldsymbol{\alpha}} = \rho\frac{\partial\varphi^{c}_{loc}(\ldots)}{\partial\boldsymbol{\beta}}. \tag{18}$$

There is a serious problem with Form (10) of the Clausius-Duhem inequality, and this is discussed next. If the gradient operator $\vec{\nabla}$ acted on Equation (17b) from the left, the left-

hand sides of Equations (17b) and (18a) would be identical. This means that also the right-hand sides of Equations (17b) and (18a) should be the same. This condition is

$$\left(\frac{1}{\rho}\,\vec{\nabla}\dot{D} = \right) \qquad \vec{\nabla}\,\frac{\partial \varphi^{c}_{loc}(\ldots)}{\partial v} = \frac{\partial \varphi^{c}_{loc}(\ldots)}{\partial \vec{\zeta}}. \tag{19}$$

Condition (19) is too restrictive for preparation of a material model expressed by potential $\varphi_{loc}(\dot{\boldsymbol{\varepsilon}}^{i}, \dot{D}, \vec{\nabla}\dot{D}, \dot{\boldsymbol{\alpha}}; \boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^{i}, D, \vec{\nabla}D, \boldsymbol{\alpha}, T)$. It is most difficult, even impossible, to write any specific complementary dissipation function $\varphi^{c}_{loc}(\boldsymbol{\sigma}, v, \vec{\zeta}, \boldsymbol{\beta}; \boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^{i}, D, \vec{\nabla}D, \boldsymbol{\alpha}, T)$ which would both satisfy Condition (19) and provide an acceptable material model. Therefore Clausius-Duhem Inequality (10) [actually Expression (11)] has to be rewritten in a form where both $\dot{D}$ and $\vec{\nabla}\dot{D}$ are not present. This is done in the following chapter.

3. THERMOMECHANICS WITH INTERNAL VARIABLES WITH GRADIENTS

This chapter shows that the introduction of the specific dissipative entropy s^{d} solves the problems caused by the gradient term in the set of internal state variables.

In order to avoid the above-described problems when the gradient term of an internal state variable is introduced, a new internal state variable s^{d} is introduced. The notation s^{d} stands for the specific dissipative entropy. It is the difference between the values of the specific entropy in non-equilibrium and equilibrium. It is a thermal internal state variable. Instead of Expression (1) the state $\mathscr{E}$ is now expressed by

$$u = u(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^{i}, D, \vec{\nabla}D, \boldsymbol{\alpha}, s - s^{d}). \tag{20}$$

State Equation (4) is now replaced by

$$s - s^{d} := -\frac{\partial \psi(\ldots)}{\partial T}. \tag{21}$$

The other state equations and the form for the specific Helmholtz free energy ψ keep their forms, viz. Forms (2), (3) and (4b). Compared to Form (10) the Clausius-Duhem inequality introduces one new term and therefore has the following appearance:

$$\boldsymbol{\sigma}:\dot{\boldsymbol{\varepsilon}}^{i} + v\dot{D} + \vec{\zeta}\cdot\vec{\nabla}\dot{D} + \boldsymbol{\beta}:\dot{\boldsymbol{\alpha}} + \rho T\dot{s}^{d} - \frac{\vec{\nabla}T}{T}\cdot\vec{q} \geq 0. \tag{22}$$

Following a concept with similar elements to those introduced by Maugin [1990, Eq.(3.5)] [7] the following can be written:

$$\vec{\nabla}\cdot(\vec{\zeta}\dot{D}) = (\vec{\nabla}\cdot\vec{\zeta})\dot{D} + \vec{\zeta}\cdot\vec{\nabla}\dot{D}. \tag{23}$$

Substitution of Manipulation (23) into Expression (22) gives

$$\boldsymbol{\sigma}:\dot{\boldsymbol{\varepsilon}}^{i} + v\dot{D} - (\vec{\nabla}\cdot\vec{\zeta})\dot{D} + \vec{\nabla}\cdot(\vec{\zeta}\dot{D}) + \boldsymbol{\beta}:\dot{\boldsymbol{\alpha}} + \rho T\dot{s}^{d} - \frac{\vec{\nabla}T}{T}\cdot\vec{q} \geq 0. \tag{24}$$

The following notation and definition are introduced:

$$\eta := v - \vec{\nabla}\cdot\vec{\zeta} \qquad \text{and} \qquad \rho T \dot{s}^{d} := -\vec{\nabla}\cdot(\vec{\zeta}\dot{D}). \tag{25}$$

Notation (25a) and Definition (25b) allow Expression (24) to be rewritten in the form

$$\boldsymbol{\sigma}:\dot{\boldsymbol{\varepsilon}}^{i} + \eta\dot{D} + \boldsymbol{\beta}:\dot{\boldsymbol{\alpha}} - \frac{\vec{\nabla}T}{T}\cdot\vec{q} \geq 0. \tag{26}$$

Clausius-Duhem Inequality (26) has the form which was aimed at. It is the sum of force times flux pairs and it does not contain the restriction the description for the material model described in Equation (19). The normality rule takes the form

$$\dot{\boldsymbol{\varepsilon}}^{i} = \rho\frac{\partial\varphi^{c}_{loc}}{\partial\boldsymbol{\sigma}} \qquad \text{and} \qquad \dot{D} = \rho\frac{\partial\varphi^{c}_{loc}}{\partial\eta} \tag{27}$$

and furthermore

$$\dot{\boldsymbol{\alpha}} = \rho\frac{\partial\varphi^{c}_{loc}}{\partial\boldsymbol{\beta}} \qquad \text{where} \qquad \varphi^{c}_{loc} = \varphi^{c}_{loc}(\boldsymbol{\sigma},\eta,\boldsymbol{\beta};\boldsymbol{\varepsilon}-\boldsymbol{\varepsilon}^{i},D,\vec{\nabla}D,\boldsymbol{\alpha},T). \tag{28}$$

4. EXAMPLE

As an example, a material model showing Hooke's law and Le Gac and Duval (1980)[8] type of creep behaviour at a virgin state is demonstrated. Material damage is traditionally described by the "effective stress concept" and by a scalar-valued damage quantity denoted by D. The damage evolution equation contains a gradient term (actually a Laplacian term) to avoid localisation.

When thermomechanics is applied for the modelling of materials, explicit forms for two functions have to be given. These functions are the specific Helmholtz free energy ψ and the specific dissipation function φ. In this work the latter is replaced by the specific complementary dissipation function φ^{c}_{loc} where the subscript loc refers to the mechanical part of the problem. The thermomechanical formulation for the mechanical part of the material model is

$$\psi = \frac{1}{2\rho}(1-D)(\boldsymbol{\varepsilon}-\boldsymbol{\varepsilon}^{i}):\boldsymbol{C}:(\boldsymbol{\varepsilon}-\boldsymbol{\varepsilon}^{i}) + \frac{k}{4\rho a}(\vec{\nabla}D)\cdot(\vec{\nabla}D) + \frac{1}{2\rho}\boldsymbol{\alpha}:\mathbf{C}^{1}:\boldsymbol{\alpha} \tag{29}$$

and

$$\varphi^{c}_{loc} = \frac{a}{\rho}\eta^{2} + \frac{2\dot{\varepsilon}_{re}\sigma_{re}}{3\rho(n+1)}\left[\frac{\langle J_{vM}(\boldsymbol{\sigma}-\boldsymbol{\beta}^{1})\rangle}{\sigma_{re}}\right]^{(n+1)} + \frac{2\tilde{C}^{11}\sigma_{re}}{3\rho(k+1)}\left[\frac{J_{vM}(\boldsymbol{\beta}^{1})}{\sigma_{re}}\right]^{(k+1)}. \tag{30}$$

In Model (29) $\mathbf{C}$ stands for the constitutive tensor for Hookean material, a is a material parameter and $\mathbf{C}^{1}$ is a fourth-order constitutive tensor. In Model (30) $\dot{\varepsilon}_{re}$ is the reference strain rate, σ_{re} is the reference stress and n, $\tilde{C}_{11}$ and k are material constants. The notations $J_{vM}(\,)$ and $\langle\,\rangle$ stand for the von Mises operator and for the Macaulay brackets, respectively. When acting on a second-order tensor $\boldsymbol{\gamma}$, the deviatoric tensor of which is $\mathbf{g}$, the von Mises

operator $J_{vM}(\,)$ gives

$$J_{vM}(\boldsymbol{\gamma}) \;:=\; \sqrt{\tfrac{3}{2}\,\mathbf{g}:\mathbf{g}}. \tag{31}$$

The Macaulay brackets $\langle\ \rangle$ have the following property:

$$\langle x\rangle \;:=\; \begin{cases} 0 & \textit{when}\quad x < 0 \\ x & \textit{when}\quad x \geq 0\ . \end{cases} \tag{32}$$

Substitution of Material Model (29) into State Equations (1) and (2) yields

$$\boldsymbol{\sigma} \;:=\; \rho\,\frac{\partial \psi(\ldots)}{\partial(\boldsymbol{\varepsilon}-\boldsymbol{\varepsilon}^{i})} \;=\; (1-D)\,\boldsymbol{C}:(\boldsymbol{\varepsilon}-\boldsymbol{\varepsilon}^{i}) \tag{33}$$

and

$$\upsilon \;:=\; -\rho\,\frac{\partial \psi(\ldots)}{\partial D} \;=\; \frac{1}{2}\,(\boldsymbol{\varepsilon}-\boldsymbol{\varepsilon}^{i}):\boldsymbol{C}:(\boldsymbol{\varepsilon}-\boldsymbol{\varepsilon}^{i}) \tag{34}$$

and further

$$\vec{\zeta} \;:=\; -\rho\,\frac{\partial \psi(\ldots)}{\partial \vec{\nabla} D} \;=\; -\,\frac{k}{2\,a}\,(\vec{\nabla} D). \tag{35}$$

and finally

$$\boldsymbol{\beta} \;:=\; -\rho\,\frac{\partial \psi}{\partial \boldsymbol{\alpha}} \;=\; -\mathbf{C}^{I}:\boldsymbol{\alpha} \qquad \textit{which yields} \qquad \dot{\boldsymbol{\beta}} \;=\; -\mathbf{C}^{I}:\dot{\boldsymbol{\alpha}} \tag{36}$$

Substitution of Forces (34) and (35) into Definition (25a) gives

$$\eta \;:=\; \upsilon - \vec{\nabla}\cdot\vec{\zeta} \;=\; \frac{1}{2}\,(\boldsymbol{\varepsilon}-\boldsymbol{\varepsilon}^{i}):\boldsymbol{C}:(\boldsymbol{\varepsilon}-\boldsymbol{\varepsilon}^{i}) \;+\; \frac{k}{2\,a}\,(\nabla^{2} D). \tag{37}$$

Substitution of Model (30) into Normality Rule (27b) provides

$$\dot{D} \;=\; \rho\,\frac{\partial \varphi^{c}_{loc}}{\partial \eta} \;=\; 2\,a\,\eta. \tag{38}$$

Substituting Expression (37) into Rate (38) gives

$$\dot{D} \;=\; a\,(\boldsymbol{\varepsilon}-\boldsymbol{\varepsilon}^{i}):\boldsymbol{C}:(\boldsymbol{\varepsilon}-\boldsymbol{\varepsilon}^{i}) \;+\; k\,(\nabla^{2} D). \tag{39}$$

Substitution of Model (30) into Normality Rule (27a) and (28a) give

$$\dot{\boldsymbol{\varepsilon}}^{i} \;=\; \rho\,\frac{\partial \varphi^{c}_{loc}}{\partial \boldsymbol{\sigma}} \;=\; \dot{\varepsilon}_{re}\left[\frac{\left\langle J_{vM}(\boldsymbol{\sigma}-\boldsymbol{\beta})\right\rangle}{\sigma_{re}}\right]^{n} \frac{\mathbf{s}-\mathbf{b}}{J_{vM}(\boldsymbol{\sigma}-\boldsymbol{\beta})}, \tag{40}$$

and

$$\dot{\boldsymbol{\alpha}} = \rho \frac{\partial \varphi^{c}_{loc}}{\partial \boldsymbol{\beta}} = \dot{\boldsymbol{\varepsilon}}^{i} + \tilde{C}^{11} \left[\frac{J_{vM}(\boldsymbol{\beta})}{\sigma_{re}} \right]^{k} \frac{\mathbf{b}}{J_{vM}(\boldsymbol{\beta})}. \tag{41}$$

In Rate Equations (40) and (41) **s** and **b** are the deviatoric parts of the stress tensor $\boldsymbol{\sigma}$ and the internal force $\boldsymbol{\beta}$, respectively. Substitution of Rate (41) into Expression (36b) leads to

$$\dot{\boldsymbol{\beta}} = m^{1} \dot{\boldsymbol{\varepsilon}}^{i} - m^{2} \left[\frac{J_{vM}(\boldsymbol{\beta})}{\sigma_{re}} \right]^{k} \frac{\mathbf{b}}{J_{vM}(\boldsymbol{\beta})}, \tag{42}$$

where the following was assumed:

$$C^{1} = m^{1} \mathbf{I} \qquad \textit{and} \qquad \tilde{C}^{11} = \frac{m^{2}}{m^{1}}. \tag{43}$$

In Expression (43) the notation **I** stands for the fourth-order identity tensor and m^1 and m^2 are material parameters.

5. POTENTIAL NUMERICAL SOLUTION PROCEDURE

This chapter sketches a potential concept for solving problems with material models having a Laplacian term in their damage evolution law.

The material model derived in the previous chapter is given by Expressions (33), (39), (40) and (42). They are

$$\boldsymbol{\sigma} = (1 - D)\, C : (\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^{i}) \tag{44}$$

and

$$\dot{\boldsymbol{\varepsilon}}^{i} = \dot{\varepsilon}_{re} \left[\frac{\left\langle J_{vM}(\boldsymbol{\sigma} - \boldsymbol{\beta}) \right\rangle}{\sigma_{re}} \right]^{n} \frac{\mathbf{s} - \mathbf{b}}{J_{vM}(\boldsymbol{\sigma} - \boldsymbol{\beta})} \tag{45}$$

and further

$$\dot{\boldsymbol{\beta}} = m^{1} \dot{\boldsymbol{\varepsilon}}^{i} - m^{2} \left[\frac{J_{vM}(\boldsymbol{\beta})}{\sigma_{re}} \right]^{k} \frac{\mathbf{b}}{J_{vM}(\boldsymbol{\beta})} \tag{46}$$

and finally

$$\dot{D} = a (\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^{i}) : C : (\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^{i}) + k (\nabla^{2} D). \tag{47}$$

For example, the ABAQUS finite element method computer program has a coupled thermal-stress analysis procedure which makes it possible to solve simultaneously a mechanical and a thermal problem which are mutually dependent. Furthermore, both mechanical and thermal problems can be described by the user's written subroutines. Thus the above-obtained constitutive equations are solved as follows: The Expressions (44)...(46) are solved as a mechanical part of the problem and a material subroutine UMAT is coded for that purpose. The heat equation with Fourier's law of heat conduction takes

the form

$$\dot{T} = \frac{r}{C} + \frac{1}{\rho C}\vec{\nabla}\cdot\left(\gamma(\dot{T})\,\vec{\nabla}T\right). \tag{48}$$

where r is the heat source per unit mass and where C is the specific heat. If the spatial derivative of the thermal conductivity $\gamma(\dot{T})$, i.e. $\vec{\nabla}\gamma(\dot{T})$, is so small compared to the other terms of Expression (48) that it can be neglected as a small quantity, Expression (48) takes the following form:

$$\dot{T} = \frac{r}{C} + \frac{\gamma(\dot{T})}{\rho C}\nabla^2 T. \tag{49}$$

According to ABAQUS TM [Version 5.8, Eq. (2.11.1-5)][9], ABAQUS applies Form (49) instead of Expression (48), although the conductivity γ can be dependent of many field variables besides temperature T ABAQUS UM [Version 5.8, p. 23.3.20-1][10]. Thus ABAQUS uses Form (49) instead of Form (48) although it is theoretically incorrect due to high values of the terms $\vec{\nabla}\gamma$.

Comparison of Rate Equations (47) and (49) shows that they have the same form and therefore Rate Equation (47) can be solved as a thermal part of the coupled thermal-stress problem.

6. CONCLUSIONS

This work demonstrated the problems which will result from the introduction of a gradient of an internal variable into the list of internal variables when the thermomechanics of internal variables is applied. As shown, these problems can be solved by a new concept which introduces a new internal state variable called dissipative entropy s^d. Dissipative entropy s^d is a thermal internal variable. The potential of using the ABAQUS coupled thermal-stress analysis procedure to solve problems having a Laplacian term in the damage evolution equation was briefly discussed.

ACKNOWLEDGEMENTS

Part of the theory within this report was prepared under the fellowship contract by the European Economic Community at the Joint Research Centre, Institute for Advanced Materials, Petten, the Netherlands. The theory was finalised in the REVISA (Reactor Vessel in Severe Accidents) project organised by the European Commission. The Commission of the European Communities is gratefully acknowledged.

REFERENCES

[1] Bažant, Z. P. 1994. Nonlocal damage theory based on micromechanics of crack interactions. Journal of Engineering Mechanics, vol. 120, no. 3, pp. 593 - 617.

[2] Ramaswamy, N. & Aravas, N. 1998. Finite element implementation of gradient plasticity models; Part I: Gradient-dependent yield functions. Computer methods in applied mechanics and engineering, no. 163, pp. 11 - 32.

[3] Santaoja, K. 2000. Thermomechanics of a gradient theory. Espoo, Finland: Helsinki University of Technology. Helsinki University of Technology Laboratory for Mechanics of Materials. Research Reports - TKK-LO-28. 57 p. + app. 9 p.

[4] Aifantis, E. C. 1992. On the role of gradients in the localization of deformation and fracture. International Journal of Engineering Sciences, vol. 30, no. 10, pp. 1279-1299.

[5] Kukudžanov, V. N., Bourago, N. G., Kovshov, A. N., Ivanov, V. L. & Schneiderman, D. N. 1995. On the problem of damage and localization of strains. Göteborg, Sweden: Chalmers University of Technology, Department of Structural Mechanics. 35 p. (Publication 95:11)

[6] Ziegler, H. 1963. Some extremum principles in irreversible thermodynamics with application to continuum mechanics. In: Sneddon, I. W. & Hill, R. (F.R.S.) Eds. Progress in Solid Mechanics. Vol. IV. Amsterdam, The Netherlands: North-Holland Publishing Company. 198 p.

[7] Maugin, G. A. 1990. Internal variables and dissipative structure. Journal of Non-Equilibrium Thermodynamics, vol. 15, pp. 173-192.

[8] Le Gac, H. & Duval, P. 1980. August. Constitutive relations for the non elastic deformation of polycrystalline ice. In: Tryde, P. Ed. Physics and Mechanics of ice. Book of papers presented at the Symposium of the International Union of Theoretical and Applied Mechanics. Berlin, Federal Republic of Germany: Springer-Verlag. Pp. 51-59.

[9] ABAQUS Theory Manual 1998 (Version 5.8). Rhode Island, USA: Hibbit, Karlsson &Sorensen, Inc.

[10] ABAQUS User's Manual 1998 (Version 5.8). Rhode Island, USA: Hibbit, Karlsson & Sorensen, Inc.

HITSAUKSEN VAIKUTUS MUOKKAUSLUJITTUNEESTA RUOSTUMATTOMASTA TERÄKSESTÄ VALMISTETUN TASALUJAN LIITOKSEN LUJUUTEEN

L. HUHTALA
Lappeenrannan teknillinen korkeakoulu, konetekniikan osasto, PL 20, 53851 LPR
A. TALJA
VTT Rakennustekniikka, PL 18071, 02044 VTT
P. YRJÖLÄ
Stala Oy, 15170 Lahti

TIIVISTELMÄ

Tässä esityksessä on tarkasteltu hitsauksen vaikutusta valmistusprosessissa muokkauslujittuneiden ruostumattomien putkipalkkien tasalujien liitosten lujuuteen. Tutkittavana kohteena oli austeniittisten neliöputkiprofiilien piena- ja päittäisliitoksen kestävyystarkastelu staattisella aksiaalisella vetokuormituksella. Liitosten kestävyyden laskennassa käytettiin kahta laajalti käytössä olevaa standardia, SFS-ENV 1993-1-4:1996 (Eurocode 3, osa 1.4) ja ANSI/ASCE-8-90. Käytettyjen standardien mitoitusperusteet ovat erilaiset. Eurocode 3-ohjeen ja ANSI/ASCE-8-90-ohjeen liitoskestävyyden laskentayhtälöitä vertailtiin keskenään. Testitulosten vertailussa kestävyysyhtälöiden välillä sovellettiin lujuutena materiaalitodistusten määrittämiä lujuusarvoja, putkiprofiilin mitattuja lujuusarvoja ja hitsiaineen tyypillisiä lujuusarvoja. Tarkoituksena oli selvittää, onko mahdollista hyödyntää valmistusprosessissa tapahtuvaa putkipalkkien lujittumista hitsausliitosten mitoituksessa, ja millä mitoitusperusteilla päästään luotettavimpaan lopputulokseen.

1. JOHDANTO

Ruostumattomien terästen eräs edullinen ominaisuus on sen kyky muokkauslujittua kylmämuovausprosessissa. Lujittuminen havaitaan materiaalin kohonneina lujuusarvoina. Kuitenkin käytössä olevien ohjeistojen mitoitusyhtälöissä lujuusarvot perustuvat nauhasta tai levystä mitattuihin arvoihin, jolloin valmistusprosessissa kylmämuovauksen seurauksena saavutettuja putken lujuusarvoja ei mitoituksessa hyödynnetä. Sekä Eurocode 3:n että amerikkalaisen ANSI/ASCE-8-90:n mukaan kylmämuokatulle materiaalille voidaan käyttää korotettuja myötörajan ja vetomurtolujuuden arvoja sillä edellytyksellä, ettei niitä hitsata tai lämpökäsitellä kylmämuokkauksen jälkeen ellei kokeellisesti voida osoittaa, että hitsaus tai lämpökäsittely eivät pienennä lujuusarvoja sovellettavia arvoja pienemmiksi. ANSI/ASCE:ssa ohjetta tarkennetaan vielä, että saavutetun lujuuden on lisättävä rakenteen kapasiteettia suunniteltua kuormitustapaa ja -suuntaa vastaan. Lisäksi lujuusominaisuudet täytyy varmistaa täysmittakaavakokein. [1]

2. NYKYISET SUUNNITTELUOHJEET

Tässä yhteydessä on tarkasteltu kahta laajalti käytössä olevaa ruostumattomien terästen hitsausliitosten mitoitusohjetta, jotka ovat eurooppalainen esistandardi SFS-ENV 1993-1-4:1996 (Eurocode 3, osa 1.4): Ruostumattomia teräksiä koskevat lisäsäännöt [2] sekä amerikkalainen kylmämuovattuja ruostumattomia teräksiä koskeva standardi ANSI/ASCE-8-90: Specification for the Design of Cold-Formed Stainless Steel Structural Members [3].

Eurocode 3:n mukaan ruostumattomien terästen hitsausliitosten mitoituksessa sovelletaan yleisille rakenneteräksille tarkoitettua ohjetta [4] ellei osassa 1.4 muuta ilmoiteta. Vetorasitetun läpihitsatun päittäisliitoksen kestävyys mitoitetaan heikoimman liitettävän osan lähtönauhan myötölujuuden perusteella. Hitsiltä edellytetään perusainetta suurempaa myötö- ja murtolujuutta. Myötöön perustuvana osavarmuuskertoimen arvona käytetään γ_{M0} = 1,1. ANSI/ASCE-standardissa päittäisliitos mitoitetaan murtumiselle sekä perusmateriaalista että hitsausliitoksesta. Perusmateriaalin kestävyys perustuu hehkutetun nauhan murtolujuuteen ja hitsin kestävyys hitsiaineen murtolujuuteen. Lujuuskertoimen arvona käytetään Φ = 0,60 [5]. Vastaavaa murtolujuusperusteista lähestymistapaa on Eurocode 3:ssa käytetty ruuviliitosten nettopoikkileikkauskestävyyden laskennassa.

Pienaliitoksen mitoitus perustuu sekä Eurocode- että ANSI/ASCE-ohjeissa perusaineen murtolujuuteen. Eurocode-ohjeissa (liitteen M mukainen mitoitusyhtälö) liitos mitoitetaan vain hitsin murtumiselle. Hitsisaineen perusainetta korkeampi lujuus otetaan mitoituksessa huomioon β_w-kertoimella. Lujuuskertoimen β_w arvoksi valitaan 1,0 kaikille ruostumattomille teräksille ellei pienempää arvoa osoiteta perustelluksi kokeellisesti. Murtoon perustuva osavarmuuskertoimen arvo on γ_{M2} = 1,25. Amerikkalaisessa ANSI/ASCE-ohjeistossa pienaliitos mitoitetaan sekä hitsiaineen vetomurtolujuuden että perusmateriaalin lujuuden mukaan. Hitsin mitoituksessa käytetty t_w on hitsin poikkipinnan tehollinen paksuus, joka vastaa Eurocode 3:ssa hitsin a-mittaa. Perusmateriaalin mukaan mitoitus tapahtuu päittäisliitosmitoituksena. Taulukossa 1 on esitetty Eurocode 3- ja ANSI/ASCE-ohjeistoissa olevat päittäis- ja pienaliitoksen mitoitusyhtälöt (1)-(6).

Suunnitteluohjeiden vertailua

Päittäisliitosmitoituksen suoraa vertailua eri laskentaohjeiden välillä ei voida tehdä, koska ohjeistojen käyttämät mitoitusperusteet eivät ole samat. Aksiaalisesti vetokuormitetun ehjän sauvarakenteen kestävyys, $\Phi_t T_n$, määritellään ANSI/ASCE-8-90 mukaan yhtälöllä (7)

$$\begin{aligned} \Phi_t &= 0{,}85 \\ T_n &= A_n F_y , \end{aligned} \qquad (7)$$

missä Φ on ehjän sauvaosan lujuuskerroin ja F_y nauhamateriaalin myötölujuus.

Taulukko 1. Päittäis- ja pienaliitoksen mitoitusyhtälöt standardien SFS-ENV 1993-1-1:1992 ja ANSI/ASCE-8-90 mukaan.

LIITOSTYYPPI	SUUNNITTELUOHJE	
	EUROCODE 3	ANSI / ASCE-8-90
Päittäisliitos	$N_{pl.Rd} = \frac{Af_y}{\gamma_{M0}}$, $\gamma_{M0} = 1{,}1$ (1)	Kestävyys ΦP_n
Ruuviliitosten nettopoikkileikkauskestävyys	$N_{u.Rd} = \frac{0.9 A_{net} f_u}{\gamma_{M2}}$, $\gamma_{M2} = 1{,}25$ (2)	$P_n = LtF_{ua}$, (3) $\Phi = 0{,}60$
Pienaliitos	$F_{w.Rd} = \frac{f_u}{\beta_w \gamma_{Mw} \sqrt{2}} aL$ (4) $\gamma_{Mw} = 1{,}25$, $\beta_w = 1{,}0$	Kestävyys ΦP_n, poikittainen kuormitus $P_n = tLF_{ua}$, $\Phi = 0{,}55$ (5) $P_n = t_w LF_{xx}$, $\Phi = 0{,}65$ (6)

Yhtälö (7) vastaa Eurocode 3:n päittäisliitoksen mitoitusyhtälöä (1). Tällöin Eurocode 3:n määrittämä kestävyys on 1,07 kertaa ANSI/ASCE:n määrittämä kestävyys, kun $A_n = A$. Vertaamalla Eurocode 3:ssa olevaa vetomurtolujuuteen perustuvaa ruuviliitosten nettopoikkileikkauskestävyyden yhtälöä (2) ANSI/ASCE-standardin päittäisliitoksen mitoitusyhtälöä (3), on Eurocode 3:n määrittämä kestävyys 1,2-kertainen, kun käytetään $A_{net} = A_{br}$.

Pienahitsien kestävyyttä vertailtaessa on liitosten suunnitellut murtumistavat asetettava vertailukelpoisiksi keskenään. Eurocode-ohjeen pienaliitoksen kestävyysyhtälöä (4) vastaa ANSI/ASCE-standardin yhtälö (6), jolloin molempien yhtälöiden käytöllä murtumisen oletetaan tapahtuvan hitsistä. Olettamalla, että hitsin lujuus on liitoksella määräävä ($F_{xx} = F_{ua}$) ja käyttämällä $\Phi = 0{,}65$, saadaan

$$\Phi P_n = 0{,}65 L t_w F_{ua}. \tag{8}$$

Eurocode 3:n mukainen yhtälö (4) saa muodon

$$F_{w.Rd} = 0{,}566 f_u t_w L, \tag{9}$$

kun käytetään $a = t_w$ ja $\beta_w = 1{,}0$. Yhtälöiden (8) ja (9) perusteella havaitaan, että ANSI/ASCE-standardin määrittämä kestävyyden arvo on n. 1,15 kertaa Eurocode-ohjeen antama kestävyys, kun vaurio on ohjattu hitsiin. Yhtälöä (5) käyttämällä murtumisen oletetaan tapahtuvan perusaineesta ja sen mukaista kestävyystarkastelua ei Eurocode 3:n pienaliitosmitoitukseen sisälly. Poikkileikkauksen kestävyys tulee tarkastettua koko rakenteen kestävyystarkastelussa.

3. KOEOHJELMA

Tutkittavina materiaaleina olivat austeniittiset ruostumattomat teräslaadut EN 1.4301 (AISI 304) ja EN 1.4404 (AISI 316L) ja niistä kylmämuovaamalla valmistettujen neliöputkiprofiilien hitsausliitokset. Testiohjelma sisälsi kokeita profiilin lähtönauhalle ja putkesta otetuille näytteille sekä täysmittakaavakokeita päittäis- ja pienaliitoksille. Testattavien kappaleiden rasitusmuotona tutkittiin staattista aksiaalista vetokuormitusta. Kokeiden tarkoituksena oli selvittää hitsatun liitoksen käyttäytymistä staattisella kuormituksella, ja tapahtuuko hitsauksen johdosta (lämmöntuonti) liitoksen haitallista pehmenemistä. Koetuloksista tutkittiin ensisijaisesti liitoksen kestävyysarvot ja vauriokohta.

Standardivetokokeet

Liitoskokeissa testattujen profiilien materiaaliominaisuudet määritettiin standardivetokokeilla, jotta todettaisiin kylmämuovauksen aikaansaama lujuuden kasvu ja sen suuruus. Profiilien nurkka-alueita ei testattu. Nauhamateriaalien mekaaniset ominaisuudet määriteltiin materiaalitodistuksen tueksi vain profiileilla RHS-1 ja RHS-2. Testaus tehtiin pituussuunnassa toisin kuin materiaalitodistuksen arvot on saatu. Myös putkesta leikatut standardivetokoekappaleet testattiin valssaussuunnassa eli valmiin putkipalkkirakenteen pituussuunnassa. Koesauvojen testaus suoritettiin standardin SFS-EN 10002-1: Metallien vetokoe -mukaan. Kokeissa mitattiin koekappaleen voima-siirtymäkäyrä ja murtovenymän arvo.

Liitoskokeet

Muokkauslujittumisen vaikutusta hitsausliitoksen kestävyyteen tutkittiin sekä päittäisliitoskokeilla että pienaliitoskokeilla taulukon 2 mukaisilla profiileilla. Pienaliitoksen a-mitan lähtöarvoksi valittiin $1.3t$, jossa t oli liitettävien osien seinämäpaksuus, ja jolla murtuminen siirrettiin pois hitsistä. Hitsausprosesseina käytettiin MIG-hitsausta ja puikkohitsausta sekä hitsiaineena austeniittista ja duplex-lisäainetta.

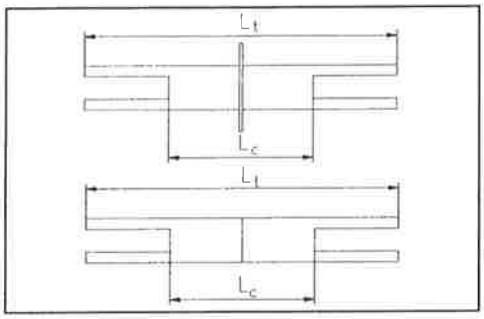
Taulukko 2. Täysmittakaavakokeissa testatut putkiprofiilit (lisäksi kaksi pienaliitosesitestausta).

Merkintä	Materiaali	Poikkipinta	Koepituus L_c [mm]	Kokonais-pituus L_t [mm]	Liitos-tyyppi	Määrä
RHS-1	EN 1.4301	100x100x3	368	788	pienaliitos	3
			360	780	päittäisliitos	4
RHS-2	EN 1.4301	100x100x6	368	788	pienaliitos	3
			360	780	päittäisliitos	4
RHS-3	EN 1.4404	50x50x1,5	744	974	pienaliitos	2
			740	970	päittäisliitos	3
RHS-4	EN 1.4301	50x50x2	744	974	pienaliitos	2
			740	970	päittäisliitos	2
RHS-5	EN 1.4301	30x30x3	744	974	pienaliitos	2
			740	970	päittäisliitos	2

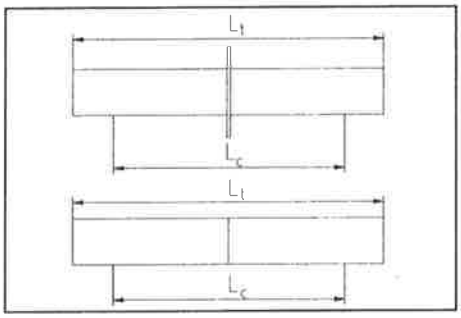
Testit suoritettiin kuvien 1-2 mukaisilla koekappaleilla. Liitoskokeiden koejärjestelyä on esitetty kuvassa 3. Kuormitusnopeus laskettiin siten, että kuormitukseen 75 %:een asti

arvioidusta murtokuormasta (F_{mit}), kuormitusnopeus oli $0{,}05F_{mit}$/min. Tämän jälkeen nopeus oli puolet tästä murtoon asti.

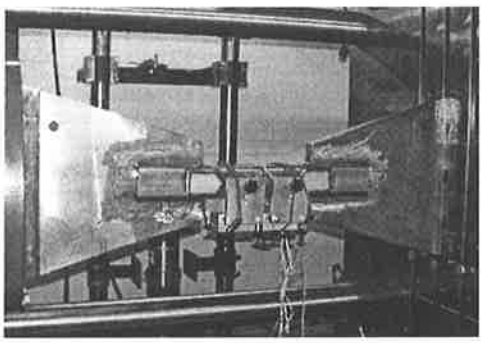
Koesauvaan liitoksen alueelle kiinnitettiin mittalaite, joka mittasi siirtymiä alkumittapituuden matkalta. Kokonaissiirtymää mitattiin kuormituskehän siirtymäanturilla. Mitatuista voimista ja alkumittapituuden siirtymistä saatiin liitoksen alueen jännitys-venymäkäyrät. Profiileihin RHS-1 ja RHS-2 liimattiin venymäliuskoja profiilin ylä- ja alapinnalle, kuva 4, liitoksen lokaalisen käyttäytymisen selvittämiseksi.



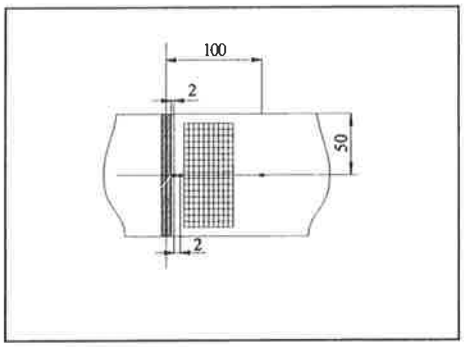
Kuva 1. Piena- ja päittäisliitoskoekappaleet RHS-1 ja RHS-2.



Kuva 2. Piena ja päittäisliitoskoekappaleet RHS-3, RHS-4 ja RHS-5



Kuva 3. Päittäisliitoksen testaus.



Kuva 4. Venymäliuskat liitosalueella.

Kovuusmittaukset

Kovuusmittauksen tarkoituksena oli pehmenemis-/lämpövyöhykkeen leveyden sekä perus- ja lisäaineen kovuuden selvittäminen. Kovuusmittaukset tehtiin hitsin poikkileikkauksesta standardien SFS 1043-1 ja SFS 288-3 mukaisesti ja kovuudet mitattiin Vickers-HV5 -kovuuksina.

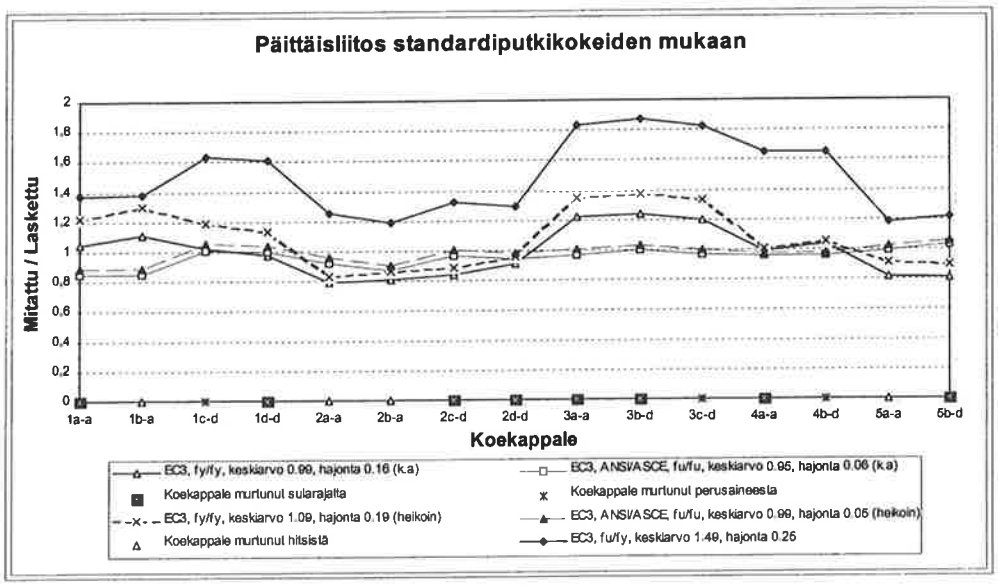
4. TULOKSET

Taulukossa 3 on vertailtu profiileista RHS-1 ja RHS-2 leikattujen koekappaleiden materiaaliparametrien arvoja lähtönauhasta tehtyjen kokeiden ja materiaalitodistuksen arvoihin. Tarkastelun kohteena on myötöraja ($R_{p0.2}$-raja), murtolujuus ja murtovenymä. Lujuusarvot on määritetty jännitys-venymäkäyrien avulla.

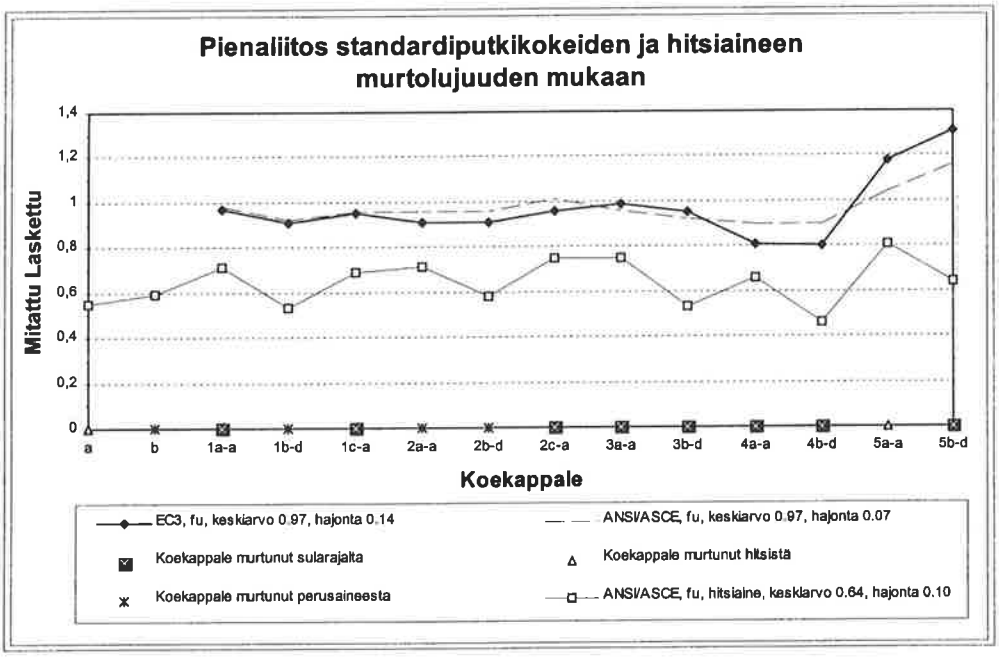
Taulukko 3. Profiilien RHS-1 ja RHS-2 putki- ja nauhamateriaalin lujuusvertailu.

Profiili	**RHS-1**			**RHS-2**		
Koekappaleen paikka	Uuma1	Laippa	Uuma2	Uuma1	Laippa	Uuma2
0.2-rajan kasvu: (b)/(a)	1,37 1,29	1,69 1,59	1,23 1,15	1,72 1,70	1,74 1,72	1,59 1,57
a) Nauhan $R_{p0.2}$ [N/mm²]	290,243,292			278,323,291		
	keskiarvo 275 mat.tod. 292			keskiarvo 297 mat.tod. 301		
b) RHS:n $R_{p0.2}$ [N/mm²]	376	465	337	512	518	473
	keskiarvo 393			keskiarvo 501		
Murtolujuuden kasvu: (b)/(a)	1,03 0,99	1,14 1,10	1,02 0,98	1,10 1,07	1,22 1,19	1,11 1,08
a) Nauhan R_m [N/mm²]	599,607,593			601,599,603		
	keskiarvo 600 mat.tod. 621			keskiarvo 601 mat.tod. 617		
b) RHS:n R_m [N/mm²]	615	683	608	659	732	664
	keskiarvo 635			keskiarvo 685		
Venymän muutos: (b)/(a)	0,96 0,86	0,97 0,87	0,96 0,86	0,80 0,85	0,65 0,69	0,82 0,88
a) Nauhan A_5 [%]	54,5 61			58,9 55		
b) RHS:n A_5 [%]	52,5	52,8	52,4	46,9	38,1	48,5
	keskiarvo 52,6			keskiarvo 44,5		
Murto- ja myötölujuuden suhteen muutos R_m / $R_{p0.2}$: (b)/(a)	0,75 0,77	0,67 0,69	0,83 0,85	0,63 0,63	0,69 0,69	0,69 0,68
a) Nauhan R_m / $R_{p0.2}$	2,18 mat.tod. 2,13			2,04 mat.tod. 2,05		
b) RHS:n R_m / $R_{p0.2}$	1,64	1,47	1,80	1,29	1,41	1,40

Päittäisliitoskokeissa mitattujen ja mitoitusyhtälöillä laskettujen kestävyyksien suhteet koekappaleittain on esitetty kuvassa 5. Mitoitusyhtälöissä on käytetty standardivetokokeissa saavutettuja putkesta leikattujen koekappaleiden lujuusarvoja (keskiarvo/heikoin). Osavarmuus- ja lujuuskertoimen arvo yhtälöissä on tällöin $\gamma_{M0} = 1,0$ ja $\Phi = 1,0$. Kuvassa on vertailtu kolmea eri suhdetta/mitoitustapaa. Kun päittäisliitosmitoitus perustuu murtolujuuteen saadaan sekä EC 3:lla että ANSI/ASCE:lla liitokselle sama kestävyys. Tällöin kokeessa mitattua murtokuormituksen arvoa on verrattu arvoon, jonka liitos kestää, kun mitoitusyhtälössä käytetään lujuutena standardivetokokeessa saavutettua putkesta leikatun koekappaleen murtolujuusarvoa. Vastaavanlainen vertaus on tehty mitatun murtokuorman ja putkimateriaalin laskettua myötörajaa vastaavan kestävyyden sekä mitatun myötäämisen johtavan kuorman ja lasketun myötörajaa vastaavan kestävyyden välillä. Koekappaleiden murtumiskohdat on jaoteltu murtuminen hitsistä, sularajalta ja perusaineesta. Kuvassa 6 on esitetty vastaavasti pienaliitosten mitattu/laskettu -suhteen arvot EC 3:n ja ANSI/ASCE:n vertailukelpoisten mitoitusyhtälöiden välillä.



Kuva 5. Päittäisliitoksen kestävyysvertailu eri mitoitusohjeistojen välillä, kun mitoituslähtökohtana on standardiputkikokeiden määrittämät lujuusarvot. Käytetyt hitsausaineet: a = austeniittinen lisäaine, d = duplex-lisäaine.



Kuva 6. Pienaliitoksen vertailu eri mitoitusohjeistojen välillä, kun murtumisen oletetaan tapahtuvan hitsistä.

5. TULOSTEN TARKASTELU

Materiaalin murtuminen on eksakti tapahtuma ja murtumisen arvo on selvästi määritettävissä. Murtolujuuteen perustuvan laskennan hajonnat olivat pienempiä kuin myötölujuuteen pohjautuvan. Näin ollen tulosten tarkastelussa keskitytään murtumistarkasteluun, etenkin päittäisliitoksen murtumiseen. Tehtyjen testien perusteella pyrittiin löytämään päittäisliitosmitoitukselle käytettäviä osavarmuuskertoimia, jotka perustuvat materiaalitodistuksen määrittelemään murtolujuuteen ja muokkauslujittuneen putkipalkin murtolujuuteen.

Standardivetokokeet

Kylmämuovausprosessin aikaansaama materiaalin lujittuminen havaittiin selvästi erityisesti myötölujuudessa paitsi hoikkaseinämäisellä haponkestävällä profiililla RHS-3. Murtolujuuden kasvu oli pientä tai sitä ei tapahtunut lainkaan. Putken lujuuden suurimmat arvot saavutettiin hitsiin nähden vastakkaisella sivulla. Standardin vaatimiin minimilujuuksiin verrattuna murtolujuus oli 16-18 % suurempi. Muokkauslujittumisesta on seurannut murtovenymän pienentymistä. Mitatut murtovenymien arvot olivat 59-97 % nauhamateriaalien arvoista.

Kimmokertoimien arvojen tarkka vertailu on vaikeaa. Putkista otettujen näytteiden arvojen hajonta oli suuri. Keskimääräinen saavutettu kimmokerroin E = 186 GPa on ohjeistojen kimmokertoimien arvoja (Eurocode 3: E = 200 GPa, ANSI/ASCE: E = 193 GPa) pienempi. Amerikkalaisen ANSI/ASCE:n arvot vastaavat lähinnä tämän työn tuloksia.

Osavarmuuskertoimet

Laskettu redusoitu osavarmuuskerroin $\gamma_{M1,red}$ määrittää varmuuden, jonka murtoon perustuva mitoitus vaatii, jotta karakteristinen kestävyys saavutetaan. Karakteristista liitoksen kestävyyttä haettaessa käytetään hyväksi Eurocode 3:ssa olevaa murtoon perustuvaa osavarmuuskerrointa γ_M = 1,25, joka sisältää materiaalista, suunnittelusta ja valmistuksesta johtuvat epävarmuustekijät. Varmuuskertoimen laskentayhtälö on muotoa

$$\gamma_{M1,red} = \frac{1}{\bar{x} - 2s} \cdot \gamma_M , \qquad (10)$$

missä $\bar{x}$ on koetulosten mitattu/laskettu -suhteen aritmeettinen keskiarvo ja s on keskihajonta.

Päittäisliitokset

Taulukossa 4 on esitetty murtolujuuteen perustuen tutkittujen päittäisliitosten vaatimia redusoitujen kertoimen arvoja karakteristisen kestävyyden takaamiseksi materiaalitodistuksen ja standardiputkikokeiden mukaan. Kun mitoitus tehdään EC3:n ja ANSI/ASCE:n mukaisesti murtolujuuden perusteella materiaalitodistuksen arvoja käyttäen, saadaan redusoidun osavarmuuskertoimen arvoksi $\gamma_{M1,red}$ = 1,49. Tarkasteltaessa liitoksen

alueelta murtuneita koekappaleita (vaurioituminen on tapahtunut suunnitellusta kohdasta eikä koejärjestelyn "puutteesta" johtuen perusmateriaalista), on osavarmuuskertoimen arvo $\gamma_{M1.red}$ = 1,47. Lähes vastaavat arvot saadaan, kun lähtökohdaksi valitaan putken keskimääräinen murtolujuus ($\gamma_{M1.red}$ = 1,49, $\gamma_{M1.red}$ = 1,46). Standardivetokokeiden mukaista minimilujuutta käyttäen on osavarmuuskerroin $\gamma_{M1.red}$ = 1,40, ja liitoksen alueelta murtuneilla koekappaleilla $\gamma_{M1.red}$ = 1,39.

Taulukko 4. Kertoimen $\gamma_{M1.red}$ arvoja eri vertailukriteerein mitattu/laskettu -suhteen mukaan.

Vertailukriteeri	Keskiarvo $\bar{x}$		Hajonta s		Kerroin $\gamma_{M1.red}$	
	Kaikki	Liitosalue	Kaikki	Liitosalue	Kaikki	Liitosalue
Materiaalitodistus						
EC3, ANSI/ASCE-8-90, f_u/f_u	0,98	0,99	0,07	0,07	1,49	1,47
Standardiputkikoe						
EC3, ANSI/ASCE-8-90, f_u/f_u (k.a)	0,95	0,97	0,06	0,05	1,49	1,46
EC3, ANSI/ASCE-8-90, f_u/f_u (heikoin)	0,99	1,02	0,05	0,06	1,40	1,39

Taulukkoon 5 on koottu lasketut ja nykyisin käytössä (ANSI/ASCE) olevat murtolujuuteen perustuvat osavarmuuskertoimet päittäisliitokselle. EC3:n mukainen osavarmuuskerroin on käänteislukuna, jotta se olisi vertailukelpoinen ANSI/ASCE:n lujuuskertoimiin nähden. EC3:n mukaisesti lasketut varmuuskertoimet antavat liitoksien kestävyydelle suurimman laskennallisen arvon. Ero kokeellisen mitoituksen avulla laskettuun lujuuskertoimeen on 5 % ja nykyisin käytössä olevaan ANSI/ASCE-standardin mukaiseen varmuuteen 12 % tai 18 % vertailusta riippuen.

Taulukko 5. Murtolujuusmitoitukseen perustuvien päittäisliitosmitoituksen osavarmuus- ja lujuuskertoimien vertailu.

Vertailukriteeri	Materiaalitodistus	Standardiputkikoe (keskiarvo)	Heikoin standardiputkikoe
Ohjeisto			
EC 3 ($1/\gamma_{M1.red}$)	0,67	0,67	0,71
Laskettu ANSI/ASCE	0,64	0,64	0,68
Nykyinen ANSI/ASCE	0,60	-	-
Suhde			
EC 3/Laskettu ANSI/ASCE	1,05	1,05	1,05
EC3/Nykyinen ANSI/ASCE	1,12	-	-
Laskettu/Nykyinen	1,07	-	-

Heikoimman standardiputkikokeen mukaan laskettu osavarmuuskerroin ($1/\gamma_{M1.red}$ = 0,71) on lähes sama kuin ruuviliitosten mitoituksessa käytettävän laskentayhtälön (2) osavarmuuskerroin 0,72, joka myös perustuu murtolujuuteen.

Pienaliitokset

Pienaliitoskokeiden tarkoituksena ei ollut vaikuttaa EC3:n ja ANSI/ASCE:n laskentayhtälöiden osavarmuus- ja lujuuskertoimiin vaan pyrittiin selvittämään putkesta mitattujen kylmämuokkauksessa saatavien lujuusarvojen säilymistä myös pienaliitoksella. Pienaliitoskokeilla myös tuettiin päittäisliitoksen redusoidun osavarmuuskertoimen muodostamista. Murtolujuuteen perustuva osavarmuuskerroin γ_M = 1,25 lienee sopiva käytettäväksi myös ruostumattomien terästen pienaliitoksen laskennassa. Kuitenkin EC3:n antama n. 6 % pienempi kestävyys ANSI/ASCE:een verrattuna, kun a-mittana käytetään 1.3 kertaa seinämän paksuutta, voisi mahdollistaa lujuuskertoimen β pienentämisen. Tämä kuitenkin vaatisi jatkotutkimusta.

Tässä työssä on varmuuskertoimen osalta tyydytty vertailemaan keskenään ANSI/ASCE:n kokeellisella mitoituksella [2] saatavaa lujuuskerrointa ja nykyisin käytössä olevaa lujuuskerrointa, jotka on esitetty taulukossa 6. Lasketut lujuuskertoimet ovat suurempia kuin käytössä olevat, mikä johtunee profiilin kohonneista lujuusarvoista. Hitsiaineen lujuuskertoimien välinen suora vertailu ei liene perusteltua, koska sen todellista lujuutta ei tiedetty. Lisäksi vertailuun taustaksi olisi vaadittu erisuuruisia *a*-mittoja.

Taulukko 6. Pienaliitoksen perusmateriaalin mukaiset lujuuskertoimet.

Vertailukohde	**Laskettu lujuuskerroin Φ**	**Käytössä oleva lujuuskerroin Φ**
Materiaalitodistus	0,66	0,55
Standardiputkikoe	0,63	-

6. JOHTOPÄÄTÖKSET

Tehtyjen testien perusteella voidaan esittää, että nykyisin käytössä oleva hehkutetun materiaalin myötölujuuteen perustuva eurooppalainen ohjeisto päittäisliitoksen kestävyyden laskennassa on konservatiivinen. Kylmämuovatun putkiprofiilin liitoksen kestävyys ei näytä pienenevän lämmöntuonnin vaikutuksesta. Kun murtuminen ei tapahdu hitsistä, liitoksen kestävyyttä kuvaa parhaiten perusmateriaalin murtolujuus. Mitoituksen lähtökohdaksi valittiin putken minimimurtolujuus. Testien perusteella piena- ja päittäisliitosten kestävyyden laskenta voidaan tehdä putkesta mitattuun vetomurtolujuuteen perustuen. Tällöin mitoituksessa on osavarmuuskertoimena käytettävä vähintään arvoa γ_M = 1.40, jota voidaan käyttää, kun hitsiaineen murtolujuus on vähintään putken materiaalin murtolujuuden suuruinen. Tehtyjen testitulosten perusteella ehdotetaan, että päittäisliitoksen kestävyys, kun murtuminen tapahtuu liitoksen alueelta, lasketaan yhtälöstä

$$N_{pl.Rd} = \frac{Af_u}{\gamma_{M1.red}},$$

missä f_u = putken minimilujuus ja osavarmuuskerroin $\gamma_{M1.red}$ = 1,40.

Mikäli lähtökohdaksi otetaan ruuviliitosten nettopoikkileikkauksen mukainen kestävyyden mitoitusarvo, voidaan päittäisliitoksen kestävyystarkastelu suorittaa vaihtoehtoisesti yhtälön (2) mukaan. Hitsausaineen murtolujuuden on oltava vähintään perusaineen murtolujuuden suuruinen.

Kokeissa käytetyt pienaliitokset oli suunniteltu niin, että murtuminen tapahtui muualta kuin hitsistä. Koska tällöin lujuuskokeiden tulokset vastasivat hyvin päittäisliitoskokeiden kestävyyksiä, pienaliitos voidaan mitoittaa perusmateriaalin murtumisen suhteen kuten päittäisliitoskin. Lisäksi on tarkistettava pienahitsin koon riittävyys Eurocode 3:n mukaisesti, mutta hitsiaineen lujuuteen perustuen.

TIEDOKSI

Tämä esitelmä perustuu Lappeenrannan teknillisellä korkeakoululla tehtyyn diplomityöhön ja sen tuloksiin. Työ oli osa laajempaa TEKES-projektia: Ruostumattomista teräksistä valmistetut kantavat rakenteet. Projektissa olivat mukana VTT Rakennustekniikka, Outokumpu Oy, Stala Oy, Rannila Steel, JA-RO Oy ja A-Insinöörit Oy, joille työn tukemisesta ja ohjauksesta esitetään kiitokset.

LÄHTEET:

1. Huhtala, L., 1999. Hitsauksen vaikutus muokkauslujittuneesta ruostumattomasta teräksestä valmistetun tasalujan liitoksen lujuuteen. Lappeenrannan teknillinen korkeakoulu, diplomityö. 141 s.

2. SFS-ENV 1993-1-1:1992. Eurocode 3: Teräsrakenteiden suunnittelu, Osa 1.1: Yleiset säännöt ja rakennuksia koskevat säännöt. Bryssel: Eurooppalainen standardisoimisjärjestö (CEN). 344 s.

3. ANSI/ASCE-8-90.1991. Specification for the Design of Cold-Formed Stainless Steel Structural Members. New York: American Society of Civil Engineers. 114 s. ISBN 0-87262-794-2.

4. SFS-ENV 1993-1-4:1996. Eurocode 3: Teräsrakenteiden suunnittelu, Osa 1.4:Yleiset säännöt: Ruostumattomia teräksiä koskevat lisäsäännöt. Bryssel: Eurooppalainen standardisoimisjärjestö (CEN). 61 s.

5. Lin, S-H., 1989. Load and Resistance Factor Design of Cold-Formed Stainless Steel Structural Members. Rolla: University of Missouri-Rolla. (Ph.D.thesis). 244 p.

EXPERIMENTAL DETERMINATION OF FATIGUE STRENGTH OF DYNAMICALLY LOADED WELDED JOINTS

Veli-Matti Lihavainen
Laboratory of steel structures
Lappeenranta University of Technology
P.O. Box 20
53851 Lappeenranta, FINLAND
veli-matti.lihavainen@lut.fi

ABSTRACT

The capacities and strengths of materials have a long time been improved by the steel industry. The higher yield and tensile strength of materials give a possibility to use higher stress level for the same thickness of plate. The fatigue phenomena will sometimes be indicated the most important failure mode. The fatigue problem has to take account into designing, especially when there are welded joints in structure. The fatigue strength of welded joints is usually determined by a fatigue life-stress range curve, so-called S-N-curve. These curves are presented as a log-log graph. The term fatigue class shows the stress range in MPa that gives a fatigue life of two million cycles at the confidence level of 97%.

The fatigue class of welded details are determined by experimental test series. Because of statistical aspects of fatigue phenomena, several tests are required for defining the fatigue strength of welded details. Based on experimental tests, the calculation of fatigue strength of one welded detail is described in this report. The terms fatigue class, fatigue limit and the effect of variable amplitude loading are briefly presented in this paper.

1. INTRODUCTION

The experimental determination of fatigue strength is usually analysed using nominal or hot spot stress approach [1]. The nominal stress approach is widely used because most design rules for steel structures contain a procedure for fatigue analysis based on this approach. The nominal stress approach yields satisfactory results with minimum calculation effort if the following conditions are fulfilled:

- there is an exact defined nominal stress, not complicated by macrogeometric effects
- the structural discontinuity is comparable with one of the classified details included in the design guides
- the detail is free from significant imperfections

The hot spot stress approach has been included in the new European pre-standard SFS-ENV 1993-1-1 [2] and the new IIW-recommendation [3]. The hot spot stress approach is used mainly for joints in which the weld toe orientation is transverse to the fluctuating stress component, and the crack is assumed to grow from the weld toe. The approach is not suitable for joints in which the crack would grow embedded defects or from the root of a fillet weld. Compared with the nominal stress approach, this approach is more suitable for use in following cases:

- there is no clearly defined nominal stress due to complicated geometric effects
- the structural discontinuity is not comparable with any classified details included in nominal stress design guides
- the offset or angular misalignments exceed any fabrication tolerance specified as consistent with the design S-N curves used in the nominal stress approach

The constant amplitude fatigue class and fatigue limit of welded details are determined by experimental test series in this research. The hot spot stress approach is used. Also the effect of variable amplitude loading is briefly studied. The results of test series are compared to hot spot S-N-curve.

2. THE S-N-CURVES OF FATIGUE STRENGTH

The definitions of basic terms are according to standard proposal [3]:

- constant amplitude loading - A type of loading causing a regular stress fluctuation with constant magnitudes of stress maxima and minima.
- variable amplitude loading - A type of loading causing irregular stress fluctuation with stress range of variable magnitude
- characteristic value - loads, forces or stresses, which vary statistically, at specified fractile
- fatigue strength - magnitude of stress range leading to a particular fatigue life
- fatigue class - stress range at the point of two million cycles at a confidence level of 97%
- fatigue limit - fatigue strength under constant amplitude loading corresponding to infinite fatigue life or a number of cycles large enough to be considered infinite by design code.
- stress range - the difference between stress maximum and stress minimum in a stress cycle
- knee point - the number of cycles which the slope of S-N-curve declines from m=3 to m=5

There are some differences between traditional nominal stress S-N-curve and hot spot stress S-N-curve. The constant amplitude fatigue limit initiates at the point of five million cycles in both cases. The knee point of nominal stress approach (denote m) is five million cycles and knee point of hot spot stress approach is twenty million cycles. For example, the S-N-curves of transverse fillet weld detail are presented in figure 1.

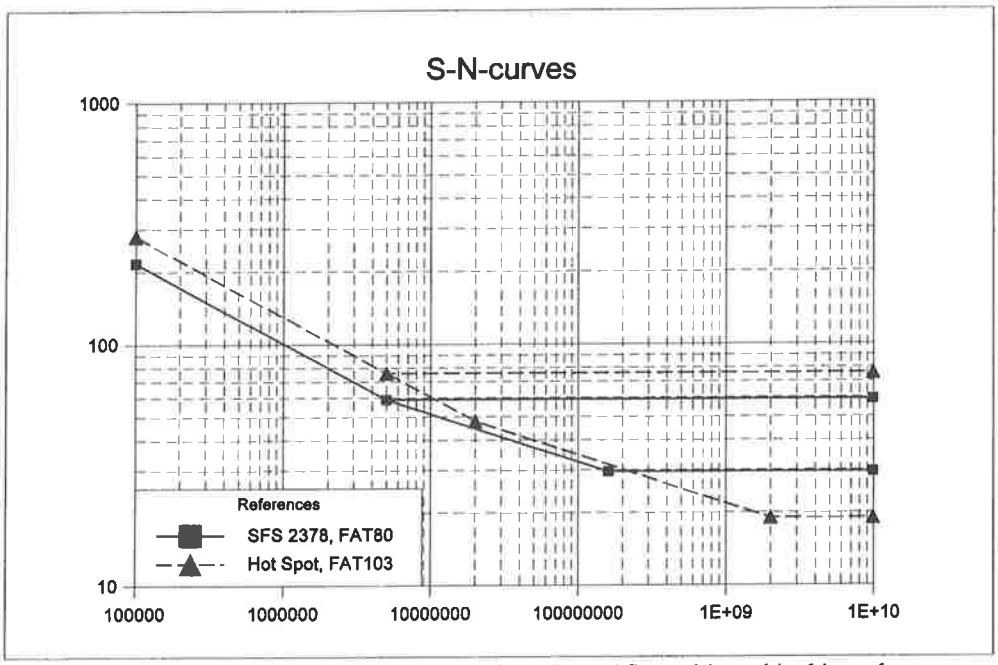


Figure 1. The S-N-curves of transverse fillet weld specimen (figure 2) used in this study.

3. TEST SPECIMEN

The tension fatigue tests were done using transverse fillet weld specimens. The material of specimens were high strength steel QSTE 460 TM made by Rautaruukki Oy.

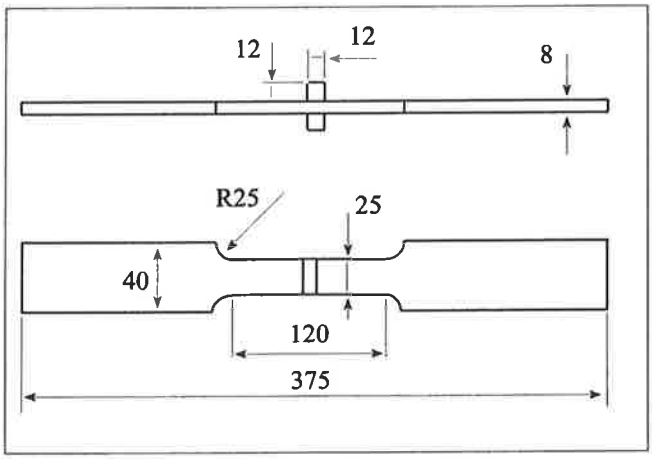


Figure 2. The transverse fillet weld test specimen.

4. DEFINITION OF S-N-CURVE OF TEST SERIES

The results based on hot spot stress range are presented in figure 3. Results are compared to hot spot S-N-curve FAT 103.

CA	constant amplitude fatigue test
VA	variable amplitude fatigue test
Broken	fatigue crack has been occurred, failure
Run-Out	fatigue crack has not been occurred, no failure

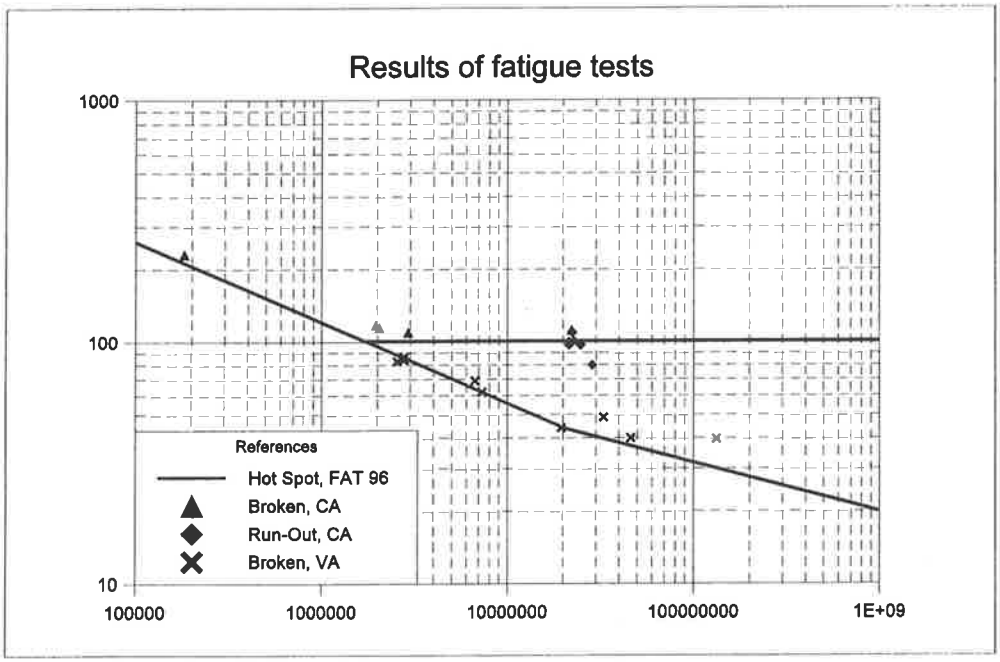


Figure 3. The results of fatigue tests.

The fatigue strength of test specimen is calculated using so-called C-value according to equation (1).

$\Delta\sigma$	nominal stress range
N	number of cycles

$$\Delta\sigma^3 \cdot N = C = \mathrm{FAT}^3 \cdot 2 \cdot 10^6 \qquad (1)$$

n	number of test pieces
s	standard deviation
FAT	fatigue strength of test specimen

The mean C-value and standard deviation of test series.

$$logC_{50\%} = \frac{\sum logC_i}{n} \qquad s = \sqrt{\frac{\sum (logC_i - logC_{50\%})^2}{n-1}} \tag{2}$$

The characteristic C-value.

$$logC_{95\%} = logC_{50\%} - s\left(1{,}64 + \frac{1{,}15}{\sqrt{n}}\right) \tag{3}$$

The characteristic fatigue class $FAT_{95\%}$ Stress range at the point of two million cycles at a confidence level of 95%.

$$FAT_{95\%} = \left(\frac{C_{95\%}}{2000000}\right)^{0{,}333} \tag{4}$$

The constant amplitude fatigue limit is calculated according to table 1. The fatigue tests have been discontinued (so-called Run-Out test), if the fatigue crack has not occurred before twenty million cycles.

Table 1. Calculation of constant amplitude fatigue limit.

Stress level i	Stress range [MPa]	Broken	Run-Out	Number of Run-Out f_i	if_i	i^2f_i
2	120	2	0	0	0	0
1	110	2	1	1	1	1
0	100	0	4	4	0	0
				$\Sigma 5 = F$	$\Sigma 1 = A$	$\Sigma 1 = B$

The average hot spot stress fatigue limit [5].

$$\Delta\sigma_{D50\%} = y_0 + d\left(\frac{A}{F} + 0{,}5\right) \tag{5}$$

y0 the lowest stress level (100 MPa)
d step between stress levels (10 MPa)

Standard deviation is given approximately by.

$$s = 1{,}62\ d\left(\frac{FB\text{-}A^2}{F^2} + 0{,}029\right) \tag{6}$$

The characteristic fatigue limit $\Delta\sigma_D$ at the point of twenty million cycles.

$$\Delta\sigma_D = \Delta\sigma_{D50\%} - 2 \cdot s \tag{7}$$

The equivalent stress range is calculated using equation (8) in variable amplitude loading.

$$\Delta\sigma_{eq} = \left(\frac{\sum_{i=j}^{k} (n_i\ \Delta\sigma_i^{m_1}) + \Delta\sigma_{th}^{(m_1-m_2)} \sum_{i=i_0}^{j-1} (n_i\ \Delta\sigma_i^{m_2})}{\sum n_i} \right)^{\frac{1}{m_1}} \tag{8}$$

$\Delta\sigma_i$	stress range at the level i
n_i	number of cycles at the level i
m_2	slope of low-gradient part of S-N-curve (= 5)
m_1	slope of high-gradient part of S-N-curve (= 3)
i_0	the lowest stress level above variable amplitude fatigue limit
j	the lowest stress level above the knee point of S-N-curve
k	the upper stress level

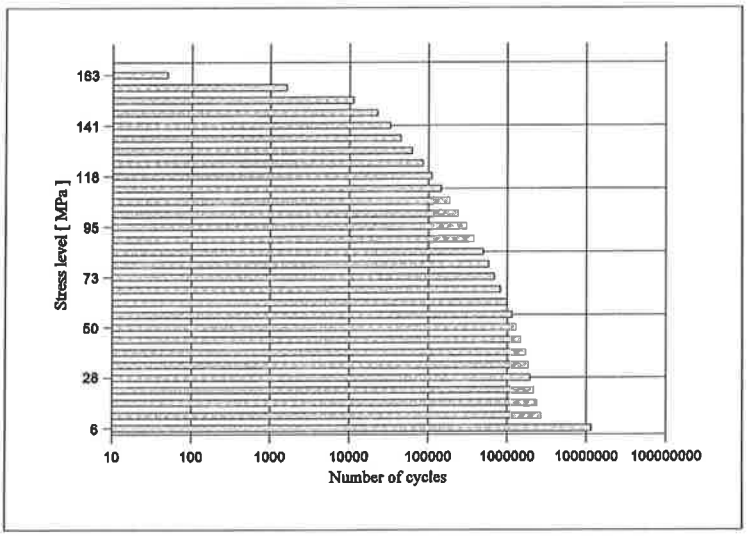


Figure 4. The stress level distribution of variable amplitude test.

5. DISCUSSION

Table 2. Comparison between hot spot -test results and hot spot S-N-curve.

	constant amplitude fatigue class $FAT_{95\%}$ $N=2\cdot10^6$	constant amplitude fatigue limit $\Delta\sigma_D$ $N=2\cdot10^7$	knee point between high- and low-gradient part of S-N-curve	variable amplitude fatigue limit $\Delta\sigma_L$
Results	96	101	44	-
S-N-curve	103	76	48	19

The constant amplitude fatigue class, fatigue limit and knee point have been determined using experimental fatigue tests and equations (1) - (8). The fatigue class and knee point of S-N-curve seems to be quite a similar than hot spot S-N-curve recommended in references.

The fatigue strength of some variable amplitude loaded, test specimen is lower than fatigue strength of constant amplitude result according to figure 3. The difference is not however substantial.

REFERENCES

[1] Erkki Niemi. Stress Determination for Fatigue Analysis of Welded Components. International Institute of welding. Document XIII-1458-92 / XV-797-92. Abington publishing. 69 p.

[2] Suomen standardoimisliitto. SFS-ENV-1993-1-1. Eurocode 3. Teräsrakenteiden suunnittelu. Osa 1-1. Yleiset säännöt ja rakennuksia koskevat säännöt.

[3] A. Hobbacher. Recommendations for fatigue design of welded joints. ISO standard proposal. International Institute of Welding. Document XIII-1539-96 / XV-845-96. 127 p.

[4] Suomen standardoimisliitto. Hitsaus. Väsyttävästi kuormitettujen teräsrakenteiden hitsausliitosten mitoitus ja lujuuslaskenta. Standardi SFS 2378. 41 s.

[5] Fatigue of welded structures. T.R.Gurney. The Welding Institute. Second edition. Cambridge University Press. 456 p.

GEOMETRIC DEPENDENCY OF FATIGUE STRENGTH IN A TRANSVERSE LOAD-CARRYING CRUCIFORM JOINT WITH PARTIALLY PENETRATING V-WELDS

T. NYKÄNEN
V-M. LIHAVAINEN
Department of Mechanical Engineering
Lappeenranta University of Technology
P.O. Box 20
FIN-53851 Lappeenranta, FINLAND

ABSTRACT

The fatigue behaviour of one-sided partially penetrating welds in a cruciform joint has been investigated using plane strain Linear Elastic Fracture Mechanics (LEFM) calculations. A maximum tangential stress criterion with the Paris crack growth law was used to predict the simultaneous growth of a root crack and a toe crack under mixed mode K_I - K_{II} conditions. The effect of weld height, h, weld flank angle, β, and crack length, w, at the weld root (the lack of penetration) on the fatigue strength is studied. The dimensions are expressed as the terms h/t and w/t, where t = 50 mm and is the main plate thickness. The base-plate thickness is T = 80 mm. The loads had degrees of bending (DOB) of -1, -1/2, 0, 1/2 and 1, where the DOB is defined as $\Delta\sigma_b/(\Delta\sigma_m + |\Delta\sigma_b|)$, where $\Delta\sigma_m \geq 0$. The nominal bending stress range is $\Delta\sigma_b$ and $\Delta\sigma_m$ is the membrane stress range in the main plate. Theoretical fatigue strengths are compared with experimental results. The dependence of the geometric thickness effect (theoretical) on plate thickness was examined with some dimensions of the joint.

KEYWORDS

Cruciform joint, fatigue strength, theoretical thickness effect.

NOMENCLATURE

a_i = initial crack depth
a_f = final crack length
a_w = weld throat thickness
C, C_{mean}, C_{char} = constant in the Paris crack growth equation (mean, characteristic)
da/dN = crack propagation rate
h = weld height
I = crack growth integral
K_I = opening mode stress intensity factor
K_{II} = sliding mode stress intensity factor
m = constant of Paris law
N = number of cycles
n = thickness effect exponent
R = stress ratio
T = base plate thickness
t, t_0 = main plate thickness and reference thickness
w = root crack length (lack of penetration)
β = weld toe angle (fillet flank angle)
ΔK = stress intensity factor range
$\Delta\sigma, \Delta\sigma_b, \Delta\sigma_m, \Delta\sigma_0$ = nominal stress range (bending, membrane, reference)
$\Delta\sigma_{char}$ = characteristic fatigue strength, FAT, ($N = 2\times10^6$)
$\Delta\sigma_{exp}$ = experimental fatigue strength
$\Delta\sigma_{mean}$ = mean fatigue strength ($N = 2\times10^6$)

INTRODUCTION

The welding process introduces inherent surface crack-like flaws at the weld toe, i.e., along the fusion line. With fatigue loading, these preexisting flaws play a dominant role and may reduce the design stresses to a fraction of those allowed in the static loading case. Due to these preexisting flaws, the initiation stage of fatigue failure is often almost nonexistent, and these flaws can be conservatively regarded as initiated cracks. In normal quality welds, these toe cracks are the most common cause of fatigue failures. Weld root failure is also possible for partially penetrating welds. Whether it is, the toe crack or the root crack that causes the failure depends on the weld size, weld geometry and penetration, and on the way of loading.

In this work, the curved crack growth path and the simultaneous growth of the toe crack and the root crack are taken into account. Besides tensile loading, bending and combined tension/bending moment loading in both directions are examined.

Fig. 1 shows the welded joint that was used in this study. For all loadings, the critical crack initiation site is either at the upper weld toe or at the weld root. The direction of crack growth depends on the type of loading and on the dimensions of the joint. For the toe crack, an initial crack length of 0.2 mm was assumed. This length is typical when arc-welding is used. A lack of penetration forms the root crack.

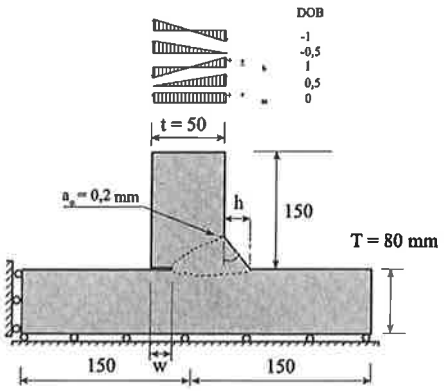


Figure 1. Fillet-welded joint.

MODELS AND METHODS

Loading and dimensions of the model

Several finite element model analyses were performed with a main plate thickness of 50 mm and a base-plate thickness of 80 mm. Five different cyclic loads combinations of tension and bending were applied at the end of the main plate, see Fig. 1. These loads had degrees of bending (DOB) of -1, -1/2, 0, 1/2 and 1, where the DOB is defined as $\Delta\sigma_b/(\Delta\sigma_m + |\Delta\sigma_b|)$, where $\Delta\sigma_m \geq 0$. The nominal bending stress range is $\Delta\sigma_b$ (on the right-sided surface of the main plate in Fig. 1) and $\Delta\sigma_m$ is the membrane stress range in the main plate. Simultaneous crack growth at the weld toe and weld root was assumed. The degree of weld penetration, weld height and weld flank angle were altered. The analysis was carried out for w/t ratios ranging from 0.004 to 1.0, h/t ratios ranging from 0 to 1.2, and for weld flank angles of 30°, 45° and 60°.

Finite element modelling and crack growth simulation

The finite element crack growth simulation programme FRANC2D/L [1] was used in the analysis. The opening mode and sliding mode stress intensity factors K_I and K_{II} were calculated using the J-integral [2] approach. The influence of K_I and K_{II} on fatigue crack growth is based on the maximum tangential stress criterion [3]. Then, the propagation path of the fatigue crack is perpendicular to the maximum principal stress. Basically the same method was used by the author earlier with reasonable success [4].

Calculation of fatigue life

The fatigue life was calculated using the Paris-Erdogan relation [5]. Paris' law for the crack growth rate is

$$\frac{da}{dN} = C\,\Delta K^m \qquad (1)$$

where da/dN is the crack growth rate per cycle, C and m are constants, and ΔK is the range of the stress intensity factor for the opening mode. The Paris law constants m = 3 and $C_{char} = 3\times10^{-13}$ ($C_{mean} = 1.7\times10^{-13}$), with da/dN in mm/cycle and ΔK in $Nmm^{-3/2}$, are recommended for the analysis of welded steel joints in Ref. [6], and are used in this study. The characteristic C_{char}-value given above corresponds to a 95% survival limit. The threshold value of the stress intensity factor was omitted. Integrating Equation (1) so that the variables, i.e., a and N, are separated produces

$$N = \int_{a_i}^{a_f} \frac{1}{C} \Delta K^{-m}\, da = \frac{\Delta\sigma^{-m}}{C} \int_{a_i}^{a_f} k^{-m}\, da = \frac{\Delta\sigma^{-m}\, I}{C} \quad \Rightarrow \quad N\Delta\sigma^3 = \frac{I}{C} \tag{2}$$

where a_i and a_f are the initial and final crack lengths, respectively. The value of the crack growth integral, I, in Equation (2) depends on the geometry of the cracked body. In the numerical integration of Equation (2), which is carryed out automatically by the FRANC2D/L program during the crack growth simulation, the final crack length a_f was considered to be reached when the increase in fatigue life N was negligible.

RESULTS

Fatigue strength

Several different models were analysed using the simulation program with a certain stress range, $\Delta\sigma = \Delta\sigma_m + |\Delta\sigma_b|$, and thus the mean crack propagation life N was determined. The stress range was then changed with Equation (2) to correspond to a fatigue life of two million cycles. The predicted mean fatigue strengths (mean fatigue classes), $\Delta\sigma_{mean}$, are presented in Tables 1, 2, 3, 4 and 5. The results for β = 45° are also presented graphically in Fig. 2.

Table 1. Mean fatigue strength $\Delta\sigma_{mean}$ for DOB = -1, ($N = 2\times10^6$).

	$\Delta\sigma_{me}$	h/t												
	MPa	0	0.1			0.3			0.6			1.2		
	β		60°	45°	30°	60°	45°	30°	60°	45°	30°	60°	45°	30°
w/t	0.004	90.0	90.5	90.0	94.3	98.0	106.3	118.2	114.5	139.2	171.0	165.7	244.0	327.0
	0.1	64.6	66.2	67.3	69.1	71.8	78.0	87.9	83.6	103.1	131.0	122.2	185.7	247.0
	0.4	29.0	29.4	31.1	31.1	33.0	37.4	44.5	42.7	55.4	75.2	70.1	117.7	167.5
	0.7	8.2	9.2	10.2	11.0	12.1	16.3	21.7	22.0	30.6	45.6	45.3	76.3	113.8
	1.0	0	0.2	0.5	0.8	2.3	4.4	6.6	7.8	15.7	22.6	29.6	53.4	82.4

Table 2. Mean fatigue strength $\Delta\sigma_{mean}$ for DOB = -1/2, ($N = 2\times10^6$).

$\Delta\sigma_{me}$ MPa		h/t												
		0	0.1			0.3			0.6			1.2		
β			60°	45°	30°	60°	45°	30°	60°	45°	30°	60°	45°	30°
w/t	0.004	76.3	74.3	76.3	76.3	75.2	76.3	77.2	77.7	80.9	84.3	86.2	95.1	103.8
	0.1	51.3	51.6	52.2	52.2	53.1	54.3	55.9	55.9	59.9	63.8	64.8	74.0	79.4
	0.4	19.6	20.2	20.6	21.4	20.9	23.2	24.7	25.2	29.6	33.9	35.3	45.1	53.1
	0.7	4.7	5.3	6.3	6.1	6.7	8.3	10.3	10.6	13.7	16.6	19.3	26.3	31.6
	1.0	0	0.2	0.5	0.4	0.4	1.9	2.7	3.3	2.7	7.6	10.3	15.5	22.1

Table 3. Mean fatigue strength $\Delta\sigma_{mean}$ for DOB = 0, ($N = 2\times10^6$).

$\Delta\sigma_{me}$ MPa		h/t												
		0	0.1			0.3			0.6			1.2		
β			60°	45°	30°	60°	45°	30°	60°	45°	30°	60°	45°	30°
w/t	0.004	71.1	63.8	63.0	62.1	61.3	59.4	57.5	59.0	57.0	55.9	58.2	59.0	61.9
	0.1	43.1	42.7	42.5	42.2	42.2	41.7	41.2	42.0	42.2	42.0	43.6	45.8	47.2
	0.4	14.7	15.2	15.5	15.1	15.2	16.5	17.1	17.9	20.0	21.2	22.8	27.6	30.4
	0.7	3.3	3.7	4.3	4.2	4.6	5.7	6.8	6.9	8.8	10.6	12.5	15.7	18.2
	1.0	0	0.1	0.4	0.3	0.7	1.2	1.7	2.2	3.6	4.7	6.1	9.0	12.5

Table 4. Mean fatigue strength $\Delta\sigma_{mean}$ for DOB = 1/2, ($N = 2\times10^6$).

$\Delta\sigma_{me}$ MPa		h/t												
		0	0.1			0.3			0.6			1.2		
β			60°	45°	30°	60°	45°	30°	60°	45°	30°	60°	45°	30°
w/t	0.004	76.3	74.9	75.8	84.3	75.2	77.7	91.6	75.5	91.5	105.7	91.3	99.0	127.0
	0.1	76.3	75.2	76.3	84.8	75.2	79.2	106.3	76.9	100.0	122.7	132.0	119.9	115.7
	0.4	59.4	57.7	58.5	60.6	55.4	58.7	54.3	59.2	58.7	58.5	65.0	68.0	73.7
	0.7	11.1	12.3	14.0	13.1	14.2	16.8	18.9	20.0	24.5	27.7	32.1	38.5	43.4
	1.0	0	0.3	0.5	0.8	1.9	3.3	4.6	5.6	9.5	11.2	14.9	21.5	26.5

Table 5. Mean fatigue strength $\Delta\sigma_{mean}$ for DOB = 1, ($N = 2\times10^6$).

$\Delta\sigma_{me}$ MPa		h/t												
		0	0.1			0.3			0.6			1.2		
β			60°	45°	30°	60°	45°	30°	60°	45°	30°	60°	45°	30°
w/t	0.004	90.0	86.6	86.4	92.8	83.9	83.6	94.4	79.9	92.0	101.0	89.6	100.0	122.7
	0.1	77.5	75.5	76.2	84.3	73.7	76.9	90.8	75.8	89.2	101.0	87.1	100.3	122.7
	0.4	36.8	36.6	37.3	44.5	38.7	45.1	69.1	47.0	72.4	101.0	68.4	96.9	123.8
	0.7	10.8	10.7	11.4	16.2	13.2	21.0	47.6	23.7	50.3	98.3	57.7	95.1	118.7
	1.0	0	0.5	1.7	6.7	4.1	12.7	42.7	14.7	48.5	98.3	53.4	94.0	119.3

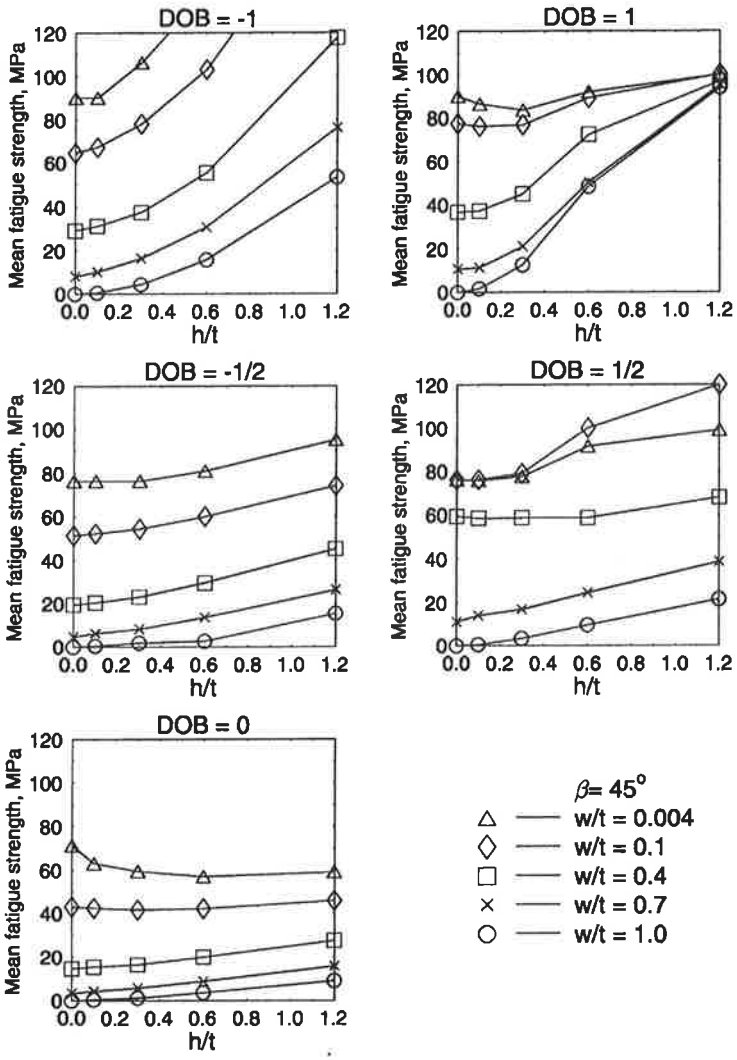


Figure 2. Theoretical mean fatigue strength, $\Delta\sigma_{mean}$ ($N = 2\times10^6$), for $\beta = 45°$.

Theoretical fatigue class

Using the $\Delta\sigma_{mean}$-values given in Tables 1, 2, 3, 4 and 5 and $C_{char} = 3\times10^{-13}$, the theoretical fatigue class (FAT) for each case analyzed can be determined. From Equation (2), we can obtain the following equation for the FAT

$$FAT = \Delta\sigma_{char} = \sqrt[3]{\frac{C_{mean}}{C_{char}}}\,\Delta\sigma_{mean} = 0.8275\times\Delta\sigma_{mean} \qquad (3)$$

with da/dN in mm/cycle and ΔK in $Nmm^{-3/2}$. (For high quality welds, the use of $C_{char} = 2.2\times10^{-13}$ might be justified, but for very poor quality welds $C_{char} = 5.4\times10^{-13}$ [7], giving $FAT = 0.9176\times\Delta\sigma_{mean}$ and $FAT = 0.6803\times\Delta\sigma_{mean}$, respectively.)

Thickness effect

The main causes of the thickness effect are the technological effect, the statistical effect and the stress gradient effect. The statistical and stress gradient effects are the main factors regarding size effects in welded joints [8]. The reason for the stress gradient effect is that a crack at the surface of a thick specimen will grow at a higher stress than a crack of the same length in a thin specimen for the same stress at the surface. Thus, the thinner specimen will have a longer fatigue life. For proportionally-scaled joints, when the crack is scaled in the same proportion as the other dimensions, the geometrical thickness effect exponent n is -1/6 in Equation (4) [4].

$$\frac{\Delta\sigma}{\Delta\sigma_o} = \left(\frac{t}{t_0}\right)^n \tag{4}$$

where $\Delta\sigma_0$ is the reference fatigue stress range for the reference thickness t_0. This geometrical size effect can be calculated using fracture mechanical models. Scaling the joints proportionally and keeping the initial crack depth at the weld toe constant, i.e., $a_i = 0.2$ mm, the $\Delta\sigma/\Delta\sigma_0$-results shown in Table 6 are obtained. The thickness ratio t/t_0 was varied from 1/2 to 4. A reference thickness of $t_0 = 50$ mm has been used because of the results presented in Tables 1, 2, 3, 4 and 5.

Table 6. Geometrical thickness effect $f(t)=\Delta\sigma_{mean}(t)/\Delta\sigma_{mean}(t_0 = 50\text{ mm})$ for proportionally scaled joints (a_i is constant at 0.2 mm), $\beta = 45°$.

	DOB		-1	-1/2	0	1/2	1
h/t = 0	t/t_0 =	1/2	1.057	1.040	1.053	1.040	1.057
$w/t_0 = 0.004$		1	1	1	1	1	1
		2	0,927	0.937	0.892	0.937	0.927
		4	0.876	0.882	0.816	0.882	0.876
h/t = 0.6	t/t_0 =	1/2	1.098	1.078	1.091	1.101	0.935
w/t = 0.4		1	1	1	1	1	1
		2	0,888	0.885	0.883	0.879	0.987
		4	0.781	0.794	0.816	0.789	0.991
h/t = 1.2	t/t_0 =	1/2	1.081	1.10	1.10	1.094	1.014
w/t = 1.0		1	1	1	1	1	1
		2	0.886	0.881	0.894	0.890	1.004
		4	0.810	0.806	0.812	0.840	0.983

The results from Table 6 are presented graphically in Fig. 3. The geometrical thickness effect is dependent on the DOB and β, and depends also on the w/t- and h/t-ratios. For comparison the thickness effect curves for n = -1/3, -1/5 and -1/6 are also presented. Theoretically, the use of a value for n of -1/6 seems to be quite safe when t ≥ 50 mm. For smaller thicknesses, its use may be very unconservative. On the basis of experimental results, a general thickness correction of n = -1/3 is proposed by Örjasäter [8]. The value of n depends on the severity of the stress concentration of the joint. For the joints with severe stress concentration like tubular joints in-plane bending the exponent is changed to n = -0.4 [8]. In practise, the use of the the commonly used 'fourth root rule' thickness correction formula (n=-1/4) or the use of more conservative value of n = -1/3 when t ≥ 50 mm might be justifiable in this case because no test data is available for the joint type studied.

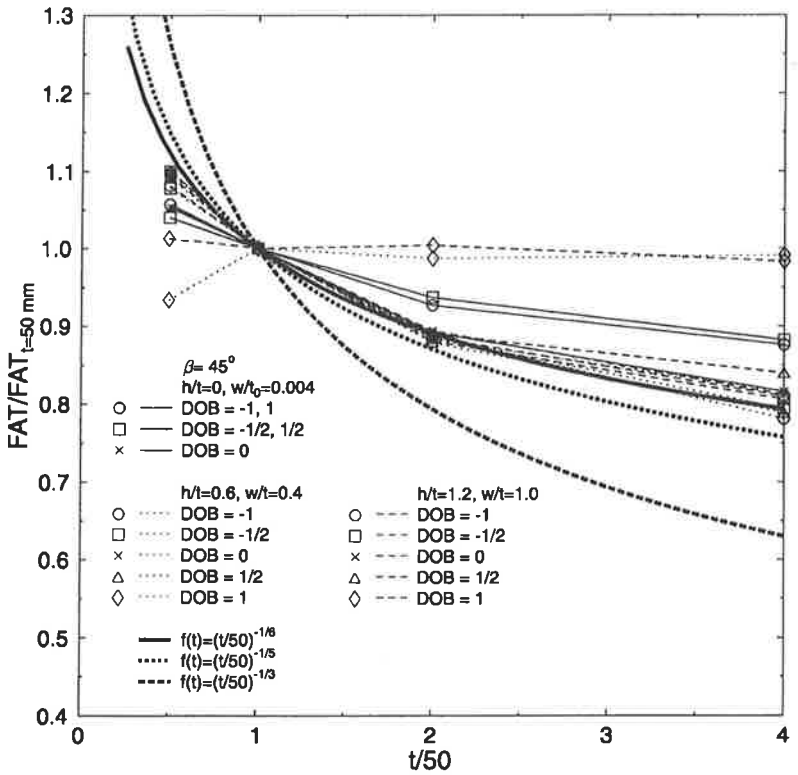


Figure 3. Geometrical thickness effect based on fracture mechanics.

Comparison with experimental results

Tests were carried out by Lihavainen and Kokkonen with constant amplitude tensile or bending loading (DOB = -1), with a stress ratio of R ≈ 0.1, and specimens fillet-welded using MAG welding and fabricated from 50 mm thick S355J0 plate steel.

The bending loading test set-up is showed in figure 4 and the mechanical properties for material S355J0 are presented in table 7. The typical fracture surface of tensile loaded test specimen is showed in figure 5. In Figure 6, the predicted fatigue strengths are compared with experimental results.

Table 7. Mechanical properties for material S355J0.

Yield strength R_{eh} [MPa]	Ultimate strength R_m [MPa]	Elongation A_5 [%]	Impact ductility 0°C [J]
335	490-630	19	27

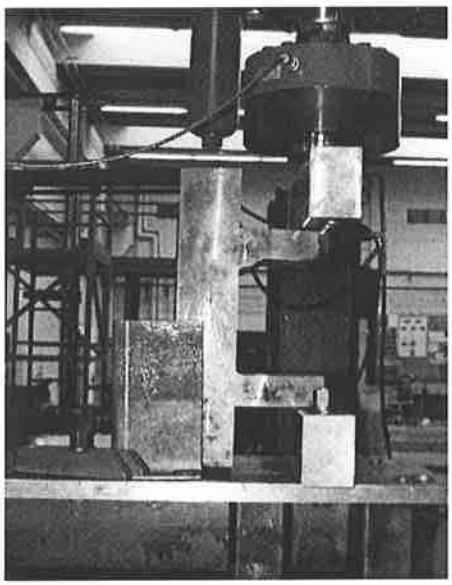


Figure 4. Bending loading test (DOB = -1).

The comparison is made using Maddox's generalized stress parameter $\Delta\sigma^*$ [5], which takes into account the different geometries of the specimen. Equation (2) can be written as

$$\Delta\sigma^{*3} N = \frac{1}{C} \tag{5}$$

where

$$\Delta\sigma^* = \frac{\Delta\sigma_{exp}}{\sqrt[3]{I}}, \qquad I = \Delta\sigma_{mean}^3 N C_{mean} \quad , N = 2\times 10^6 . \tag{6}$$

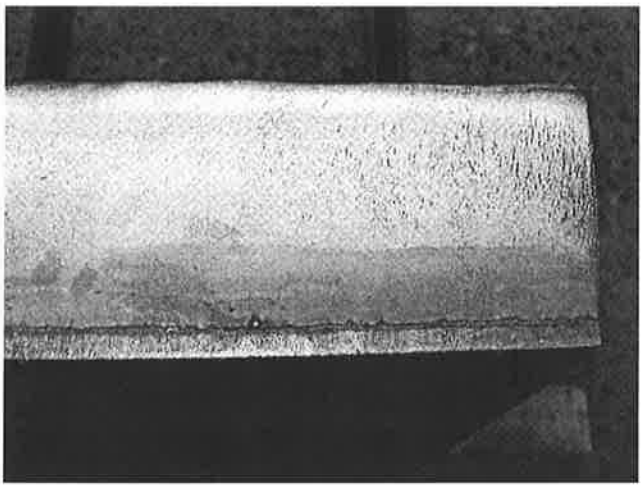


Figure 5. Fracture surface of tensile loaded test specimen.

Agreement with experimental results is quite good. It will be seen that the theoretical lower-bound line follows the experimental results reasonably well over the studied range of fatigue life.

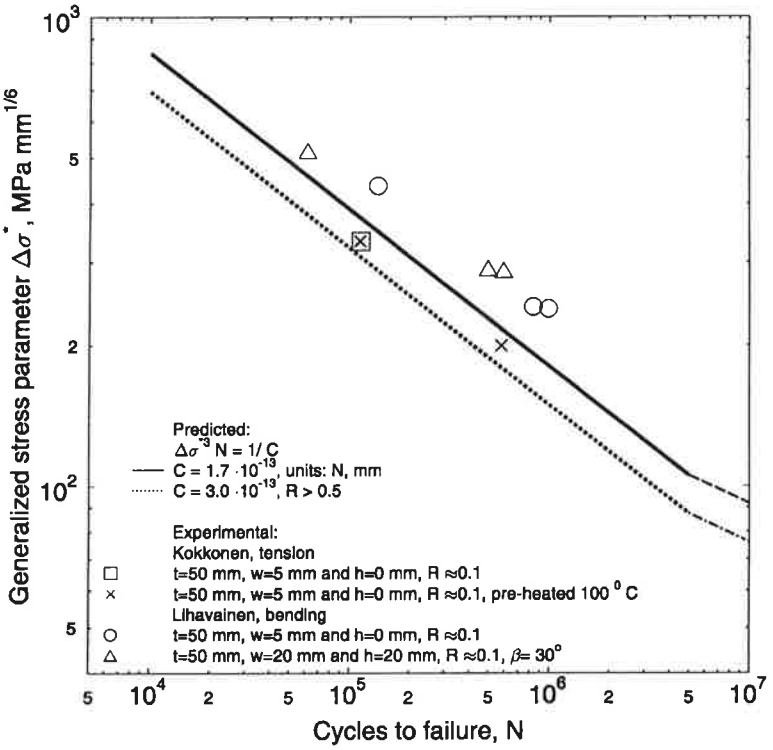


Figure 6. Fatigue test results.

The experimental fatigue strengths are compared with predicted fatigue strengths also in Figure 7. The bending fatigue strength values are 14 - 35% higher than the predicted values. On average, the experimental fatigue strengths are 10% lower for tensile loading, and 25% higher in the bending loading case, compared with the predicted fatigue strengths.

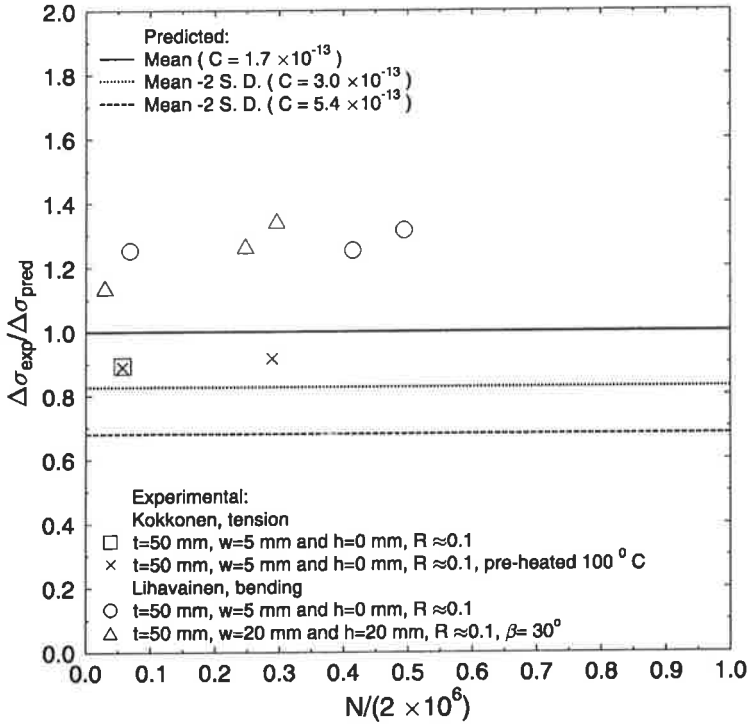


Figure 7. Comparison of experimental fatigue strength with predicted fatigue strength.

DISCUSSION AND CONCLUSIONS

The fatigue behaviour of one-sided partially penetrating welds in cruciform joints was investigated using plane linear elastic fracture mechanics calculations. The J-integral method, maximum tangential stress criterion, and Paris' crack growth law were used to predict the fatigue strength. The predicted fatigue strengths correlate reasonably well with experimental results. On average, the experimental fatigue strengths were 10% lower for tensile loading, and 25% higher in the bending loading case, compared with the predicted fatigue strengths.

It was assumed that the fatigue life of the joint could be described by the propagation of preexisting straight-fronted cracks. The initiation period and the semi-elliptical crack-front shape of an initial fatigue crack was not taken into account. This should lead to conservative fatigue strength estimates.

The joint was assumed to be in the "as welded" condition, then the whole stress intensity factor range is regarded as effective. However, for example the root crack in case of DOB = 1 does not grow in the model because of the local compression-to-compression stresses are keeping the crack closed. In practice, there may exist high tensile residual stresses or reaction stresses at the vicinity of the crack tip if the structure is highly redundant. The fatigue behaviour can then be expressed in terms of stress range alone [5]. To take the compression-to-compression loading into account, we may define DOB for $\Delta\sigma_m \geq 0$ as $\Delta\sigma_b/(|\Delta\sigma_m| + |\Delta\sigma_b|)$ and for $\Delta\sigma_m < 0$ as $-\Delta\sigma_b/(|\Delta\sigma_m| + |\Delta\sigma_b|)$ and the fatigue strength $\Delta\sigma_{mean,(DOB=-1 \text{ and } 1)} = \min\{\Delta\sigma_{mean,(DOB=-1)}, \Delta\sigma_{mean,(DOB=1)}\}$ and $\Delta\sigma_{mean,(DOB=-1/2 \text{ and } 1/2)} = \min\{\Delta\sigma_{mean,(DOB=-1/2)}, \Delta\sigma_{mean,(DOB=1/2)}\}$. The effective stress range is now $\Delta\sigma = |\Delta\sigma_m| + |\Delta\sigma_b|$.

An initial crack depth of 0.2 mm was assumed for the surface cracks. Identical fatigue behaviour was assumed for both the weld metal and the parent material. These assumptions mean that only the geometrical effects were considered. Based on theoretical analyses using fracture mechanics, the following conclusions can be drawn:

- The theoretical fatigue strength values for the weld root crack length to plate thickness ratios of w/t = 0.004 - 1.0, weld height to plate thickness ratios of h/t = 0 - 1.2, and weld flank angles of $\beta = 30°$, 45° and 60° were calculated for tensile loading, bending and combined tension/bending moment loading in both directions.
- The fatigue strength becomes greater with decreasing values of w/t and β and with increasing values of h/t.
- The fatigue strength is higher in the bending loading case.
- The fatigue strength of combined tension and bending cannot be interpolated linearly from the results of pure tension and pure bending. This is because of different crack growth paths and because of different crack initiation points.
- The geometrical thickness effect is dependent on the degree of bending and β. It also depends on the w/t- and h/t-ratios.

Acknowledgements - This paper forms part of the research project "The Fatigue Strength of Thick Walled Structures" which is carried out at the Department of Mechanical Engineering in Lappeenranta University of Techology. The author would like to express his sincere gratitude to the organizer of the project, Mr. Timo Björk, who is working at The Centre for Training and Development, which is part of the Lappeenranta University of Technology. The author also wishes to thank Metso Corporation for their financial support.

REFERENCES

1. James M. and Swenson D., "FRANC2D/L: A Crack Propagation Simulator for Plane Layered Structures," available from http://www.mne.ksu.edu/~franc2d/.
2. Dodds and Vargas, Numerical evaluation of domain and contour integral for nonlinear fracture mechanics: formulation and implementation aspects, Report from University of Illinois at Urbana-Champaign, Dept. of Civil Engineering, 1988.
3. Erdogan F. and Sih G. C., On the crack extension in plates under plane loading and transverse shear, *ASME J Basic Engng*, 85, pp. 519-527, 1963.
4. Nykänen T., On fatigue crack growth simulation in a transverse double lap joint, IIW document XIII-1583-95.
5. Gurney T. R., Fatigue of welded structures, The Welding Institute, Cambridge University Press, Cambridge 1979.
6. Hobbacher A., Recommendations for fatigue design of welded joints and components, IIW document XIII-1539-96/XV-845-96.
7. Personal communications with Professor E. Niemi, Lappeenranta University of Technology, 1993.
8. Örjasäter O., Effect of plate thickness on the fatigue of welded components, IIW-JWG XIII-XV-118-93.

FATIGUE STRENGTH OF THE SOCKET WELDS

I. POUTIAINEN

Department of Mechanical Engineering
Lappeenranta University of Technology
P.O.BOX 20
FIN-53851 Lappeenranta, FINLAND

ABSTRACT

Socket welds are widely used to connect pipes of small diameter. Many failures have been reported in these connections after a period of service. This paper presents a project that was carried out in Lappeenranta University of Technology. The aim of the project was to determine behaviour of socket welds under dynamic loading. Fatigue properties were examined by conducting a crack growth simulation for two critical locations. Calculations were verified with the fatigue test series of 24 test specimens. The test program was carried out with pulsating bending moment of a cantilever pipe. The effect of a multiaxial fatigue was studied with a combined bending and torsion test.

1. INTRODUCTION

Socket welded structures are often used in a processing industry to build pipelines of small diameter. This welded structure enables easy mounting of pipe connections. Thus these joints have been used in various locations in a pipe line, e.g. pipe-to-pipe and pipe-to-valves joints. Unfortunately, socket welds has not a good reputation. Socket welds have been present in many pipe line failures. Pipeline failure can be dangerous if this happens e.g. in nuclear power or oil refinery plants.

Pipelines are often exposed to various loads which can cause failure of pipe connections. In the case of fatigue failure of socket welded details, vibrations of a pipeline are found to be most damaging.

Socket welded details have two potential regions where the failure can occur. The first one is the weld toe, where a crack grows right through the pipe. The other one is the root of the weld, where a crack grows through the weld material. The root failure has found to

be more common. The weld root is in many way troublesome. The quality of the root is difficult to inspect after the welding and a growing root crack is impossible to detect by visual inspection before a leakage. Also, traditional fatigue improvement methods can not be applied to the root region.

In this paper a test program of one socket welded detail is reported. Test results are compared with the results reported in Ref. [1] and [2]. Finite element analyses were performed to verify strain gauge measurement of the test specimen and to assess fatigue strength of the weld toe. Hot spot stress approach was applied to weld toe. The weld toe and the weld root was analysed by using a fracture mechanics based crack growth simulation.

2. MATERIALS AND METHOD

The analysed structure in this project was a cantilever socket welded pipe joint (Fig. 1). Material of the pipe was API 5 L-B, yield stress of 340 MPa. The pipe was attached to a ¾" half-coupling by using two pass SMAW (shielded metal arc welding). The filler material was OK 48.00.

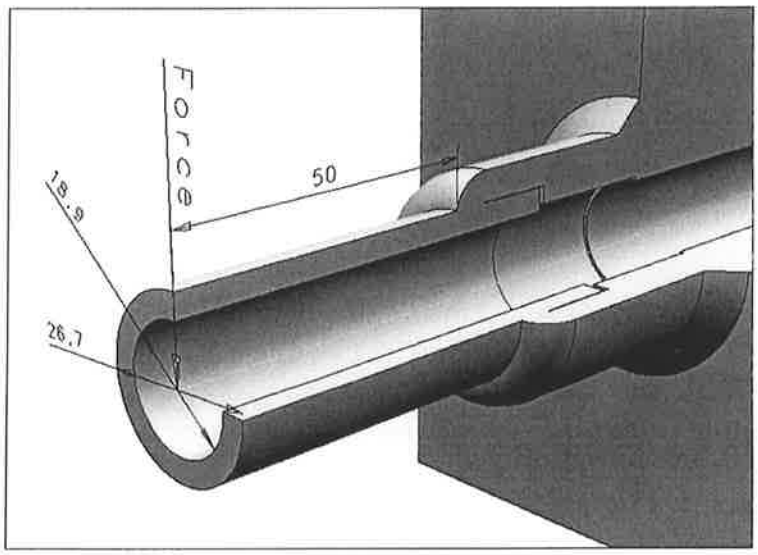


Figure 1. Socket welded detail.

2.1 Finite element model

A quarter model of the structure was modelled (Fig. 2). Parabolic solid elements were used. A symmetry restraints were applied to vertical plane and antimetric restrains to horizontal plane. The end of the coupling was fully restraint

The element size near the weld toe was 0,2*t* (*t* = 3,91 mm). Stresses were plotted from the distances 0,4*t* and 1,0*t* from the weld toe. These stresses were afterwards used to calculate structural stress in the weld toe [3].

Model was constructed, solved and analysed by using *IDEAS Master Series ™ 7.0* package.

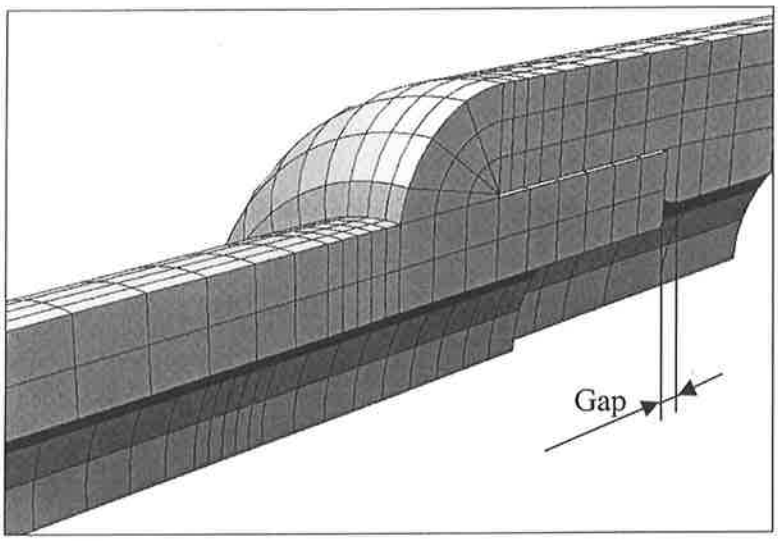


Figure 2. Finite element model of the socket welded joint.

2.2 Crack growth simulation

The fatigue strength of the weld root was examined by simulating crack growth in the root area. Simulation was performed by using FRANC3D (FRacture ANalysis Code for 3 Dimensions) program, developed by The Cornell Fracture Group [4].

2.3 Fatigue test

Fatigue tests were conducted to a total of 24 test specimens. The main object of these tests, was to determine a fatigue limit for this detail. A fatigue limit is a value for stress range, under which failure does not occur in constant amplitude loading. In this study, the fatigue limit was determined at 10 million cycles.

When installing a pipe to a coupling, the pipe should not touch the bottom of the coupling. This might produce undesirable weld cracking in the weld root. According to most construction codes there should be about 1.6 mm gap between the pipe end and the bottom

of the coupling. Here, a 2 mm gap was adopted to the test specimens. However, six test specimens were manufactured and tested with zero gap.

The test specimens were loaded with a pulsating (R=0.1) bending of a cantilever socket welded pipe (Fig. 3). The pulsating force was placed to the distance of 50 mm from the weld toe. Four fatigue tests were performed by using combined bending and torsion loading. In this case, a cylinder force was moved 45 mm away from the pipe centreline.

During the fatigue test a force range, strains near the weld toe (4 mm from the weld toe) and displacements of the cylinder were recorded. A 10% increase in the cylinder displacements was chosen for the criteria of the failure.

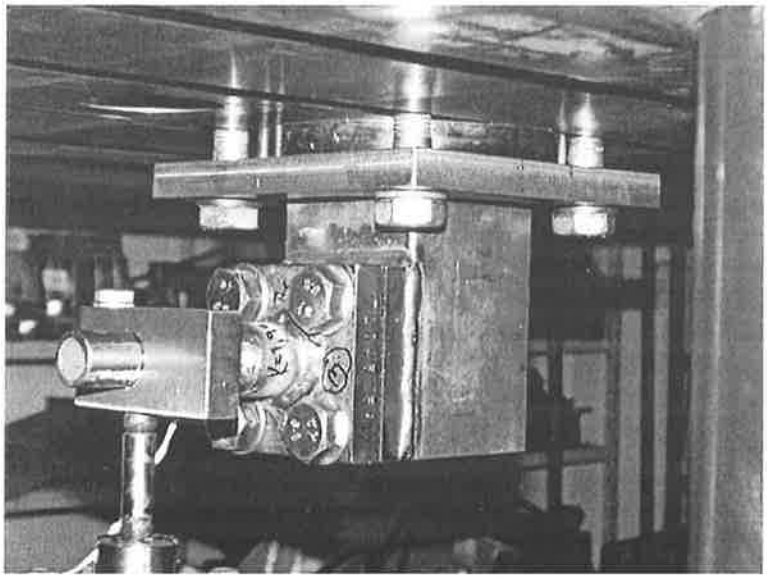


Figure 3. Fatigue test arrangement.

3. RESULTS

3.1 Finite element results

The weld toe is one critical location in the socket welded structure. Here, the weld toe was analysed by using both the structural stress approach and the fracture mechanics calculations. The structural stress were obtained from the surface stresses at maximum tension stress area. Here, critical site was on the top part of the pipe (force downwards).

Structural (hot spot) stress is obtained by extrapolating the axial top surface stresses at the distances 0,4*t* and 1,0*t* (Fig. 4) to the weld toe. The stress concentration factor for the structural stress in this detail is 1.7.

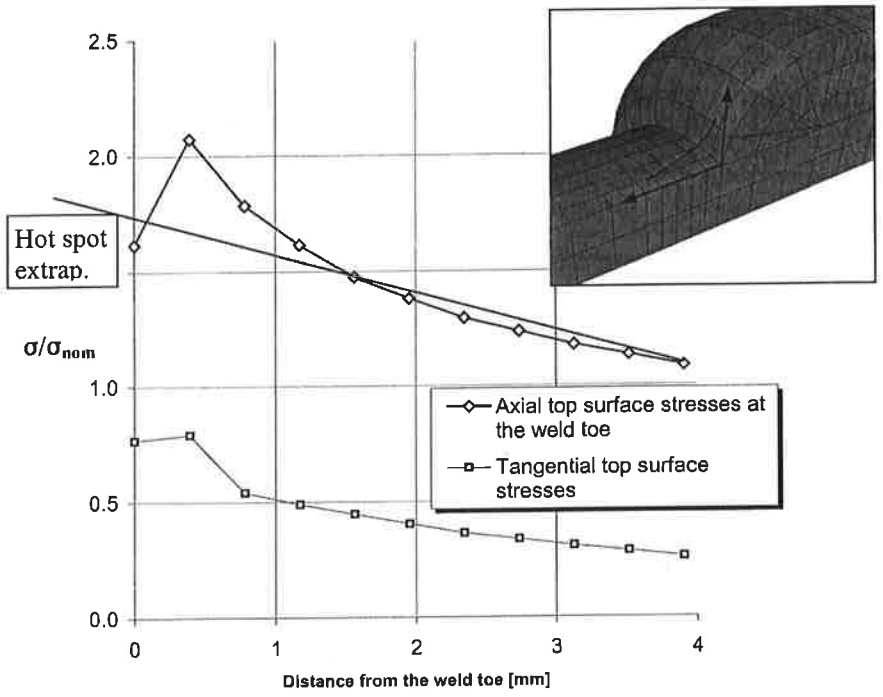


Figure 4. Stresses at the weld toe in bending load.

3.2 Crack growth simulation

The crack growth simulation was adopted both to the weld root and the weld toe. In both locations a initial crack of size 0.1 mm was modelled in the worst possible direction i.e. perpendicular to maximum principle stress. Crack were then let to grow to size where it nearly penetrated through the surface. In the simulation model, bending loading was set to create 150 MPa nominal stress to the weld toe.

Fatigue life was calculated by using the Paris law with the stress intensity history, obtained from the crack growth simulation, and constants m=3 and C_{mean}=1.7E-13 [5].

In the weld root case (Fig. 4), a leak would occur after 1,89 million cycles. The initial crack in the weld toe would grow through the pipe thickness after about one million cycles.

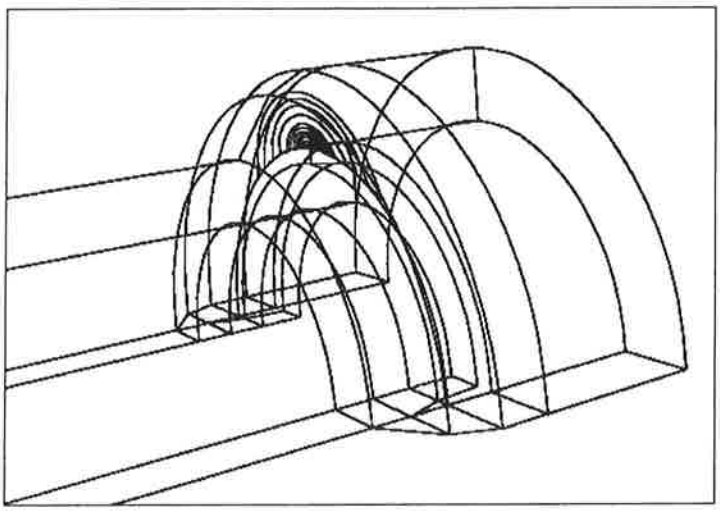


Figure 5. Crack propagation through the weld.

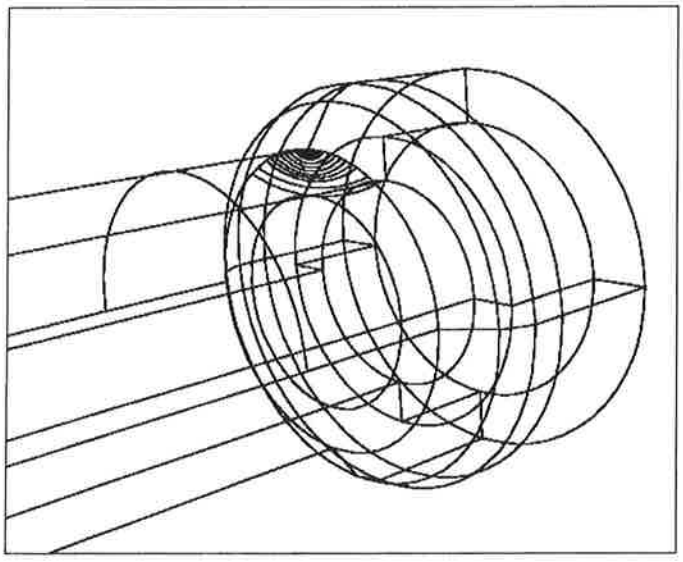


Figure 6. Crack propagation through the pipe thickness.

3.3 Fatigue test results

The results of the fatigue tests are collected in Fig. 7. The stress range in Fig. 7 is a nominal stress at the weld toe. A mean hot spot –curve (FAT 135) is converted to mean nominal stress -curve of this structure by dividing it with the stress concentration factor 1.7.

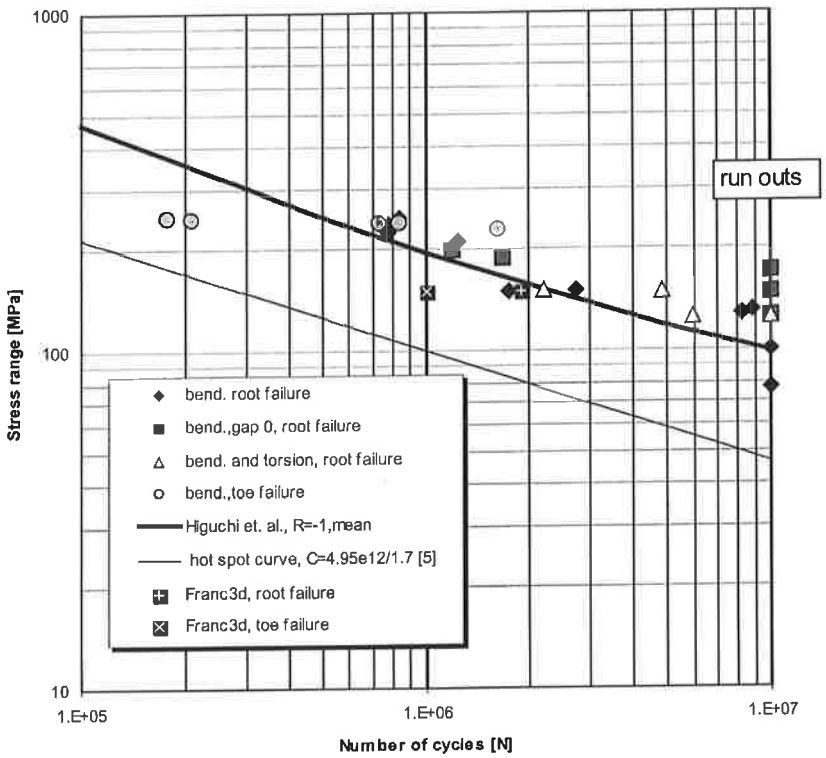


Figure 7. Collection of the fatigue test results.

4. DISCUSSION AND CONCLUSIONS

The toe failure occurred only with the high stress range values (Fig. 7). In these stress levels, the hot spot stress exceeded the yield stress of the pipe material and the failure was expected. At lower stress range levels the failure initiated from the weld root.

The crack growth simulations predicted that the toe failure would be more crucial. As the test results reveals, the weld toe calculations underestimates the true fatigue strength. However, the weld root results are in the same line with simulation results. The root weld analysis had fewer geometrical uncertainties than in the weld toe case. This is also a reason that root failure occurred with a relatively small scatter. The weld toe calculations are very sensitive to a local geometry of the weld toe and to the size of a initial crack. The hot spot stress method also predicts lower fatigue life to weld toe. This indicates that the weld toe, in this case, has exceptionally good fatigue strength.

are equal in every direction. The axial strength of the joint (pullout strength) does not normally reach the level of shear strength and depends on it. Thereby, this kind of joint is most suitable in applications where the loading of the joint comes mainly from shear forces. A typical joint made by method in question is illustrated in Fig. 1.

Fig. 1. A typical joint made by noncutting, single stroke clinching method.

The diameter of the round clinch (or the diameter of the die) was chosen to be eight millimetres. Moreover, the thickness of the sheets to be joined was taken to be one millimetre and the material of the sheets was hot-zinced, high strength structural steel of grade 550 MPa (Ragal 550 S). The depth of the die was chosen to be 1.4 mm and the diameter of the punch was taken to be 5.6 and 5.4 mm respectively. Based on the tests, these combinations gave the best shear strength with the material used and with one-millimetre sheet gauge and therefore were the most interesting ones.

The main objective of this work was to establish so-called C-value (Fig. 2) using numerical simulation of the clinch process. This value is needed, when we want to determine an optimal value for the punch stroke that can be illustrated with parameter X (Fig. 2). The X-value is needed when setting process parameters for the clinching machine and it is depending among other things on the thickness of the sheets to be joined and the pair of tools (punch-die combination). Many tests have been carried out to determine the C-value with different tool combinations by Varis [1]. Therefore, it was under interest to find out if the determination could be done quite easily using finite element method and MARC analysis software and hence replace the tests at least partly. Furthermore, the workability of the model based on the C-value in order to determine the X-measure wanted to be verified. Finally, also to gain a better understanding of the clinch process by simulating it numerically was one of the interests.

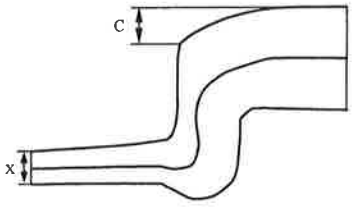


Fig. 2. Definition of the C- and X-values.

2 FE-ANALYSIS USING MARC ANALYSIS SOFTWARE

The simulation was carried out using MARC analysis software, version K7.3.2. Pre- and postprocessing was performed using Mentat pre- and postprocessor, version 3.3.0r2 [2]. The definition of contacts is rather easy in MARC software.

2.1 The element model

Because of the axisymmetric nature of the clinched joint, only a plane model with four-node axisymmetric quadrilateral elements was needed. This element is number 10 in the element library of MARC [3]. This low-order element is preferred over higher-order elements when contact analysis is involved. An example of the element model is shown in Fig. 3. Tools were modelled as rigid bodies and sheets as deformed bodies on the contrary. Since the axisymmetry was utilised, only a half-model was created. The symmetry boundary condition was modelled as a rigid body with symmetry flag set. This makes the rigid body (in this case, a plane) to act as a symmetry plane, i.e. the nodes residing on that plane are kept on it with no friction. The outer edges of the sheets were fixed both in radial and axial directions.

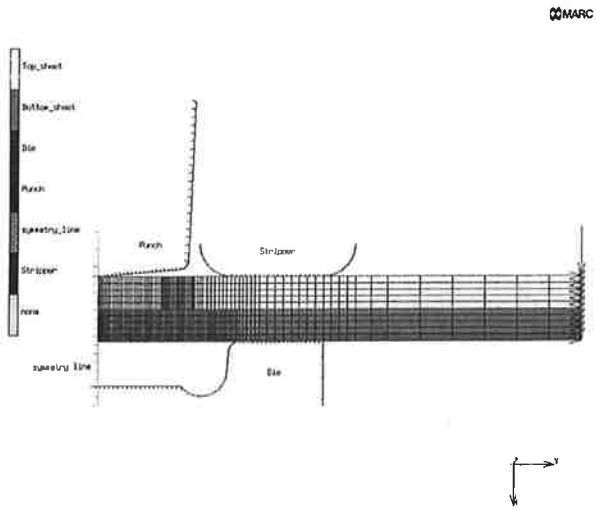


Fig. 3. An example of the element model.

There is a possibility to remesh the model during the analysis in MARC software. This is done by using a parameter REZONE [4]. Before this can be done, enough space must be allocated for the new nodes and elements at the beginning of the input data file with the SIZING parameter. The space allocation must be done also in contact definition sections. Rigid bodies cannot be redefined. The basic idea of the rezoning is that the element mesh can be straightened and elements and nodes can be added or deleted in the areas needed during the analysis. This is often necessary in problems like this in which strains are quite large and elements distort easily, in which case the accuracy of the analysis suffers and convergence problems can arise. One analysis can contain several REZONE increments.

In this work, the rezoning procedure was tried at first, but it proved to be not entirely trouble-free in this kind of analysis that contains rather complicated contact analysis (two sheets with a contact between and in addition a contact between tools and sheets). The rezoning could namely not be made to work properly, but the contact analysis failed in post-rezone increments. Therefore, this technique was abandoned and the excessive element distortion could be avoided by applying an adequate fine mesh to critical areas at the beginning of the analysis at once. This procedure was quite workable with a die depth in question (1.4 mm).

The definition of the contact and friction is rather easy in MARC; elements belonging to each deformable body must be defined and stated to be in a set with a given name. The sheets to be joined are deformed bodies in this case. The definition of rigid (non-deformable) bodies is easy too. In a plane model these can be defined as curves, which can be straight line segments, circular arcs, splines or nonuniform rational bspline surfaces (NURBS). The geometry of the rigid body can be modelled in Mentat using its geometry definition commands, but these primitives are always converted to NURBS when writing the input data file for MARC. The direction of the curves is important since the contact occurs only in one direction. Mentat has an option that shows the directions, and a flip-function can be used in order to change the directions if needed. The user must just remember to check these directions. The "wrong" side of the rigid body is shown as streaked in Fig. 3. If the analysis contains many bodies involving contact between them, a contact table can be used in order to define the complex contact situation. The contact table defines between which bodies there can be a contact as well as the friction coefficient between bodies. The separation of the contact can be defined to base on nodal forces or nodal stresses. In this case, nodal stresses were used, which is recommended, since by that means the effect of the element size can be eliminated.

There are several friction models in MARC software that can be used, among other things shear, Coulomb and stick-slip models [5]. In this work, the most common model was used and the friction was modelled as Coulomb friction. In Coulomb friction model of MARC, the friction coefficient between surfaces and a parameter RVCNST, is needed. Physically, the value of RVCNST is the value of the relative velocity when sliding occurs. A very large value of RVCNST results in a reduced value of the effective friction. A very small value results in poor convergence. The value of this parameter is recommended to be between 1% and 10% of a typical relative sliding velocity. By default, the value of RVCNST is 1.0; in this case a value 0.1 was used.

The loading of the model was defined by giving a linear motion for the punch. The punch was given a velocity of one m/s and consequently time in the analysis means directly the distance the punch has moved. The spring-loaded stripper around the punch was modelled as force driven rigid body. The stripping force depends on pre-stressing force (4756 N) and the movement of the punch with the spring constant being 580 N/mm. The very first models did not contain the stripper at all.

2.2 The material model

The material model has rather significant importance in simulation of this kind. In this case, available material data were the proof strength $R_{p0.2}$, the tensile strength R_m and the percentage elongation after fracture A_{80}. These values were measured in both rolling direction and transverse to that. Stress-strain curves were not available and thus the material model had to be built solely based on the data listed in Table 1. The anisotropy of the material was not considered, on the contrary, it was modelled as isotopic. The value for the modulus of elasticity and the Poisson's ratio were E = 210 MPa and 0.3 respectively.

TABLE 1.
MEASURED MECHANICAL PROPERTIES OF RAGAL 550 S.

Direction	$R_{p0.2}$ [MPa]	R_m [MPa]	A_{80} [%]
Rolling	638	663	10.4
Transverse	692	719	4.7

Different material models were tried. These models can be divided mainly in two categories: such in that the work hardening is not taken into account and such in that it is considered as linear (or piece-wise linear) curve. When the work hardening was not considered, only yield stress was needed in addition to Young's modulus and Poisson's ratio. Stress-strain data is given to MARC software as slope-breakpoint pairs or alternatively as σ-ε pairs that are converted to first ones by the software. Data points are always given as ε-σ pairs in Mentat preprocessor. It is also important to notice that when giving the stress-strain data, only plastic part of the strain should be given. This means that the first point of the data has following values: yield stress for stress and zero for strain.

The option PLASTICITY,3 was used as analysis parameter in FE-analysis. This option turns on parameters LARGE DISP (large displacement or buckling i.e. geometric stiffness matrix), UPDATE (updated Lagrange procedure) and FINITE (finite strain plasticity). The FINITE parameter enables rather large strains, up to 3%, per one increment. When using these parameters, it must be taken into account that the slope of the work hardening is defined as a ratio of true stress change and true plastic strain. Therefore, the data of σ-ε curve must be given as true stresses and logarithmic plastic strains, which means that the conventional stress-strain curve obtained from uniaxial test must be converted to right form.

2.2.1 True stress and logarithmic strain

True stress means simply the ratio of load P and current cross-sectional area A rather than the original area A_i, which is normally used in stress-strain curves obtained from uniaxial tensile tests [6].

$$\tilde{\sigma} = \frac{P}{A} \tag{1}$$

Hence, true stress and engineering stress are related by

$$\tilde{\sigma} = \sigma\left(\frac{A_i}{A}\right) \tag{2}$$

True strain is obtained by integrating the change of strain between initial and final length

$$\tilde{\varepsilon} = \int_{L_i}^{L} \frac{dL}{L} = \ln\frac{L}{L_i} \tag{3}$$

where L_i is the initial length and $L = L_i + \Delta L$ is the final length. Noting that $\varepsilon = \Delta L / L_i$ is the engineering strain leads to a relationship between true strain (logarithmic strain) and engineering strain.

$$\tilde{\varepsilon} = \ln\frac{L_i + \Delta L}{L_i} = \ln\left(1 + \frac{\Delta L}{L_i}\right) = \ln(1+\varepsilon) \tag{4}$$

For materials that behave in a ductile manner, plastic strains are large compared with elastic strains. Thus, it is reasonable to approximate the volume as constant.

$$A_i L_i = AL \tag{5}$$

This gives

$$\frac{A_i}{A} = \frac{L}{L_i} = \frac{L_i + \Delta L}{L_i} = 1 + \varepsilon \tag{6}$$

By substituting the above into Eqs. (2) and (3) two additional equations relating true and engineering stress and strain is obtained:

$$\tilde{\sigma} = \sigma(1+\varepsilon) \tag{7}$$

$$\tilde{\varepsilon} = \ln\frac{A_i}{A} \tag{8}$$

Equations (4) and (7) were used to convert the stress-strain curve in this study. These equations are valid only up to the beginning of the necking, but since the information about the tensile tests of the material was not precise enough, this was acquiesced in. Another way to perform the conversion was to use Eqs. (1) or (2) for stress and Eq. (8) for strain, which are valid up to breaking, but the use of these equations requires the measurement of the cross-sectional area during the test. However, this kind of information was not in use.

3 RESULTS AND DISCUSSION

The development of the C-value as a function of the punch stroke was observed as a result of the numerical simulation. This could then be compared with experimental data. The curve based on FE analysis does not contain the springback, which is included in the experimental data, but usually this action is rather negligible. It should also be mentioned that by using FE analysis, the X-measure could be established directly, thus leaving the determination of the C-value unnecessary. However, the C-value is needed in analytical model [1], hence these values were determined. The influence of the material properties and friction was studied by variating the both friction and material data.

The deformed model with different punch strokes is presented in Fig. 4. The contour of the joint agrees quite well with a real joint.

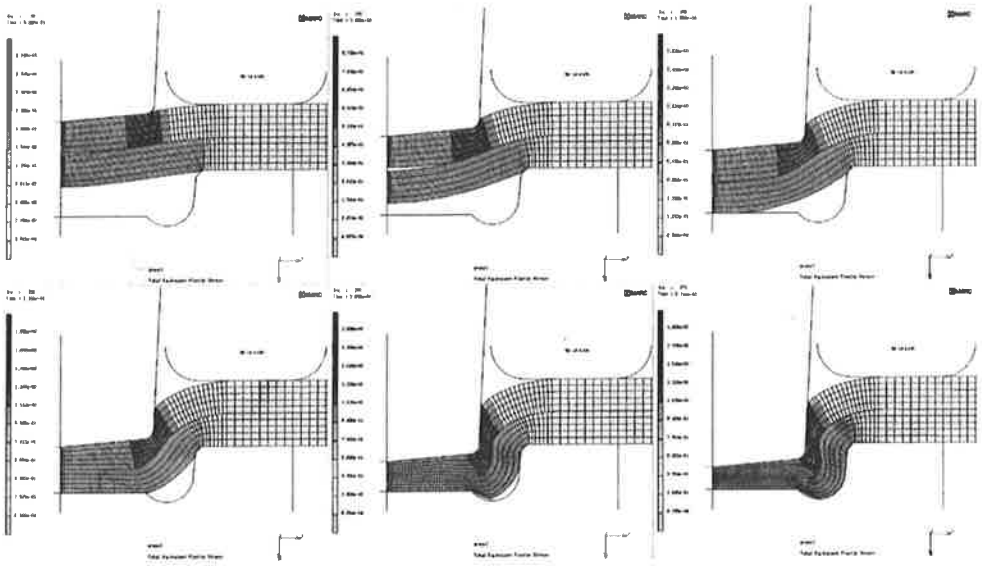


Fig. 4. Deformed structure with different punch strokes. The stroke beginning from upper left corner: 0.5, 1.0, 1.5, 2.0, 2.5 and 2.71 mm respectively.

The influence of the different material properties and friction is presented in Fig. 5. As can be seen, the friction does not have a significant influence on the C-value, hence it could be ignored when defining the C-value. However, the friction does have an influence on the force need, especially in the end of the stroke when the material is filling the groove in the bottom of the die. On the left side of the Fig. 5, a clear division in two categories can be seen depending on the fact if the work hardening is modelled or not. In model 56-14-m10h, the work hardening was modelled in such way that after the logarithmic strain corresponding with test strain A_{80}, the work hardening slope was bisected.

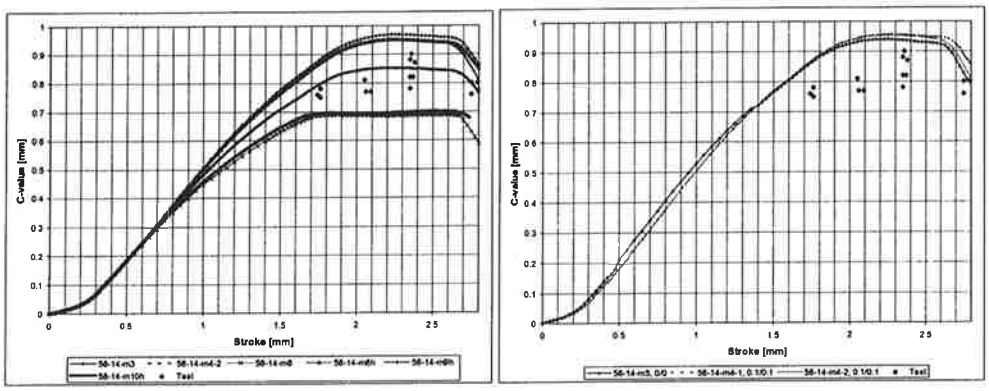


Fig. 5. The influence of different material properties and friction on the C-value. On the left different material models and on the right the influence of the friction.

A summary of C-value curves with different values for material properties, friction and punch diameter is presented in Fig. 6. The essential information about different models is presented in Table 2. The C-value is between 0.68…1.02 in all calculated models, as can be seen in Fig. 6. In other words, the range of C-value is about 0.3 mm, regardless of the modelling of work hardening. The C-value begins to decrease when the stroke is between 2.56…2.7 mm. This means that so called counter piping occurs i.e. the material begins to climb upwards between the tools and is detached from the top face of the die. This phenomenon can be clearly seen when examining the deformation plots increment by increment. Generally, no counter piping is allowed, but usually some kind of compromise must be done. However, no more counter piping is allowed that is needed to fill the die. If the counter piping is too extensive reduces this the strength of the joint and hence this can be regarded as one criteria of the quality of the joint.

TABLE 2.
DESCRIPTIONS OF STUDIED ELEMENT MODELS

Model	Work Hard.	Yield Stress	Friction Coeff.	Strip-per	Notice
56-14	N	667	0	N	
56-14-m2	Y	667	0.13/0.15	N	Friction coeff. between contacts tool-sheet/sheet-sheet
56-14-m3	Y	667	0.1	N	Same friction coefficient in all contacts
56-14-m4-2	Y	640	0.1	N	
56-14-m4h	Y	640	0.1	Y	Last breakpoint at strain corresponding with $A_{80}/2$
56-14-m5	Y	640	0	N	
56-14-m6	N	640	0.1	N	
56-14-m7h	Y	640	0.1	Y	Stress-strain curve as engineering stress-strain curve
56-14-m8h	Y	640	0.1	Y	Young's modulus E=200 MPa
56-14-m9h	Y	640	0.1	Y	Work hardening slope = 0 after ε corresponding with A_{80}
56-14-m10h	Y	640	0.1	Y	Work hard. slope bisected after ε corresponding with A_{80}
54-14	N	667	0	N	
54-14-m2	Y	667	0.13/0.15	N	

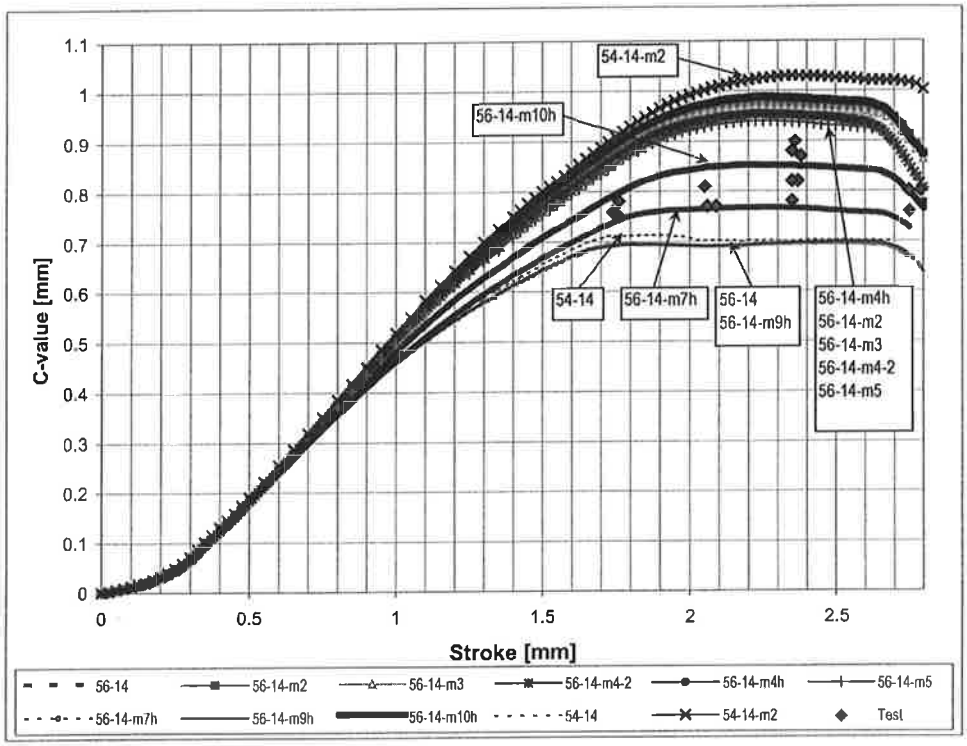


Fig. 6. A summary of C-value curves with different values for material properties, friction and punch diameter.

4 CONCLUSIONS

The finite element method and MARC software can rather easily be used to simulate the clinching process that produces an axisymmetric joint using noncutting, single stroke method. The optimal penetration of the punch can be established experimentally, but by simulating the process using FE analysis can the amount of this work be reduced and suitable tool pairs can be defined for materials used at a time. The numerical simulation helps also to understand the process itself better. MARC software is quite suitable for this kind of analysis in consequence of easy contact definition.

ACKNOWLEDGEMENTS

This work is done in a project *Design of Sandwich Panel Constructions*, which belongs to a technology programme KENNO - *Lightweight Panels*, funded by National Technology Agency (TEKES) and several companies from industry. All involved participants are gratefully acknowledged.

REFERENCES

[1] Varis, J. A novel procedure for establishing clinching parameters for high strength steel sheet, Dissertation (in print), Lappeenranta, 2000.

[2] Mentat 3.3-MARC K7.3: New Features, MARC Analysis Research Corporation, Palo Alto, USA. 1998.

[3] MARC® Volume B: Element Library, Version K7.3. MARC Analysis Research Corporation, Palo Alto, USA, 1998.

[4] MARC® Volume C: Program Input, Version K7.3. MARC Analysis Research Corporation, Palo Alto, USA, 1998.

[5] MARC® Volume A: Theory and User Information, Version 7.3. MARC Analysis Research Corporation, Palo Alto, USA, 1998.

[6] Dowling, N. E. Mechanical Behavior of Materials. Engineering Methods for Deformation, Fracture, and Fatigue. Second Edition, Prentice-Hall, Inc., New Jersey, USA, 1999.

TELESKOOPPIPUOMIN MALLINTAMINEN TASOSSA

H. MARJAMÄKI
VTT Automaatio
PL 1307
33101 Tampere

TIIVISTELMÄ

Tässä esityksessä esitetään eräs menetelmä, jolla teleskooppipuomin mallinnus voidaan tehdä tasossa elementtimenetelmää käyttäen. Puomi kootaan käyttäen Reissnerin kinemaattiseen malliin perustuvia palkkielementtejä. Toisiinsa liittyvät liukuvat puomin osat on liitetty käyttäen erityiselementtejä, joissa palkkielementtiin liittyy liuku-jousi-elementti, joka on kytketty orjuutustekniikkaa käyttäen palkkielementin neutraaliakselille. Teleskooppiliike on toteutettu käyttäen hydraulisylinterielementtiä, jonka hydrauliset vapausasteet on kytketty puomin vapausasteisiin. Sylinteri liikuttaa vain ensimmäistä jatkopuomia ja seuraava liukuva puomijäsen on kytketty edelliseen ketjuja kuvaavalla välittäjäelementillä.

Syntyvä hydro-mekaaninen laskentamalli muokataan tavalliseksi differentiaaliyhtälöryhmäksi ottamalla tilamuuttujaan puomin siirtymävapausasteiden lisäksi myös sen solmupisteiden nopeudet. Yhtälöryhmä ratkaistaan adaptiivista aika-askelta käyttävällä implisiittisellä Runge-Kutta menetelmällä.

JOHDANTO

Monet työkoneet, joiden tehtävänä on nostaa tavaraa tai ihmisiä, on varustettu teleskooppipuomilla. Teleskooppipuomilla tarkoitetaan palkkirakennetta, jossa on tietty määrä sisäkkäin asetettuja palkkeja eli teleskooppijaksoja, jotka liukuvat toisiinsa nähden. Teleskooppiliike saadaan aikaiseksi tavallisesti hydraulisylinterillä, joka on kytketty ensimmäisen ja toisen jakson välille. Mikäli jaksoja on enemmän kuin kaksi, niin muita jaksoja liikutetaan yleensä sisään- ja ulosvetoketjujen välityksellä.

Tällaisen puomiston mallinnus FEM -valmisohjelmalla on hankalaa ja edellyttää sidosehtojen antamista toisiinsa nähden liukuvien teleskooppijaksojen välillä, myöskään hydraulisylintereiden kuvaamiseen ei ole valmiita elementtejä. Seuraavassa esityksessä on esitetty vaihtoehtoinen tapa mallintaa teleskooppipuomi käyttäen välitysketjuja, hydraulisylinteriä [5] sekä liukumista kuvaavia erityiselementtejä.

Mekaanisen järjestelmän laskentamalli on ohjelmoitu käyttäen suurten siirtymien kinematiikkaa. Edelleen mekaanisen järjestelmän ainemalli on oletettu lineaariseksi. Liikettä toteuttavat hydraulisylinterit on liitetty suoraan syntyvään laskentamalliin.

Kytketty hydromekaaninen tehtävä ratkaistaan yhtenäisenä järjestelmänä, jonka etuina hajautettuun järjestelmään verrattuna on yksinkertaisuus, numeerinen stabiilius ja laskentatehokkuus. Yhtenäisen stabiilin järjestelmän aikaintegrointi on numeerisesti ehdoitta stabiili, jos aikaintegrointialgoritmi on ehdoitta stabiili. Näin asian laita ei ole hajautetulle järjestelmälle.

RAKENNEOSIEN MALLINNUS

Palkki- ja sauvaelementit

Palkkielementteinä on käytetty lähteen [2] mukaisia elementtejä. Elementissä on kolme solmuvapausastetta. Palkkielementti perustuu Reissnerin kinemaattiseen malliin.

Siirtymämittauksen vaihto

Niin sanottujen offset-elementtien, eli palkkielementtien, joiden solmut on siirretty pois neutraaliakselilta ja liukuelementtien sisäisten voimien, tangentiaalisen jäykkyys- ja massamatriisin laskeminen voidaan suorittaa seuravassa esitettävällä orjuutusperiaatteella [3].

Olkoon kahden eri siirtymämittauksen välillä tunnettu yhteys:

$$\mathbf{u} = \mathbf{f}(\mathbf{v}) \tag{1}$$

Siirtymän **u** variaatio voidaan lausua siirtymän **v** variaation avulla

$$\delta\mathbf{u} = \frac{\partial \mathbf{f}}{\partial \mathbf{v}}\,\delta\mathbf{v} = \mathbf{B}(\mathbf{v})\,\delta\mathbf{v} \tag{2}$$

Oletetaan edelleen, että derivoïtuvan kuvauksen **B** rangi on täysi. Koska molemmissa mittausjärjestelmissä tehty virtuaalinen työ tulee olla yhtä suuri, on vastaavien voimamittausten välillä on yhteys:

$$\mathbf{F}_v = \mathbf{B}^T\mathbf{F}_u \tag{3}$$

Derivoimalla voimien välinen yhteys (3) siirtymäkentän **v** suhteen, saadaan

$$\frac{\partial \mathbf{F}_v}{\partial \mathbf{v}} = \frac{\partial \mathbf{B}^T}{\partial \mathbf{v}}\mathbf{F}_u + \mathbf{B}^T\frac{\partial \mathbf{F}_u}{\partial \mathbf{u}}\mathbf{B} \tag{4}$$

Elementin solmuarvoista laskettu kineettinen energia, T nopeuskentän arvolla du/dt on

$$T = \frac{1}{2}\dot{\mathbf{u}}^T \mathbf{M} \dot{\mathbf{u}} \qquad (5)$$

Derivoimalla siirtymien välinen yhteys (1) saadaan mittausjärjestelmien nopeuksien välille yhteys

$$\dot{\mathbf{u}} = \frac{\partial \mathbf{u}}{\partial t} = \frac{\partial \mathbf{f}}{\partial \mathbf{v}} \frac{\partial \mathbf{v}}{\partial t} = \mathbf{B}\dot{\mathbf{v}} \qquad (6)$$

Liike-energia voidaan lausua mittauksessa **v**

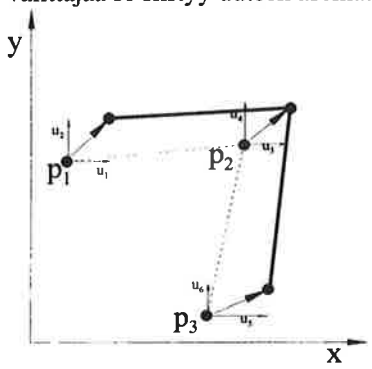
$$T = \frac{1}{2}(\mathbf{B}\dot{\mathbf{v}})^T \mathbf{M}(\mathbf{B}\dot{\mathbf{v}}) = \frac{1}{2}\dot{\mathbf{v}}^T \mathbf{B}^T \mathbf{M} \mathbf{B} \dot{\mathbf{v}} \qquad (7)$$

jolloin siirtymämittauksen **v** massamatriisi saadaan lausuttua siirtymämittauksen **u** avulla

$$\mathbf{M}_v = \mathbf{B}^T \mathbf{M}_u \mathbf{B} \qquad (8)$$

Teleskooppipuomin ketju

Kuvataan teleskooppipuomissa olevaa kahta ketjua, jotka välittävät teleskooppisylinterin liikkeen teleskooppijatkeille yhdellä elementillä, joka vetosuunnassa vastaa köyttä. Elementti ottaa kuitenkin köydestä poiketen vastaan myös puristuspuolen kuormitukset. Tämä käyttäytyminen on puhtaasti tarkoituksenmukaisuuskysymys ja tarvittaessa elementti voitaisiin korvata kahdella vain vetokuormituksen vastaanottavalla elementillä. Kuvassa 1 on esitetty välittäjäelementin, joka kulkee lepotilassa pisteiden $\mathbf{p}_i$ (i=1,2,3) kautta, periaatekuva. Kuormitettaessa välittäjää se siirtyy uuteen asemaan, jota kuvaa kunkin pisteen siirtymä $\mathbf{u}_i$.



Kuva 1. Teleskooppipuomin ketjua kuvaava laskentamalli

Ketjun omasta massasta aiheutuva ekvivalenttinen solmukuormitus on siirtymäriippuvainen. Solmulle 2 tuleva kuormitus voidaan kuitenkin olettaa olevan puolet koko ketjun massan aiheuttamasta kuormituksesta. Loppukuormitus voidaan jakaa solmuille 1 ja 3 osapituuksien mukaan.

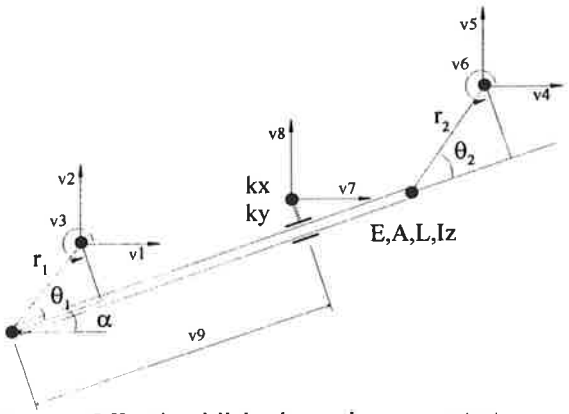
Välittäjän tangentiaalinen jäykkyysmatriisi saadaan derivoimalla sisäisen voiman vektori solmusiirtymillä

$$\mathbf{k}_t = \frac{\partial \mathbf{q}_{int}}{\partial \mathbf{u}} \qquad (9)$$

Ketjun massamatriisi on myös siirtymäriippuva. Tässä esityksessä on käytetty keskitettyä massamatriisia, jossa keskisolmulle kohdistuu puolet massavaikutuksesta. Loppumassa on jaettu osapituuksien suhteessa.

Liukuelementti

Puomien liukuvien osien kuvaamisessa käytetään kuvan 2 mukaista liukuelementtiä. Elementin sisäiset voimat saadaan yhteyttä 4 käyttäen. Elementillä on 9 vapausastetta, joista viimeinen on dimensioton välillä [0,1] oleva liu'un asemaa kuvaava vapausaste.



Kuva 2. Offset-jousi-liukuelementin vapausasteet

Elementit sisäiset voimat lasketaan käyttäen isäntäelementtinä offset-palkkielementtiä, johon orjaelementin eli liukujousen sisäiset voimat lisätään käyttäen yhteyttä (3). Vapausasteelle 9 tulee vielä sisäiseksi voimaksi jousivoiman komponentti vapausasteen suuntaan.

Liukuelementin jäykkyysmatriisi saadaan taas orjuutustekniikkaa käyttäen. Jousielementin jäykkyys summataan isäntäelementtinä olevaan offset-palkkielementin jäykkyyteen käyttäen yhteyttä (4).

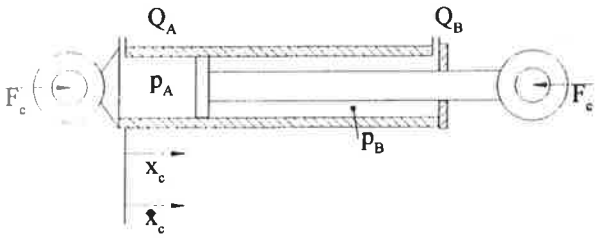
Massamatriisina voidaan käyttää solmuille 1 - 6 kuuden vapausasteen palkin massamatriisia, joka saadaan tavallisen palkkielementin massamatriisista käyttäen yhteyttä (8). Vapausasteille 7-9 massamatriisin alkioiksi on käytetty liukuun osallistuvien liukupalojen arvioitua massaa.

Hydraulisylinterin mallintaminen

Staattisella kuormituksella hydraulisylinteri voidaan kuvata sauvaelementtiä käyttäen. Sylinterin eri asemat voidaan kuvata sauvan muuttuvana pituutena. Mikäli halutaan mallintaa myös sylinterin mahdollinen nurjahdustilanne, niin sauvaelementti voidaan korvata palkkielementeillä.

Ratkaistaessa laskentamalli dynaamisella kuormituksella voidaan hydraulisylinteri mallintaa käyttäen erityiselementtiä. Mekaaniseen järjestelmään sylinterielementti kytketään jakamalla sylinterin silmukoissa vaikuttava voima mekaanisen järjestelmän vapausasteille.

Tarkastellaan kuvan 3 mukaista hydraulisylinteriä, jossa kammiopaineet p_A ja p_B oletetaan vaikuttavan koko kammiotilavuuden alueella.



Kuva 3. Hydraulisylinterin periaatekuva

Hydraulisylinterin kammiopaineiden p_A ja p_B muutoksille saadaan yhteys

$$\begin{Bmatrix} \dot{p}_A \\ \dot{p}_B \end{Bmatrix} = B \begin{Bmatrix} \frac{Q_A - A_A \dot{x}_C}{V_A + A_A x_C} \\ \frac{-Q_B + A_B \dot{x}_C}{V_B - A_B x_C} \end{Bmatrix} \qquad (10)$$

missä alaindeksi A viittaa sylinterin männän ja B varren puolelle ja

Q	on sylinteriin tuleva tai lähtevä tilavuusvirta
A	on sylinterin painepinta-ala
V	paineenalainen alkutilavuus
B	paineväliaineen kokoonpuristuvuuskerroin

Sylinterin silmukoihin vaikuttava voimaresultantti saadaan vähentämällä kitkavoima painevoimien erotuksesta

$$F_C = A_A p_A - A_B p_B - F_\mu \qquad (11)$$

Kitkavoimalle voidaan käyttää erilaisia laskentamalleja, joista tässä työssä on käytetty lähteen [6] mukaista kitkamallia. Silmukoihin vaikuttava voima jaetaan mekaanisen järjestelmän vastaaville vapausasteille kohdistuviksi.

Laskentamallin ratkaiseminen staattisella kuormituksella

Puomirakenteeseen liittyvät ulkoiset kuormitukset oletetaan tunnetuksi. Kuormitus koostuu työlaitteen aiheuttamista kuormista, puomirakenteen omasta massasta, hitauskuormista, tuulikuormista sekä annetuista ulkoisista kuormista.

Kokoamalla puomirakenteen tangentiaalinen jäykkyysmatriisi saadaan rakenteen siirtymä alkuasennosta määritettyä elementtimenetelmällä [1]. Mikäli siirtymät alkuasemasta ovat suuria on tasapainoyhtälö epälineaarinen. Tavanomaisissa puomirakenteissa voidaan tasapaino yleensä saavuttaa Newton' in menetelmällä laskemalla kullakin iterointikierroksella rakenteen siirtymämuutos ulkoisen kuormituksen ja puomirakenteen sisäisistä voimista laskettujen voimien erotuksesta eli epätasapainosta ja jatkamalla iterointia kunnes epätasapaino on riittävän pieni.

Lisäksi voidaan käyttää suoralta hakua tai kaarenpituusmenetelmää, mikäli ratkaistaan rakenteen stabiliteetin menettämistilanteita. Edelleen joissain laskentatilanteissa voidaan käyttää myös siirtymäohjausta.

Laskentamallin ratkaiseminen dynaamisella kuormituksella

Tässä työssä laskentamallin vaste dynaamisella kuormituksella ratkaistaan kirjoittamalla laskentamallin liikeyhtälöt tavalliseksi differentiaaliyhtälöryhmäksi. Ratkaistava tilasuure, $\mathbf{x}$ sisältää tällöin laskentamallin hydraulisten vapausasteiden ja siirtymäkoordinaattien lisäksi siirtymäkoordinaattien aikaderivaatat eli nopeudet. Näin syntyvä yhtälöryhmä on standardimuotoa:

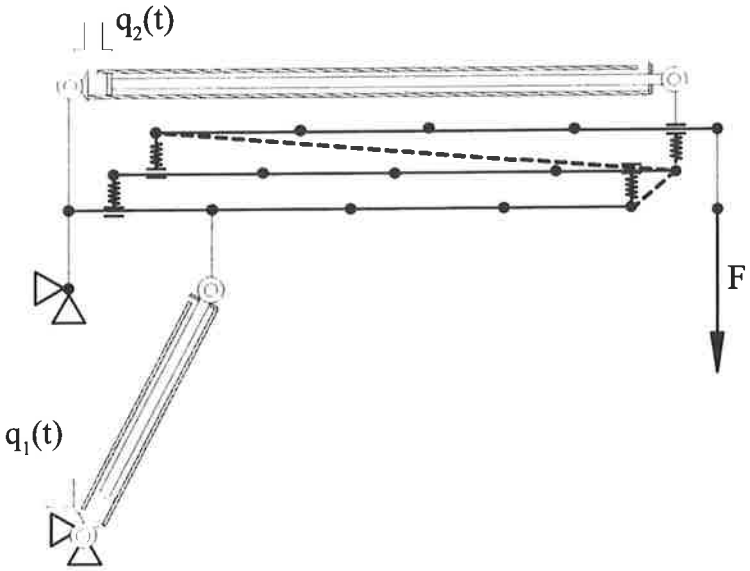
$$\begin{cases} \mathbf{M}\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t) \\ \mathbf{x}(t_0) = \mathbf{x}_0 \end{cases} \qquad (12)$$

ja se voidaan ratkaista kankeisiin (stiff) differentiaaliyhtälöryhmiin soveltuvalla implisiittisellä aikaintegrointimenetelmällä [4]. Tässä työssä on käytetty adaptiivista aika-askelta käyttävää, Rosenbrock- menetelmää [7], jossa aika-askeleen aikana syntyvää virhettä arvioidaan toisen ja kolmannen kertaluvun ratkaisun erotuksen normilla.

Ratkaisussa kannattaa hyödyntää Jacobin matriisin harva rakenne sekä mekaanisen järjestelmän tangentiaalisten alimatriisien symmetria.

LASKENTAESIMERKKI

Laskentamenetelmän testaamiseksi laadittiin elementtimetelmäohjelmisto, johon mallinnettiin kuvan 4 mukainen teleskooppipuomi. Puomissa on kaksi teleskooppijatketta ja sitä voidaan nostaa nostosylinterin avulla. Teleskooppiliike toteutetaan kuvassa ylimpänä olevalla sylinterillä sekä katkoviivalla esitetyllä ketjua kuvaavalla välittäjäelementillä. Kuvassa olevat teleskooppiliikkeeseen liitttyvät elementit on piirretty selvyyden vuoksi eri korkeuksille. Laskentamallissa ne sijaitsevat samalla suoralla.

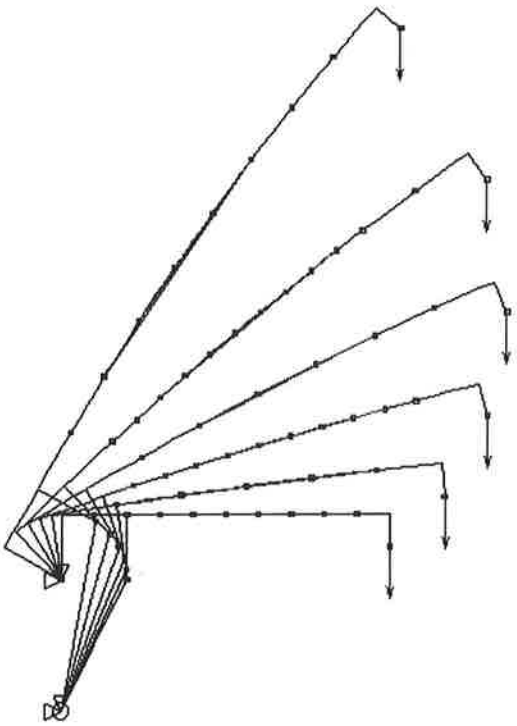


Kuva 4. Teleskooppipuomin laskentamallin periaatekuva

Laskentatilanteessa nostosylinterille annettiin 5 sekuntia kestävä pulssimainen tilavuusvirta q_1 siten, että se nostaisi joustamattoman puomin 60 asteen kulmaan vaakatasosta. Samaan aikaan teleskooppisylinterille annetaan tilavuusvirta, joka vastaa koko teleskooppiliikettä. Puomin päähän sijoitettiin lisäksi vakiona pysyvä 40000 N pistevoima.

Laskentatulokset

Kuvassa 5 on esitetty teleskooppipuomin siirtymäkenttä sekunnin välein laskennan alkamisesta.



Kuva 5. Esimerkkirakenteen siirtymäkenttä sekunnin välein laskennan alkuhetkestä

JOHTOPÄÄTÖKSET

Koska laskentamallista syntyvä differentiaaliyhtälöryhmä ratkaistaan ilman algebrallisia sidosehtoja, niin ratkaisualgoritmista saadaan verrattain nopea. Lisäksi hydraulisylintereiden vaikutukset kokonaisjoustoon saadaan mallinnettua suoraan.

Luontevana jatkona edellä esitetylle tasomekanismeihin soveltuvalle laskennalle on ohjelmoida vastaava kolmiulotteisena laskettava joustava mekanismi.

LÄHTEET

1. Bathe, K-J. 1996. Finite Element Procedures. Prentice-Hall Inc, New Jersey. 1039 s.

2. Crisfield, M.A., 1991. Non-linear Finite Element Analysis of Solids and Structures, Vol. 1: Essentials, J. Wiley & Sons, New York. 345 s.

3. Crisfield, M.A. 1997. Non-linear Finite Element Analysis of Solids and Structures, Vol. 2: Advanced Topics, J. Wiley & Sons, New York.

4. Hairer, E., Wanner, G., 1991. Solving Ordinary Differential Equations II, Stiff and Differential-Algebraic Problems, Springer Series in Computational Mathematics 14, Springer-Verlag, Berlin, 601 s.

5. Mäkinen, Ellman and Piche 1997. Dynamic simulations of Flexible Hydraulic-Driven Multibody Systems using Finite Strain Beam Theory, Linköping CP.

6. Olsson, H. 1996. Control systems with friction, Lund Institute of Technology, 172 s.

7. Piché, R. An L-stable Rosenbrock method for step-by-step time integration in structural dynamics, Computer Methods in Applied Mech. Engrg. 126 (1995) 343 – 354.

SLIDING BEARING WEAR DAMAGE SIMULATION AT MICROSTRUCTURAL LEVEL

H. MARTIKKA* and M. KUOSA†
*Faculty of Construction Design, Department of Mechanical Engineering
† Department of Energy Technology
Lappeenranta University of Technology, P.O.Box 20,
FIN-53851 Lappeenranta, FINLAND

ABSTRACT

Background for this study is an increasingly important need to control the tribological behaviour of machine element contacts for instance in sliding bearings and gears. In this study results of studying dry friction mechanisms of metals, steel and lead bronze, are reported. The simulations using LS-DYNA FEM virtual models of the microstructural mechanisms give detailed and visual scenarios of the contact phenomena at the surface roughness level and of the large deformation process involved. Several microstructural design variables like roughness topography, nonlinear material behavior parameters are used to minimize wear. The results show that simulations can be advantageously used to predict and optimize sliding bearing behaviour on micro and macrolevels. They can also be used to explain the relationship between acoustic emission signal and phenomena and damage in sliding contacts.This information is technically and economically important in many installations.

1. INTRODUCTION

Background for this study is need to control the tribological and total behaviour of some industrial machinery.This has guided attention on getting control over the basic microstructural mechanisms influencing wear. The method is to simulate the behaviour of contacting surfaces with or without lubricating films. One tool is the LS-DYNA [1] program for simulating metal deformation, cutting and fracturing processes and for analyzing nonlinear response of three-dimensional inelastic structures using contact-impact of complex forms. By this method the microstructure at surface roughness level is modelled using LS-DYNA program to simulate the elasto-plastic-fracture impacts causing wear damage and fracture behaviour. At present only dry friction contact results are reported. Also some basic dislocation mechanics is used to explain damage mechanisms. The next goal in future is to use the present methods to study elastohydrodynamic lubrication with various oils and various surface layers on the base materials.

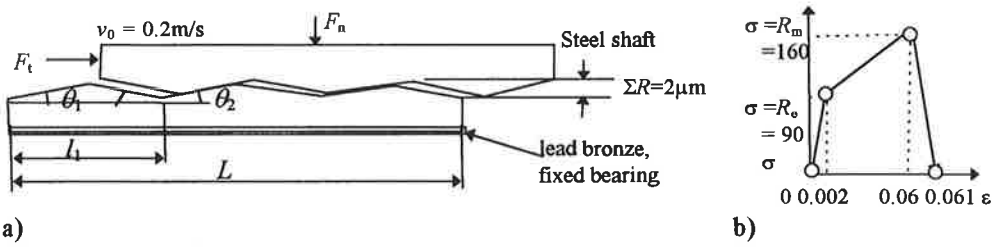
2. MATERIALS AND METHODS

The material used and the FEM input data are presented in the following.

2.1 The model used in FEM and in analytical simulations

2.1.1 The geometry of the model

The geometry of the model is shown in Fig. 1a.



a) b)

Figure 1. a) Main dimensions of the system. Angles vary from 0, ±5°,±10°. Length of the model is L = 0.185mm, wave length $l_1 = 2\Sigma R / \tan\theta \approx 40\mu m$ .Initially the upper body is in contact with the lower. It has a large mass and its is initial velocity is v_o = 0.2m/s horizontally , b) Material model.

2.1.2 Materials

The bearing is lead bronze,LPT120 Promet 420 Cu76 Pb20 Sn4. Its yield 0.2% limit is R_{p02} = 90MPa, ultimate tensile strength R_m = 160 MPa, elongation to fracture δ_5 = 6 %, brinell hardness HB = 500 MPa. Elastic modulus is E_2 = 70000MPa, Poisson's ratio ν = 0.35.The shaft material is quenched and hardened steel MoC410M, with yield 0.2% limit R_{p02} = 690MPa, ultimate tensile strength R_m = 950 MPa, ν = 0.3.

The surface is modelled using the following assumption.

The surface contains a large number of asperities in the form of cones of equal base angle. In the FEM models these are nodelled as a saw tooth profile with randomly varying base angle ranging rationally within θ = 0, ±5°,±10°

2.1.3 FEM model and results

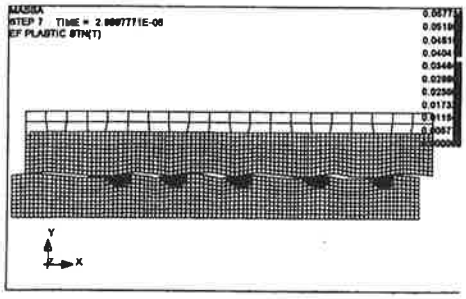
The material model of LS-DYNA is model 24 piecewise linear isotropic plasticity. Deviatoric stresses s_{ij} are determined that satisfy the yield function [1]

$$\phi = \tfrac{1}{2} s_{ij} s_{ij} - \tfrac{1}{3}\sigma_y^{\,2} \le 0 \quad , \quad \sigma_y = \beta\left[\sigma_0 + f\left(\varepsilon^p_{eff}\right)\right] \quad , \quad \beta = 1 + \left(\frac{\dot{\varepsilon}}{C}\right)^{1/p} \tag{1}$$

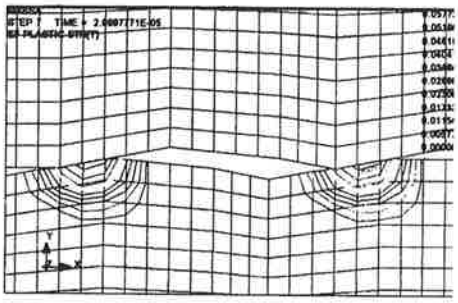
Here σ_y is the yield stress, f is the hardening function (Figure 1), $\dot{\varepsilon}$ is strain rate. Stress-strain and strain rate curves are user given. Now the failure criterion is based on plastic strain. When strain .06 is exceeded then the element is eroded away.In Figure 2 some results are shown. In Figure 2d the maximum shear stress contours are shown with max values of 90 MPa.The material is still elastic.The maximal shear stresses close to surface are important for the zero wear model according to McGregor [2].

Figure 2. FEM model results. The upper mass is given an initial velocity v_o and large kinetic energy $K = \frac{1}{2}Mv_o^2$. It is transformed at plastic frictional contacts to material damage and heat.The critical strain for erosion away is $\varepsilon_{p,Rm} = .06$. Upper material is steel and lower is lead bronze.

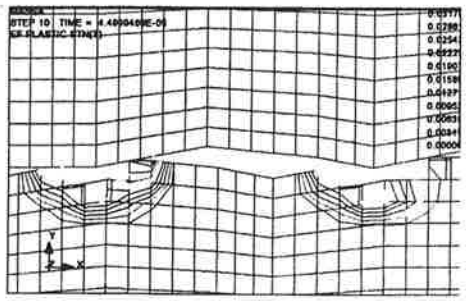
a) & b) The effective plastic strain below 0.06, at step 7.
c) Effective plastic strain at step 10 showing surface wear as erosion of elements,
d) Maximum shear stress contours, max value is 90MPa in bronze at step 7.
e) Surface wear profiles at step 9 and 14.
f) Plastic strain at step 25 showing large wear.



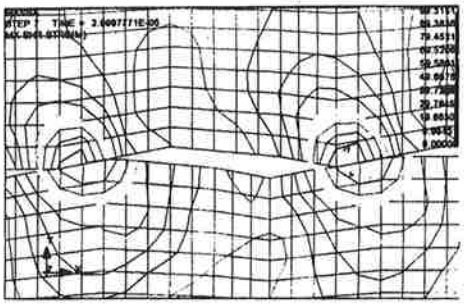
a)



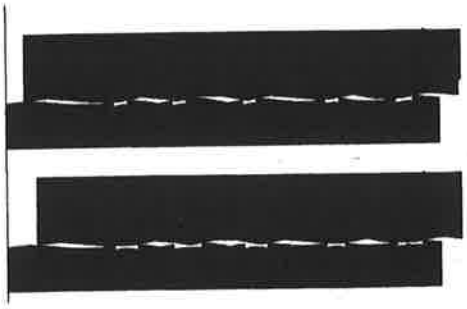
b)



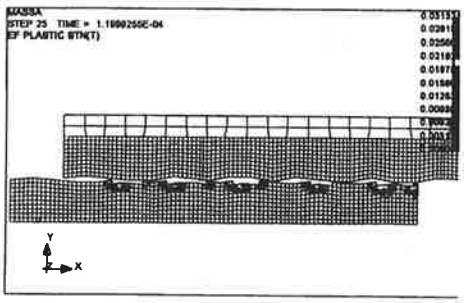
c)



d)



e)



f)

3. MECHANICS OF CONTACT AND SURFACE ROUGHNESS

The effects of surface topography on wear may be considered using the model of two contacting surface profiles shown in Figure 3. This is modified from Stout & al. [3] and Tsukizoe [4].Stout & al. [3] have observed that in cases when the harder surface has a clearcut surface profile then it will truncate the softer profile and impose its own profile on the softer. But if the harder surface profile has a helical or irregular texture then it will continue to wear away evenly the softer one.

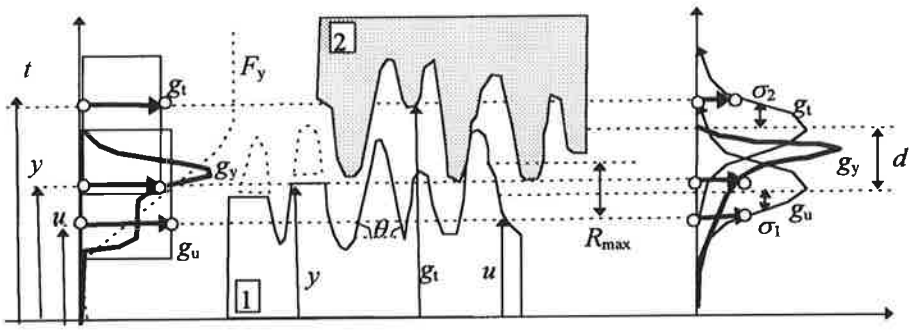


Figure 3. Simple wear models of truncation of a harder body 2 on by a softer 1.The unworn surface profile height is u and distribution is g_u , the harder truncating profile is g_t and the worn profile is g_y

Let us denote the distances measured from a reference level by u,t,y and the respective probability density functions ,or pdf's, and cumulative density functions , cdf's, as

- u is distance to the unworn profile curve , its pdf is g_u and cdF is F_u
- t is distance to the truncated profile curve , its pfd is g_t and cdf is F_t
- y is distance to the worn surface profile curve ,its pdf is g_y and cdf is F_y

The probability that a truncation takes place within a coordinate range y, $y+dy$, is [3]

$$\Pr\{y < Y < y+dy\} = g_y(y)dy \quad , \quad \Pr\{-\infty < Y < y\} = \int_{-\infty}^{y} g_y(y)dy = F_y(y) \tag{3}$$

$$g_y(y)dy = \Pr\{\ [y < u < y+dy] and [t > y]\ \ or\ \ [u > y] and [y < t < y+dy]\ \} \rightarrow$$
$$g_y(y) = g_u(y)[1-F_t(y)] + g_t(y)[1-F_u(y)] \tag{4}$$

As shown in Figure 3 the worn profile has been formed assuming two models. One simple model is based on rectangular distributions and the other has gaussian distributions. The truncated distributions are similar in form with peak close to the surface.

According to Tsukizoe [4] the base angle θ of conical asperities can be expressed as

$$\tan\theta = \tfrac{\pi}{2}\tan\theta_p \quad , \quad \tan\theta_p = \frac{R_{max}}{3.5R_{max}+15} \tag{5}$$

here $\tan\theta_p$ is mean value of the tangent of slope angle on profile curve of surface.
R_{max} is maximum height of asperities defined in JIS (Japanese Industrial Standard).

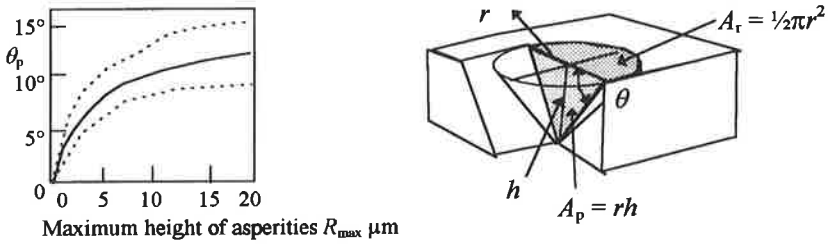


Figure 4. a) Dependence of the means slope of the profile curve θ_p on the maximum height of asperities R_{max}, [4] , b) geometry of a conical asperity.

In the case of n identical conical asperities the coefficient of friction is

$$\mu = \frac{F_p}{F_N} = \tan\theta_p = \frac{A_p\sigma_0}{A_r\sigma_0} = \frac{n\cdot r^2\tan\theta}{n\cdot\frac{1}{2}\pi r^2} = \frac{2}{\pi}\tan\theta \tag{6}$$

A slip line field model for contact is shown in Fig.5.

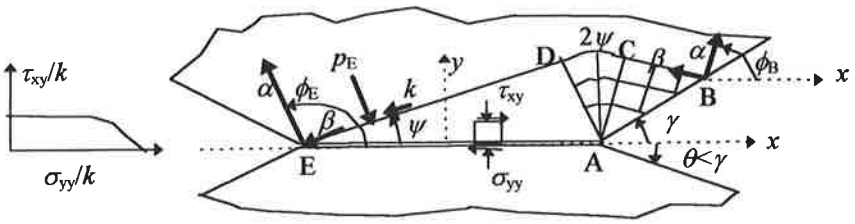


Figure 5. A slip line field model of plastic contact between surface profiles, a) model, b) yield surface, according to McClintock and Argon [5].

The yield surface is described by the shear and normal stress and the coefficient of friction is obtained by choosing the contact surface defined by condition $\psi = 0$.

$$\sigma_y = \left[1+\tfrac{\pi}{2}-2\gamma+2\psi+\sin 2\psi\right]k \quad , \quad \tau_{xy} = \cos 2\psi\cdot k \quad , \quad k = R_e/2$$
$$\psi\to 0 \quad , \quad \to \mu = \frac{\tau_{xy}}{\sigma_y} = \frac{1}{1+\frac{\pi}{2}-2\gamma} \tag{7}$$

Here the slope angle γ is about the same as the base angle θ. Now θ = 0.1 rad and coefficient of friction is μ = 0.4.Several conditions for predicting whether contact is plastic or not have been developed. Mikic and Roca [6] have presented the following criterion

$$\psi_3 = \frac{E'}{H}\tan\theta \quad , \quad \frac{1}{E'} = \frac{1-\nu_1^2}{E_1} + \frac{1-\nu_2^2}{E_2} \tag{8}$$

If $\psi_3 \geq 0.25$ then at least 2% of asperities will deform plastically. Here E' is equivalent elastic modulus.Now E'=76000MPa, H = 270MPa, $\tan\theta$ =.1,then Ψ_3=28 indicating plastic deformation.

Scuffing is affected by surface topography. Dyson [7] has presented a flow function F criterion for estimating risk for scuffing. The model is applicable on a condition $d/\sigma < 3$. here d is the mean line separation of two surfaces and σ is the composite surface finish.

$$F = 1-0.9e^{-0.58d/\sigma} \quad , \quad \sigma = \left[\sigma_1^2+\sigma_2^2\right]^{1/2} \tag{9}$$

If $d/\sigma > 3$ then the flow function F = 1 and it is the same as for smooth surfaces.

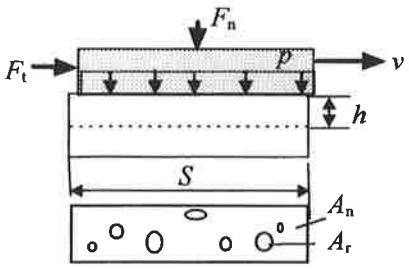
4. GENERAL WEAR MODELS

The generally used linear wear model is as follows

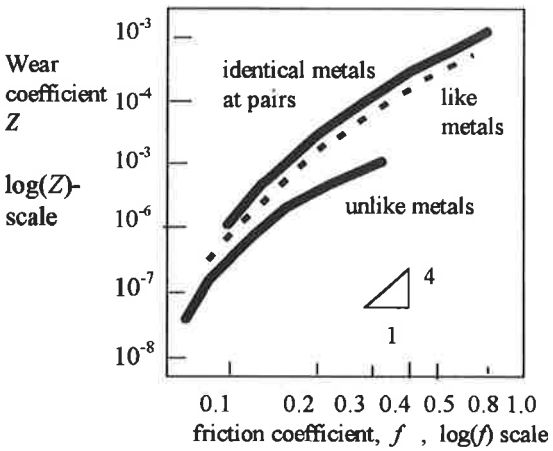
$$\frac{V}{S} = Z\frac{F_n}{H} \quad , \quad V = A_n h \quad h = \dot{h}t \quad , \quad S = vt \quad , \quad F_n = pA_n \quad \rightarrow \quad \dot{h} = pv\frac{Z}{H} \tag{10}$$

where

V is worn out volume (m^3), F_n is normal compressive force (N)
A_n is nominal area (m^2), A_r is the true contact area (m^2)
p is nominal pressure on nominal area (Pa), h is worn out layer thickness (m)
t is time (s), S is sliding distance (m), v is sliding velocity (m/s)
Z is wear coefficient, H is hardness or the wearing away surface



a)



b)

Figure 6 .Wear models. a) Contact model . b) The relationship between the wear factor Z and the coefficient of friction according to Rabinowicz [8].

5. ADHESIVE WEAR MODELS

Now in the special case of conical asperities the wear model can be derived as follows.
The minimum distance for separation of a cone is the diameter $d = 2r$ of the base of the cone. The number of such contacts per slide travel S is $S/2r$. The number of all true microcontact areas is n when A_r is the total true contact area. The probability that one contact will cause the formation of a wear particle is K.The volume of one wear particle is V_k. The total worn out volume is

$$V = K\frac{S}{2a}nV_k \quad , \quad n = \frac{A_r}{\pi a^2} \quad , \quad n = \frac{F_n}{3\sigma_0 \pi a^2} \rightarrow H = 3\sigma_0 \rightarrow V = K \cdot S \cdot V_k \frac{F_n}{2H\pi a^3} \tag{11}$$

In the case of a conical wear particle

$$V_k = \tfrac{1}{3}\pi a^2 h = \tfrac{1}{3}\pi a^2 a\tan\theta \quad , \quad H = 3\sigma_0$$

$$\frac{V}{S} = K\frac{F_n}{6\sigma_0\pi a^3}V_k = \tfrac{1}{3}K\tfrac{1}{2}\tan\theta\frac{F_n}{H} = Z\frac{F_n}{H} \qquad \rightarrow Z = \tfrac{1}{3}K\cdot\tfrac{1}{2}\tan\theta = c\cdot\mu \tag{12}$$

According to Rabinowicz [8] (Fig. 6.20, Fig.6.21) the wear coefficient Z depends on the friction coefficient for metals (m) and nonmetals (n),Table 1. Essentially, friction is not a temperature dependent quantity.

$$\dot{h}_k = pv\frac{Z_k}{H_k} \quad , \quad Z_k = A_k\mu^{B_k} \tag{13}$$

Table 1. Coefficients of the wear rate equations fitted to data from [8].

k = m:metals	$A_m = 3.7\cdot10^{-3}$	$B_m = 3.7$	$H_m = H$
k = n: nonmetal	$A_n = 26\cdot10^{-6}$	$B_n = 2.05$	$H_n = 0.2H$

6. WEAR AND FRICTION IN A DRY FRICTION BRAKING

In a case when lubrication is insufficient, as in start-ups of sliding bearings or at brakings, the wear may be great. A model is shown in Figure 7. A rotor is in normal operation supported by force field bearings in axial directions. To slow it down an axial retaining bearing is used. Let us consider a ring of radius r with normal force N acting on it.
An approximate solution is obtained as follows with numerical estimates

$$\ddot{q} = -K = -\frac{\mu N r}{J} \quad , \quad \rightarrow \dot{q}(t) = \dot{q}(0) - Kt, \quad \dot{q}(0) = \omega_0, \quad T = \frac{\omega_0}{K} \quad , \quad v = \dot{q}r \tag{14}$$

Substituting this into the equation of motion gives

$$\dot{h} = pv\frac{Z}{H} = (\omega_0 - Kt)rp\frac{Z}{H} \quad \rightarrow \quad h(t) = \left(\omega_0 t - \tfrac{1}{2}Kt^2\right)rp\frac{Z}{H} \tag{15}$$

At the end the braking at full stop the thickness of the worn out layer is

$$h(T) = \tfrac{1}{2}\frac{\omega_0^2}{K}rp\frac{Z}{H} = \frac{\tfrac{1}{2}J\omega_0^2}{A}\frac{A_k\mu^{B_k\text{-}1}}{H} \tag{16}$$

A feasibility study of various material pair alternatives may also be done.
Input data: $\mu = 0.35$, $J = 0.18$ kgm^2, ω_o=3351 rad/s $m = 85$kg, mean ring radius $r = 0.04$m gives calculated values : $K = 66$, stop time $T = 3551/66 = 50$s.
Final wear per one braking :With pressure $p = 0.61$ MPa the wear is $h(T) = 20\mu$m.
The allowed wear per one deceleration is $h = 100\mu$m. The calculated wear at room temperature for metal/metal pair is about $h = 20\mu$m, Figure 7.

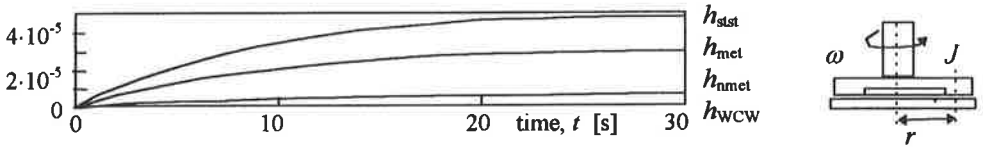


Figure 7: Weared out surface layer thicknesses h (m) for
h_{met} : for metals, h_{nmet} : for nonmetals, h_{wcwc} for WC/WC pairs,
h_{stst} for steel/steel pairs.

7. SCUFFING AND ACOUSTIC EMISSION

It has been proposed that the AE power should be proportional to the friction power.Lingard and NG [9] observe that in unlubricated friction the total AE count C would be more likely related to the total frictional work W_f than to wear parameters.They conclude that in dry friction AE rates vary systematically with loads, sliding speeds and material combinations, but for the experimental conditions used by them no consistent relationship between the AE and the absolute wear rate was found.

$$C = bW_f^{\,a} \quad , \quad W_f = \mu N \cdot s \quad , \quad N = pA \quad , \quad \mu p = \tau_{xz} \tag{17}$$

AE count rate is recorded often as pulse counts per second. Now it is

$$\dot{C} = abW_f^{\,a-1}\frac{d}{dt}(\mu N \cdot s) = ab(\mu N)^a \dot{s} \cdot s^{a-1} \tag{18}$$

here C is the total AE count, W_f is frictional work, μ is coefficient of friction, N is normal force , s is slip distance v is slip velocity. The factor b depends on the particular transducer and amplifier characteristics. The exponent a in the case of bronze on En24 steel varied between 1.1 to 1.68 with mean about 1.4.

Mirabile [10] has studied the connection of acoustic energy and dislocation mechanisms of plastic deformation and fracture in tensile testing. A test is used in which a steel ball is dropped form height h. The maximum amplitude in a volt channel is recorded

$$E_{AE} = E_{gr} \rightarrow 5 \cdot 10^{-4} \cdot V^2 \left[\frac{J}{Volt^2}\right]\left[Volt^2\right] = mgh[J] \tag{19}$$

In metals the plastic yielding takes place discontinuously in the grain size scale. Dislocation groups may stop at inclusions , interfaces and grain boundaries and forming a super dislocation of Burgers vector nb .This dislocation will have energy E [10]

$$nb = \frac{\sigma_{AE}}{2G}d \quad , \quad E \propto \frac{1}{4\pi}G(nb)^2 d \tag{20}$$

Here typical values are burgers vector $b = 2.5 \cdot 10^{-10}$ m, G is shear modulus , $G = 70000/2.6$ MPa, d is grain size, $d = 20 \cdot 10^{-6}$m , σ_{AE} is stress at which AE begins, now it may be the yield strength of the bronze σ_{AE} = 90MPa.

8. DESIGN FOR ZERO WEAR

A zero wear model predict wear life in cycles of sliding between two bodies. McGregor [2] has derived the following model. Wear cycles for parts 1 and 2 are M_1 ja M_2 with slip distance L_1 and L_2

$$M_2 = \frac{\pi d_1}{L}M_1 \quad \frac{L_1}{L_2}M_1 \quad \rightarrow M_2 L_2 = M_1 L_1 \tag{21}$$

The condition for achieving zero wear is that at the test stage life is $M \geq M_0 = 2000$.

Since the shear stresses under surface are decisive then the endurance condition may be expressed as

$$\tau_{max} \leq \tau_{M0} \quad , \quad \tau_{M0} = g_r \tau_y \quad , \quad \tau_y = 0.58R_e \tag{22}$$

Here R_e is yield strength and τ_y is yield strength in shear. The following relationship connects shear strength and wear life

$$M\tau_M^{\,9} = M_0 \tau_{M0}^{\,9} \rightarrow \tau_M = \left[\frac{M_0}{M}\right]^{\frac{1}{9}}\tau_{M0} = \left[\frac{2000}{M}\right]^{\frac{1}{9}} g_r \tau_y \tag{23}$$

where

g_r is wear and lubrication factor.Experimentally it has been observed to have discrete values

$g_r = 1$ in hydrodynamic lubrication

$g_r = 0.2$ in dry friction , $g_r = 0.2$ in boundary lubrication and for soft materials

$g_r = 0.54$ in boundary lubrication and for rather wear resistant materials.

In the case of two contact bodies this may be applied as follows:

$$\tau_{max} = 0.304 p_0 \quad , \quad p_0 = 0.591\left[\frac{FE}{dw}\right]^{\frac{1}{2}}$$

$$\frac{1}{d} = \frac{1}{D_1} + \frac{1}{D_2} \qquad (24)$$

$$L = 2b \quad , \quad b = 1.08\left[\frac{Fd}{Ew}\right]^{\frac{1}{2}}$$

where L is length of contact area , w is width
d is effective diameter in contact.

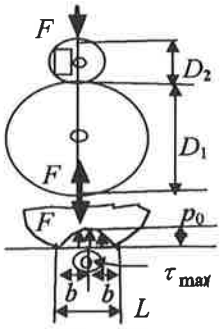


Figure 9.

9. PRINCIPLES OF MATERIALS DESIGNING OF MICROSTRUCTURE FOR OPTIMUM WEAR RESISTANCE

According to Hornbogen [11], the wear resistance is approximately a linear function of the relative area of a planar section of the microstructure covered by particles $f_A = A_p / A$. The specific wear rate can be expressed as follows

$$\dot{w} = Z\frac{p}{H} = \frac{1}{J} = \dot{w}_M + f_A\left(\dot{w}_p - \dot{w}_M\right) \quad , \quad f_A = \frac{A_p}{A} = \left(\frac{2r}{2P}\right)^2 = \left(\frac{r}{P}\right)^2 \quad ,2P = d + L \quad (25)$$

here $\dot{w}_p, \dot{w}_M$ are the specific wear rates for pure particles and for pure matrix, A_p is the area covered by particle of an area A.If the particles are cubical with side $2r$ then the wear rate is isotropic.Using an empirical model for wear resistance gives

$$\dot{w}^{-1} = a' + b'\left(\frac{F(f)}{d}\right)^{1/2} , F(f) = \frac{d}{L} = \frac{1}{0.82 f^{-1/3} - 1} , f = \frac{\frac{4}{3}\pi r^3}{(2P)^3} = \frac{\pi}{6}\left(\frac{r}{P}\right)^3 = \frac{\pi}{6}\left(\frac{d}{d+L}\right)^3 \quad (26)$$

where $d = 2r$ is diameter of hard particles, L is the surface to surface spacing and f is their volume fraction.Wear damage can be minimised using these models. Axen [12] suggests that the wear resistance can be maximized as follows

- By increasing matrix hardness
- By increasing sufficiently fracture toughness,
- By using the constraints on the dimensions of the microstructure: $L/D < 1$, $D/d < 1$ and $L < D < d$, where D is width of wear grooves [12].
- By using plate-like pearlitic microstructure which gives better wear resistance than globulitic particles obtained by quenching and tempering [13],[14].

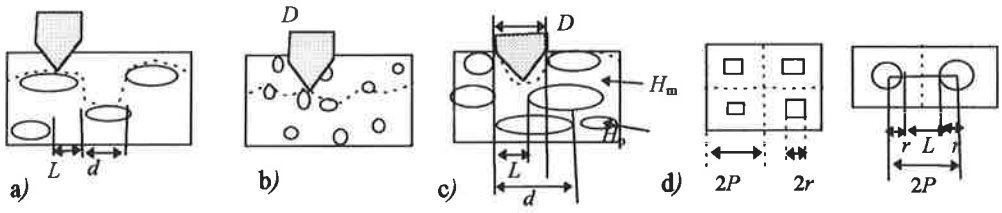


Figure 9. Wear models. Large wear with a) and b), c) an idea how wear possibly could be minimised with an optimal choice of the microstructural design variables.

10. DISCUSSION

In mechanical machines the contact between machine elements with or without lubrication is important. These kinematic joints can be optimally designed starting to design simultaneously the microstructure and the macrostructure. New virtual model simulation tools are proving to be cost effective.

11. CONCLUSIONS

The following conclusions may be drawn.

- In industry there is an increasing need to control the tribological behaviour of machine element contacts especially in sliding bearings.
- The simulations using LS-DYNA FEM virtual models of the microstructual mechanisms of dry friction give detailed and visual scenarios of the contact phenomena at the surface roughness level and of the large deformation process involved.
- Simulations can be advantageously used to predict and optimize sliding bearing behaviour on micro and macrolevels and also to explain the relationship between acoustic emission signal and phenomena and damage in sliding contacts.
- Experimental microscopic and metallographic analysis is needed to verify results.

Acknowledgements
The expert assistance of Mr. Reijo Hämäläinen in LS-DYNA simulations is gratefully acknowledged.

REFERENCES

[1] LS-DYNA version 950 , university licence version, Theoretical manual, Livermore Software Technology Corporation, Livermore.
[2] McGregor S. W., Handbuch der analytischen Verschleissberechnung. New York, Plenum Press, 1964.
[3] Stout K.J., Watson &G King T.G., The microgeometry of lubricated wear classification and modeling.
[4] Tsukizoe T. The effects of surface topography on wear.
[5] McClintock, F. A. & Argon A. S., Mechanical behavior of materials, Addison Wesley, 1966.
[6] Mikic, B.B., & Roca, R. T., International Journal of Heat and Mass Transfer, Vol.17, 1974, p.205.
[7] Dyson A, Institution of mechanical engineers.Proceedings, Vol.190,1976.
[8] Rabinowicz,E., *Friction and wear of materials*, second edition, John Wiley & Sons, New York, p.307, 1995.
[9] Lingard S. & NG K. K., An investigation of acoustic emission in sliding friction and wear of metals, Wear, 130 (1989) 367-379.
[10] Mirabile M., Acoustic emission energy and mechanisms of plastic deformation and fracture, Nondestructive testing , April 1975, pp. 77-85
[11] Hornbogen E, Metallurgical aspects of wear, *Foertschrittberichte der* VDI -Z Reihe 5 Nr.24.
[12] Axen, N.,*Abrasive wear of metal matrix composites*, PhD Thesis, Materials Science Division, Uppsala University 1993.
[13] Jacobson, S. & Hogmark, S., *Tribologi*, Liber Utbildning AB, 1996,p.260.
[14] Whitehouse D. J., Surface topography and quality and ist relevance to wear, pp.11-51.

A PATCH RECOVERY METHOD FOR 2D STOKES AND NAVIER-STOKES PROBLEMS

M. PERÄLÄ[1]
Helsinki University of Technology
Laboratory of Structural Mechanics
PL 2100, 02015 ESPOO

ABSTRACT

A patch recovery method for steady, incompressible 2D Stokes and Navier-Stokes flows, which uses local polynomial representations containing information from the field equations and the boundary conditions, is introduced. The method has been implemented to a finite element code, which uses quadratic, C^0-continuous velocities, discontinuous pressure and penalty function method for the incompressibility constraint. Numerical comparison of the present and the SPR method [1], is given using two model problems.

1. INTRODUCTION

Patch recovery methods [1]-[6] combined with so called Zienkiewich-Zhu error estimate [7] have proved out to be efficient tools for *a posteriori* error analysis of finite element results. In Ref. [5] the patch recovery method of Ref. [3] has been applied to the 2D Stokes flow. Generalization to Navier-Stokes flow was also discussed. The purpose of this paper is to study this subject in more detail. Both 2D Stokes and Navier-Stokes flows are considered. In addition to strain rate recovery also pressure recovery is dealt with. In addition to field equations also boundary conditions are included into the method.

The patch recovery method is based on a stream function/pressure formulation of the field equations and the boundary conditions. In this method, the stream function $\psi(x,y)$ and the pressure $p(x,y)$ are represented locally within each patch using a complete polynomial of degree q for the stream function and degree $q-2$ for the pressure, respectively. By demanding that the polynomial representations of the stream

[1] Address: Pinkelikatu 6 B 36, 90520 OULU.

function and the pressure satisfy approximately the differential equations (equations of motion) of the original boundary value problem (expressed in terms of the stream function and pressure) a set of constraint equations between the unknown parameters of these representations are found. With the help of the constraint equations based on the field equations part of the original parameters can be eliminated and a set of independent parameters established. New polynomial representations for the stream function and the pressure in terms of these independent parameters are finally formed.

Using the relations between the velocities and the stream function $u(x,y) \equiv \partial\psi/\partial y$ and $v(x,y) \equiv -\partial\psi/\partial x$, the velocities can be expressed in terms of the independent parameters as polynomials of degree $q-1$. The corresponding polynomial representations of degree $q-2$ for the strain rates can be obtained using the relations $d_x \equiv \partial u/\partial x$, $d_y \equiv \partial v/\partial y$ and $d_{xy} \equiv \partial u/\partial y + \partial v/\partial x$. The polynomial representations of the velocities, pressure and the strain rates contain thus "built-in" information from the field equations.

The unknown parameters are obtained by least squares fitting of: a) the representation of the velocities to the corresponding finite element nodal values and b) the representation of the pressure to the corresponding finite element pressure values at the Gauss points of the patch. In patches, whose assembly nodes are located on the boundary of the domain, information from the boundary conditions cause additional constraint equations between the unknown parameters. These constraint equations are included into the least squares fitting procedure using Lagrange multipliers. The final, C^0-continuous field for the strain rates and the pressure are calculated using the patchwise representations for them and averaging at the system nodes.

The patch recovery procedure has been implemented to a finite element code, which uses quadratic, C^0-continuous velocities and discontinuous pressure. Corresponding to isoparametric quadratic triangles and quadratic Lagrange quadrilaterals (for the velocities), respectively, the code uses constant and linear pressure. The discrete incompressibility condition is taken into account using the penalty function method. The final unknowns of the problem are thus the nodal velocities u_i and v_i.

2. EQUATIONS OF STOKES AND NAVIER-STOKES FLOWS

Equations of motion are in vector form as

$$\eta\nabla^2\mathbf{u} - \nabla p + \rho\mathbf{b} = \mathbf{0} \qquad \text{in} \quad \Omega \tag{1}$$

for Stokes and

$$\eta\nabla^2\mathbf{u} - \rho(\mathbf{u}\cdot\nabla)\mathbf{u} - \nabla p + \rho\mathbf{b} = \mathbf{0} \qquad \text{in} \quad \Omega \tag{2}$$

for Navier-Stokes flow, where $\mathbf{u}(x,y)$ is unknown velocity vector, $p(x,y)$ is unknown pressure, η is viscosity, ρ is density and $\mathbf{b}(x,y)$ is known volume force vector. Equation of continuity (incompressibility condition) is

$$\nabla\cdot\mathbf{u} = 0 \qquad \text{in} \quad \Omega. \tag{3}$$

By using the stream function $\psi(x,y)$, a set of three field equations of Stokes flow (equations (1) and (3)) can be written as a single field equation

$$R^{\mathrm{S}}(\psi) \equiv \eta\nabla^4\psi - \rho(\frac{\partial b_y}{\partial x} - \frac{\partial b_x}{\partial y}) = 0 \qquad \text{in} \quad \Omega. \tag{4}$$

The corresponding nonlinear field equation for Navier-Stokes flow is

$$R^{\mathrm{NS}}(\psi) \equiv \eta\nabla^4\psi + \rho\Big[\frac{\partial\psi}{\partial x}\frac{\partial}{\partial y}(\nabla^2\psi) - \frac{\partial\psi}{\partial y}\frac{\partial}{\partial x}(\nabla^2\psi)\Big] - \rho(\frac{\partial b_y}{\partial x} - \frac{\partial b_x}{\partial y}) = 0 \quad \text{in } \Omega. \tag{5}$$

In equations (4) and (5) $R(\psi)$ is used to denote the residual of the equation. Superscript S refers to Stokes and NS to Navier-Stokes flow.

3. LOCAL POLYNOMIAL REPRESENTATIONS WITH "BUILT-IN" FIELD EQUATIONS

We express the stream function $\psi(x,y)$ locally using a complete polynomial of degree q as

$$\tilde{\psi}(\lambda,\mu) = \sum_{i=0}^{q}\sum_{j=0}^{i}\lambda^{i-j}\mu^j\psi_{ij}, \tag{6}$$

where ψ_{ij} are unknown parameters, $\lambda = (x - x_0)/h$, $\mu = (y - y_0)/h$ are dimensionless coordinates, x_0 and y_0 are coordinates of the patch assembly node and h is characteristic length of the patch. We further express the known volume forces $b_x(x,y)$ and $b_y(x,y)$ using similar polynomials of degree $q-3$ with corresponding known coefficients b_{xij} and b_{yij}. Equation (6) can be expressed in matrix form as

$$\tilde{\psi}(\lambda,\mu) = \mathbf{Q}^q(\lambda,\mu)\Psi^q, \tag{7}$$

where

$$\mathbf{Q}^q = [\,1 \quad \lambda \quad \mu \quad \cdots \quad \lambda^q \quad \lambda^{q-1}\mu \quad \cdots \quad \lambda\mu^{q-1} \quad \mu^q\,] \tag{8}$$

and

$$\Psi^q = [\,\psi_{00} \quad \psi_{10} \quad \psi_{11} \quad \cdots \quad \psi_{q0} \quad \psi_{q1} \quad \cdots \quad \psi_{q(q-1)} \quad \psi_{qq}\,]^T. \tag{9}$$

Substituting polynomial representation of the stream function (6) and the corresponding representations of volume forces into the residuals (4) or (5) of the field equation results to

$$\tilde{R}^k(\lambda,\mu) = \sum_{i=0}^{q-4}\sum_{j=0}^{i}\lambda^{i-j}\mu^j R^k_{ij}, \qquad k = \mathrm{S},\mathrm{NS}. \tag{10}$$

By demanding that polynomial representation (10) of the residual of the field equation should vanish with all values of λ and μ results to equations $R^{\mathrm{S}}_{ij}(\Psi^q) = 0$ (linear) and $R^{\mathrm{NS}}_{ij}(\Psi^q) = 0$ (nonlinear). They are $(q-3)(q-2)/2$ constraint equations between

$(q+1)(q+2)/2$ unknown parameters ψ_{ij} and exist if $q \geq 4$. Constraint equations can be expressed in matrix form as

$$\mathbf{C}\Psi^q + \mathbf{b}^q = \mathbf{0}. \tag{11}$$

Note that matrix $\mathbf{C}$ depends on the unknown parameters Ψ^q in the case of Navier-Stokes flow. With the help of the $(q-3)(q-2)/2$ constraint equations the unknown parameters ψ_{ij} can be expressed in terms of $4q-2$ remaining ones, which can be considered as independent parameters a_i, $i = 0, \cdots, 4q-3$. The original parameters ψ_{ij} can thus be expressed in the terms of the independent parameters a_i (see Ref. [3]) in matrix form as

$$\Psi^q = \mathbf{S}^q\mathbf{a}^q + \mathbf{T}^q. \tag{12}$$

Substituting the result (12) into the original polynomial representation (7) gives

$$\tilde{\psi}(\lambda,\mu) = \mathbf{N}^q(\lambda,\mu)\mathbf{a}^q + \Psi_0^q(\lambda,\mu), \tag{13}$$

where

$$\mathbf{N}^q(\lambda,\mu) = \mathbf{Q}^q(\lambda,\mu)\mathbf{S}^q, \qquad \Psi_0^q(\lambda,\mu) = \mathbf{Q}^q(\lambda,\mu)\mathbf{T}^q. \tag{14}$$

Equation (13) is a local polynomial representation of degree q for the stream function, which contain information from the field equations "built-in".

We further express pressure $p(x,y)$ locally using a complete polynomial of degree $q-2$ as

$$\tilde{p}(\lambda,\mu) = p_0 + \sum_{i=1}^{q-2}\sum_{j=0}^{i} \lambda^{i-j}\mu^j p_{ij}, \tag{15}$$

where p_0 and p_{ij} are unknown parameters. Representation (15) can be written in matrix form as

$$\tilde{p}(\lambda,\mu) = p_0 + \mathbf{Q}_1^{q-2}(\lambda,\mu)\mathbf{P}_1^q, \tag{16}$$

where the matrix $\mathbf{Q}_1^{q-2}$ is similar to matrix $\mathbf{Q}^q$ except that the first term has been dropped out and

$$\mathbf{P}_1^q = \begin{bmatrix} p_{10} & p_{11} & \cdots & p_{(q-2)0} & p_{(q-2)1} & \cdots & p_{(q-2)(q-3)} & p_{(q-2)(q-2)} \end{bmatrix}^T. \tag{17}$$

Writing residuals of the equations of motion (1) or (2) (in terms of the stream function and the pressure) using a complete polynomial representation of degree $q-2$ and demanding that these residuals should vanish with all values of λ and μ we get relation between the parameters ψ_{ij} and p_{ij}. Relation can be written in matrix form as

$$\mathbf{P}_1^q = \mathbf{C}_\psi^q\Psi^q + \mathbf{b}_p^q. \tag{18}$$

In case of the Navier-Stokes flow, matrix $\mathbf{C}_\psi^q$ depends on the unknown parameters Ψ^q. By using the equations (12) and (18) , we can write the polynomial representation (16) of the pressure in terms of the independent parameters $\mathbf{a}^q$ as

$$\tilde{p}(\lambda,\mu) = p_0 + \mathbf{K}_p^q(\lambda,\mu)\mathbf{a}^q + T_p^q(\lambda,\mu), \tag{19}$$

where

$$\mathbf{K}_p^q(\lambda,\mu) = \mathbf{Q}_1^{q-2}(\lambda,\mu)\mathbf{C}_\psi^q\mathbf{S}^q, \qquad T_p^q(\lambda,\mu) = \mathbf{Q}_1^{q-2}(\lambda,\mu)(\mathbf{C}_\psi^q\mathbf{T}^q + \mathbf{b}_p^q). \tag{20}$$

Local polynomial representations for the velocities $u(x,y) \equiv \partial\psi/\partial y$ and $v(x,y) \equiv -\partial\psi/\partial x$ are obtained straightforwardly using equation (13) and they are in matrix form as

$$\tilde{u}(\lambda,\mu) = \mathbf{N}_u^q(\lambda,\mu)\mathbf{a}^q + u_0^q(\lambda,\mu) \tag{21}$$

and

$$\tilde{v}(\lambda,\mu) = \mathbf{N}_v^q(\lambda,\mu)\mathbf{a}^q + v_0^q(\lambda,\mu). \tag{22}$$

Correspondingly, we get local polynomial representations for the strain rates $d_x \equiv \partial u/\partial x$, $d_y \equiv \partial v/\partial y$ and $d_{xy} \equiv \partial u/\partial y + \partial v/\partial x$ by simply differentiating representations (21) and (22) . Representations can be written in matrix form as

$$\tilde{d}_x(\lambda,\mu) = \mathbf{N}_x^q(\lambda,\mu)\mathbf{a}^q + d_{x0}^q(\lambda,\mu), \tag{23}$$

$$\tilde{d}_y(\lambda,\mu) = \mathbf{N}_y^q(\lambda,\mu)\mathbf{a}^q + d_{y0}^q(\lambda,\mu) \tag{24}$$

and

$$\tilde{d}_{xy}(\lambda,\mu) = \mathbf{N}_{xy}^q(\lambda,\mu)\mathbf{a}^q + d_{xy0}^q(\lambda,\mu). \tag{25}$$

Equations (19) , (21) , (22) and (23) - (25) are the local polynomial representations for the pressure, velocities and strain rates, which contain "built-in" information from the field equations. The polynomial degrees of these representations are $q-2$, $q-1$ and $q-2$, respectively.

4. INFORMATION FROM THE BOUNDARY CONDITIONS

Dirichlet (velocity) boundary conditions and Neumann (traction) boundary conditions in terms of the stream function and pressure are

$$\begin{aligned} R_{Dx} &\equiv \frac{\partial\psi}{\partial y} - \bar{u} = 0 \qquad \text{on} \quad \Gamma_u, \\ R_{Dy} &\equiv -\frac{\partial\psi}{\partial x} - \bar{v} = 0 \qquad \text{on} \quad \Gamma_v \end{aligned} \tag{26}$$

and

$$\begin{aligned} R_{Nx} &\equiv -\eta n_y\frac{\partial^2\psi}{\partial x^2} + 2\eta n_x\frac{\partial^2\psi}{\partial x\partial y} + \eta n_y\frac{\partial^2\psi}{\partial y^2} - n_x p - \bar{t}_x = 0 \qquad \text{on} \quad \Gamma_{tx}, \\ R_{Ny} &\equiv -\eta n_x\frac{\partial^2\psi}{\partial x^2} - 2\eta n_y\frac{\partial^2\psi}{\partial x\partial y} + \eta n_x\frac{\partial^2\psi}{\partial y^2} - n_y p - \bar{t}_y = 0 \qquad \text{on} \quad \Gamma_{ty}, \end{aligned} \tag{27}$$

where $\bar{u}(s)$ and $\bar{v}(s)$ are the prescribed velocities and $\bar{t}_x(s)$ and $\bar{t}_t(s)$ are the prescribed tractions, respectively. Our aim is to form constraint equations based on

boundary conditions (26) and (27) . We express known functions $\bar{u}(s), \cdots, \bar{t}_y(s)$ on the boundary using a complete polynomial of degree q (see Ref. [3]). Substituting polynomial representations of the stream function, pressure and the known functions into the residuals (26) and (27) of the boundary conditions results to

$$\tilde{\mathbf{R}}_D(\sigma) = \sum_{i=0}^{q-1} \sigma^i \mathbf{R}_{Di}, \qquad \tilde{\mathbf{R}}_N(\sigma) = \sum_{i=0}^{q-2} \sigma^i \mathbf{R}_{Ni}, \tag{28}$$

where $\sigma = (s - s_0)/h$ is dimensionless coordinate and s_0 is s-coordinate of the patch assembly node.

By demanding that the polynomial representations (28) should vanish with all values of σ results in Dirichlet case to q matrix equations $\mathbf{R}_{Di}(\Psi^q) = \mathbf{0}$, $i = 0, \cdots, q-1$ and in Neumann case to $q-1$ matrix equations $\mathbf{R}_{Ni}(\Psi^q) = \mathbf{0}$, $i = 0, \cdots, q-2$. These equations are additional linear constraint equations between the unknown parameters Ψ^q (in addition to equations (11)), which should hold, if the patch assembly node is on the boundary of the domain. Constraint equations can be written in matrix form as

$$\mathbf{C}_D^q \Psi^q + \mathbf{d}_D^q = \mathbf{0} \tag{29}$$

and

$$\mathbf{C}_N^q \Psi^q + \mathbf{n}_0 p_0 + \mathbf{n}_1^q \mathbf{p}_1^q + \mathbf{d}_N^q = \mathbf{0}, \tag{30}$$

where

$$\mathbf{p}_1^q = [p_1 \quad p_2 \quad \cdots \quad p_{q-2}]^T. \tag{31}$$

Pressure terms p_k, $k = 1, \cdots, q-2$ can be written in terms of coefficients p_{ij}, $i = 1, \cdots, q-2$; $j = 0, \cdots, i$. Relation can be written in matrix form as

$$\mathbf{p}_1^q = \mathbf{C}_p^q \mathbf{P}_1^q, \tag{32}$$

where matrix $\mathbf{C}_p^q$ depends on the terms λ_i and μ_i, which further depend on the boundary curve (details are given in Ref. [3], Appendix A.3). With the help of equations (18) and (32) , constraint equation (30) can be written as

$$\mathbf{C}_N^{*q} \Psi^q + \mathbf{n}_0 p_0 + \mathbf{d}_N^{*q} = \mathbf{0}, \tag{33}$$

where

$$\mathbf{d}_N^{*q} = \mathbf{n}_1^q \mathbf{C}_p^q \mathbf{b}_p^q + \mathbf{d}_N^q \tag{34}$$

and

$$\mathbf{C}_N^{*q} = \mathbf{C}_N^q + \mathbf{n}_1^q \mathbf{C}_p^q \mathbf{C}_\psi^q. \tag{35}$$

By using the equations (12) we can write the constraint equation based on Dirichlet boundary conditions (29) and Neumann boundary conditions (33) in terms of the independent parameters $\mathbf{a}^q$ such as

$$\mathbf{G}^q \mathbf{a}^q = \mathbf{H}^q \tag{36}$$

and

$$\mathbf{G}^q\mathbf{a}^q + \mathbf{n}_0 p_0 = \mathbf{H}^q, \tag{37}$$

where

$$\mathbf{G}^q = \begin{cases} \mathbf{C}_D^q\mathbf{S}^q & \text{(Dirichlet boundary)} \\ \mathbf{C}_N^{*q}\mathbf{S}^q & \text{(Neumann boundary)} \end{cases} \tag{38}$$

and

$$\mathbf{H}^q = \begin{cases} -\mathbf{d}_D^q - \mathbf{C}_D^q\mathbf{T}^q & \text{(Dirichlet boundary)} \\ -\mathbf{d}_N^{*q} - \mathbf{C}_N^{*q}\mathbf{T}^q & \text{(Neumann boundary).} \end{cases} \tag{39}$$

5. LEAST SQUARES FITTING IN A RECOVERY PATCH

The unknown parameters $\mathbf{a}^q$ and p_0 are obtained by least squares fitting of: a) the polynomial representations $\tilde{u}(\lambda,\mu)$ and $\tilde{v}(\lambda,\mu)$ to the corresponding nodal values $\hat{u}_i$ and $\hat{v}_i$ of the original finite element solution at the nodes of the patch and b) the polynomial represention $\tilde{p}(\lambda,\mu)$ to the corresponding finite element solution $\hat{p}_k$ at the Gauss points of the patch. The corresponding least squares functions are thus

$$\Pi = \sum_i \Big\{ \big[\tilde{u}(x_i,y_i) - \hat{u}_i\big]^2 + \big[\tilde{v}(x_i,y_i) - \hat{v}_i\big]^2 \Big\} \tag{40}$$

and

$$\Pi_p = \sum_k \big[\tilde{p}(x_k,y_k) - \hat{p}_k\big]^2. \tag{41}$$

If patch assembly point (x_0,y_0) is on the boundary of the domain, the information from the boundary conditions is taken into account by using Lagrange multipliers method in connection with the least squares function. The modified least squares function is thus (Neumann boundary)

$$\begin{aligned} \Pi^* = &\sum_i \Big\{ \big[\tilde{u}(x_i,y_i) - \hat{u}_i\big]^2 + \big[\tilde{v}(x_i,y_i) - \hat{v}_i\big]^2 \Big\} + \\ &+ \sum_k h^2 \big[\tilde{p}(x_k,y_k) - \hat{p}_k\big]^2 + \boldsymbol{\lambda}^T(\mathbf{G}^q\mathbf{a}^q + \mathbf{n}_0 p_0 - \mathbf{H}^q), \end{aligned} \tag{42}$$

where $\boldsymbol{\lambda}$ is vector of Lagrange multipliers. The least squares fitting of velocities and pressure are now coupled, which is due to the constraint equation (37) . In the case of Dirichlet boundary the modified least squares function does not contain the pressure. The pressure fitting is made using the equation (41) .

In case of the Navier-Stokes flow the least squares fitting procedure is nonlinear and must be solved iteratively.

6. NUMERICAL EXAMPLES

6.1 Example problems and typical grids

Based on the well known analogy between incompressible plane elasticity and Stokes flow, infinite plate with a circular hole under unidirectional tension (Fig. 1) was considered as a test problem with known analytical solution. This problem is referred as "circular hole problem" in the following. Figure 2 shows typical finite element grids ($h/a = 0.5$) used in experimental convergence study of this problem.

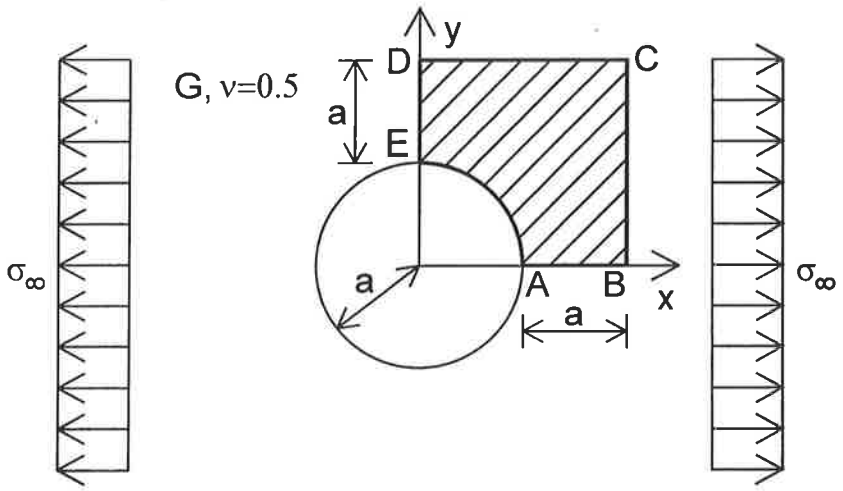


Fig. 1: Circular hole problem.

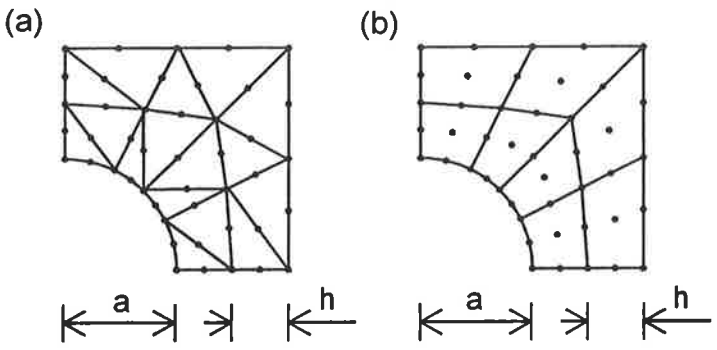


Fig. 2: Typical grids of (a) quadratic triangles and (b) quadratic Lagrange quadrilaterals.

Second problem, which is referred as "converging channel problem", in which fluid flows in the converging channel towards the origin (Fig. 3a) was considered as a test problem for Navier-Stokes flow. The Reynolds number 20 was used for the flow. Figure 3b shows a finite element grid ($h/a = 0.5$) used in experimental convergence study.

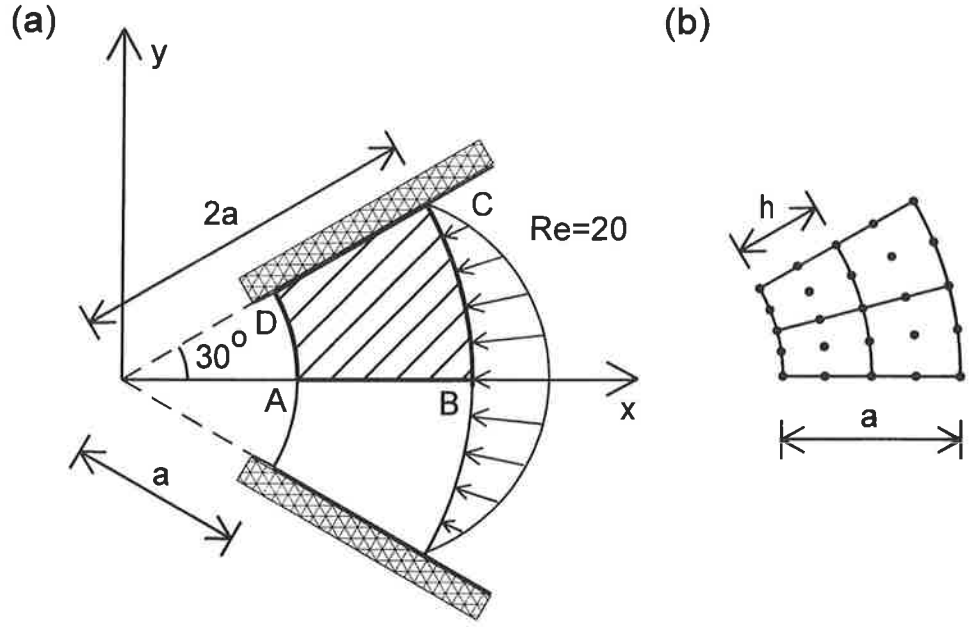


Fig. 3: (a) Converging channel problem and (b) typical grid of quadratic Lagrange quadrilaterals.

6.2 Numerical results

Figs. 4-6 present experimental convergence study of the relative error in energy (η_E) and relative error of L_2-norm of pressure (η_{L_2}) of the two example problems. Figs. 7 and 8 present corresponding effectivity indices (θ). Two different element types: quadratic triangles and quadratic Lagrange quadrilaterals were used in the analysis of the "circular hole problem". The "converging channel problem" was analyzed with quadratic Lagrange quadrilaterals. The polynomial degrees $q = 4$ and $q = 5$ were used in connection with quadratic triangles and quadratic quadrilaterals, respectively. Each of the figures 4-6 show error of the smoothed solution obtained with the present method with (BUIB) and without (BUI) information from the boundary conditions. Error of the smoothed solution obtained with the superconvergent patch recovery (SPR) method of reference [1] and error of the original finite element solution (FE) are also shown for comparison. SPR method does not recover the pressure and thus comparison between the present and SPR method is not possible. Thus the results of pressure recovery of the present method is compared to the results of the original

finite element (FE) results. Figs. 4-6 show that the present method gives clearly better smoothed solution than the SPR method for the error in energy. The present method gives higher convergence rate for the smoothed solution than SPR in case of the quadratic triangles.

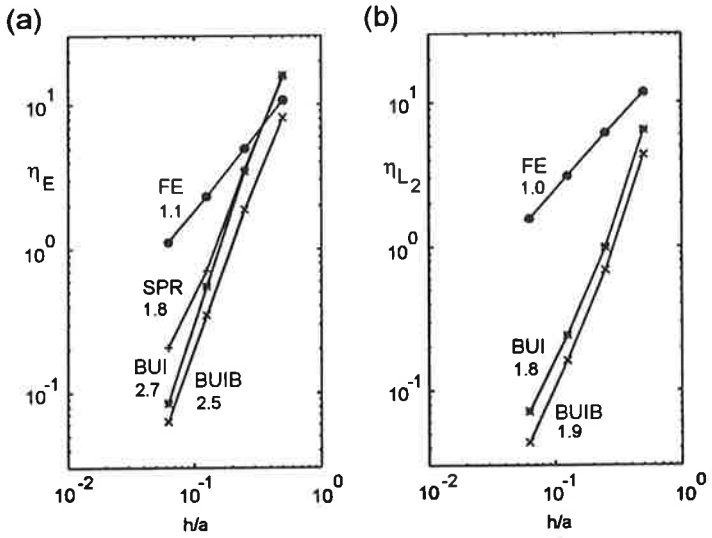


Fig. 4: Circular hole problem. Quadratic triangles: (a) error in energy (η_E) and (b) error of L_2-norm of pressure (η_{L_2}).

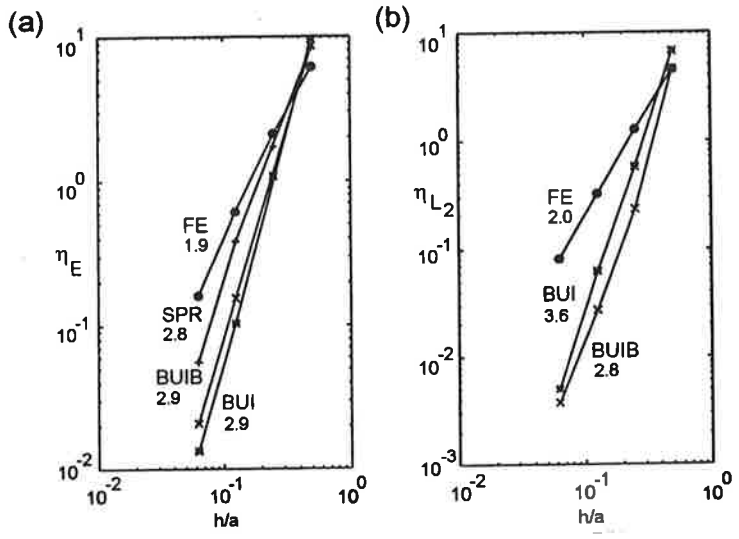


Fig. 5: Circular hole problem. Quadratic quadrilaterals: (a) error in energy (η_E) and (b) error of L_2-norm of pressure (η_{L_2}).

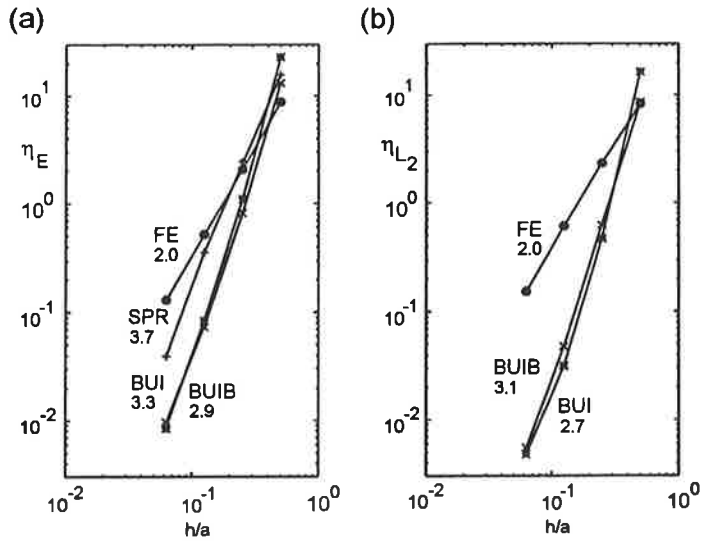


Fig. 6: Converging channel problem. Quadratic quadrilaterals: (a) error in energy (η_E) and (b) error of L_2-norm of pressure (η_{L_2}).

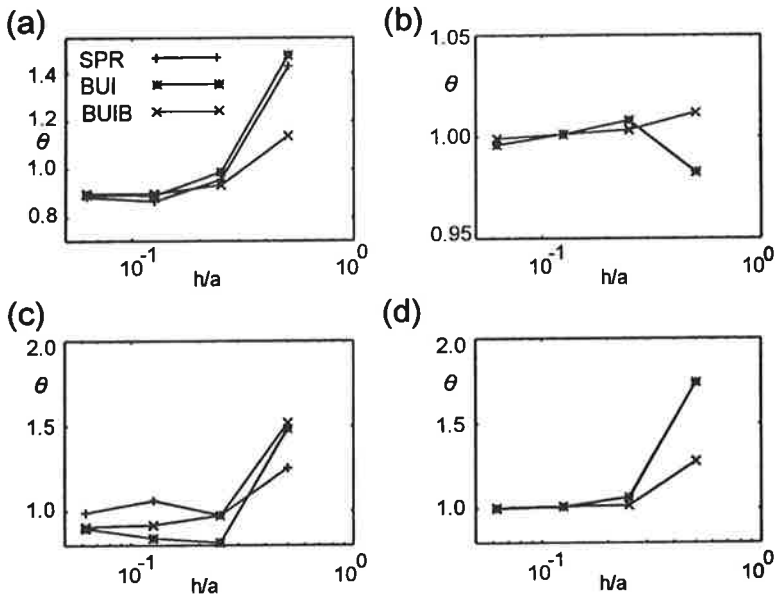


Fig. 7: Effectivity indices of the circular hole problem. Quadratic triangles: (a) error in energy and (b) error of L_2-norm of pressure. Quadratic quadrilaterals: (c) error in energy and (d) error of L_2-norm of pressure .

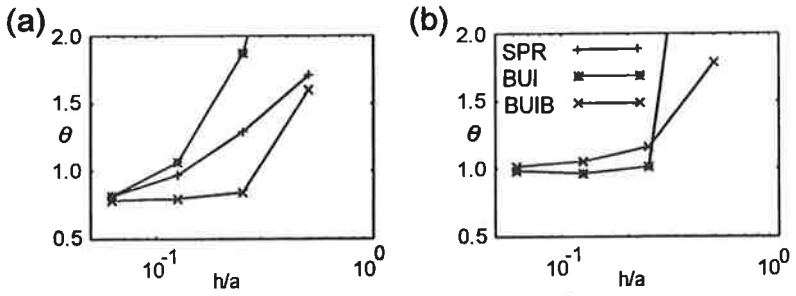


Fig. 8: Effectivity indices of the converging channel problem. Quadratic quadrilaterals: (a) error in energy and (b) error of L_2-norm of pressure.

Figs. 4-8 show that the presented pressure recovery works well, the effectivity indices are very near unity.

7. CONCLUSIONS

A patch recovery method for 2D Stokes and Navier-Stokes flows has been presented. The method covers both the strain rate and the pressure recovery. Numerical results obtained show, that the method works well and seems to give results, which are better than those of the SPR method of reference [1].

REFERENCES

[1] O.C. Zienkiewich and J.Z. Zhu: *The superconvergent patch recovery and a posteriori error estimates. Part 1: The recovery technique*, Int. J. Numer. Meth. Engng. **33** (1992), 1331-1364.

[2] J. Aalto: *Built-in field equations for recovery procedures*, Computers and Structures, **64** (1997), 157-176.

[3] J. Aalto and M. Åman: *Polynomial representations for the patch recovery procedures*, Computers and Structures, **73** (1999), 119-146.

[4] J. Aalto and M. Perälä: *Built-in field equations for patch recovery procedures using weihted residuals*, Computers and Structures, **73** (1999), 91-118.

[5] J. Aalto and R. Hannila: *A patch recovery method for 2D Stokes flow*, Proceedings fo the Eleventh Nordic Seminar on Computational Mechanics (NSCM XI), Stockholm, (1998).

[6] N.-E. Wiberg, F. Abdulwahab and S. Ziukas: *Enhanced superconvergent patch recovery incorporating equilibrium and boundary conditions*, Int. J. Numer. Meth. Engng. **37** (1994), 3417-3440.

[7] O.C. Zienkiewich and J.Z. Zhu: *A simple error estimator and adaptive procedure for practical engineering analysis*, Int. J. Numer. Meth. Engng. **24** (1987), 337-357.

STAATTINEN TASAPAINOTUS SIJOITTELUONGELMANA

Markus Aho
Tampereen teknillinen korkeakoulu
Teknillinen mekaniikka
PL 589
33101 Tampere

TIIVISTELMÄ

Tässä artikkelissa tarkastellaan eripainoisten massojen sijoittelua osaksi jäykkää systeemiä sekä optimaaliseen sijoitteluun liittyvän optimointiongelman muodostamista painovoimien aiheuttamaa ulkoista momenttia minimoitaessa. Muodostettu diskreetti sijoitteluongelma ratkaistaan geneettisellä algoritmilla, joka hyödyntää sijoittelutehtävän erityisrakenteen. Lasketut testiesimerkit osoittavat, että kyseisellä algoritmilla on käyttöä, kun kohtuullisen hyvän ratkaisun löytyminen on tärkeää eikä tarkkaa ratkaisua ole perinteisillä algoritmeilla tehokkaasti saavutettavissa. Tehtävälle saadaan ratkaisu aina myös suurilla muuttujien lukumäärillä.

1. JOHDANTO

1.1 Sijoitteluongelmat

Monissa sovellutuksissa joudutaan sijoittelemaan erilaisia kappaleita ja massoja. Toisinaan massojen sijoittelujärjestyksestä aiheutuvalla voimajakaumalla on konstruktion toiminnan kannalta suuri merkitys. Tällöin ollaan kiinnostuneita erityisesti massojen optimaalisesta järjestelystä. Syntyvät optimointiongelmat ovat sijoitteluongelmia (*Assignment problems*, katso [1], [2] ja [3]), jotka voidaan esittää seuraavassa muodossa

$$\min \; F(a_{11}, a_{12}, \ldots, a_{1n}, a_{21}, a_{22}, \ldots, a_{2n}, \ldots, a_{n1}, a_{n2}, \ldots, a_{nn}) \tag{1}$$

$$\text{s.e.} \quad \sum_{j=1}^{n} a_{ij} = 1, \qquad i = 1, 2, \ldots, n \tag{2}$$

$$\sum_{i=1}^{n} a_{ij} = 1, \qquad j = 1, 2, \ldots, n \tag{3}$$

$$a_{ij} \in \{0, 1\} \tag{4}$$

Kohdefunktio (1) muodostetaan tehtäväkohtaisesti ja se on muuttujien a_{ij} funktio. Muuttujia tehtävässä on yhteensä n^2 kappaletta, kun n on sijoiteltavien objektien ja sijoittelupaikkojen lukumäärä. Ensimmäinen rajoitusyhtälöryhmä (2) takaa, että jokainen i

sijoitetaan täsmälleen yhteen paikkaan summaamalla yli kaikkien mahdollisten paikkojen. Toinen rajoitusyhtälöryhmä (3) varmistaa, että kuhunkin paikkaan j voidaan sijoittaa täsmälleen kerran. Kyseessä on kokonaislukuoptimoinnin tehtävä, jossa kukin muuttuja a_{ij} saa vain binääriarvoja muuttujarajoitusten (4) mukaisesti.

1.2 Sijoitteluongelman yleinen ratkaiseminen

Edellä esitetyn sijoitteluongelman matemaattinen ratkaiseminen riippuu merkittävästi sovellettavan kohdefunktion $F(a_{ij})$ luonteesta. Lineaarisen positiivis- ja kokonaislukukertoimisen ($c_{ij} \in Z^+ \quad \forall i, j$) kohdefunktion

$$F(a_{ij}) = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} a_{ij} \qquad (5)$$

tapauksessa löytyy useita erittäin tehokkaita erityisalgoritmeja (esimerkiksi *Hungarian method*, katso [1], [3] ja [4]). Sen sijaan yhtä tehokkaita erityisalgoritmeja ei löydy kohdefunktion kertoimien ollessa erimerkkisiä ja tuotaessa tehtävään lisää rajoitusehtoja eikä epälineaarisen kohdefunktion tapauksessa. Tarkan optimiratkaisun hakemiseen on tällöin sovellettava yleisalgoritmeja, kuten leikkaustasomenetelmiä (*Cutting-plane algorithms*) ja luettelointimenetelmiä (*Enumerative methods*) tai niiden yhdistelmiä. Tärkein luettelointimenetelmä on *Branch-and-bound* –menetelmä. Kyseisiä menetelmiä on esitelty tarkemmin lähteissä [1], [2], [3] ja [5]. Nämä algoritmit ovat yleensä paljon suoraviivaista läpikäyntiä tehokkaampia, mutta silti kokonaislukutehtävien tarkka ratkaiseminen voi vaatia runsaasti resursseja. Tavallista kokonaislukutehtäville onkin tarvittavien resurssien eksponentiaalinen kasvu tehtävän muuttujien lukumäärän kasvaessa. Luonteenomaista sijoitteluongelmille on, että niille löytyy useita paikallisia minimejä.

Edellä lueteltujen tarkkojen menetelmien lisäksi voidaan optimointitehtävälle hakea lähellä optimiratkaisua oleva ratkaisu stokastisilla menetelmillä. Ne soveltuvat tilanteisiin, joissa kohtuullisen hyvän ratkaisun löytyminen on tärkeää, eikä tarkkaa ratkaisua ole perinteisillä algoritmeilla tehokkaasti saavutettavissa [4]. Kaksi merkittävintä todennäköisyyspäättelyyn perustuvaa menetelmää ovat geneettinen algoritmi (*Genetic algorithm* eli GA) ja jäähdytysmenetelmä (*Simulated annealing* eli SA), joista ensimmäistä sovelletaan tässä esityksessä sijoitteluongelmien ratkaisemiseen. Jäähdytysmenetelmiä on käsitelty esimerkiksi lähteissä [1], [2] ja [6].

1.3 Sijoitteluongelman ratkaiseminen geneettisellä algoritmilla

Geneettiset algoritmit ovat luonnon evoluutiomallia mukailevia, heuristisia optimointimenetelmiä. Ne eivät tarvitse herkkyysanalyysiä vaan yksilöitä eli ratkaisuja verrataan keskenään elinkelpoisuusarvon (*fitness-value*) avulla, joka on tässä esityksessä suoraan kohdefunktion arvo käyvillä ratkaisuilla. Epäkäypiä ratkaisuja sakotetaan tavalla tai toisella rajoitusyhtälöiden rikkomisesta. Ongelman muuttujat koodataan sopiviin

esitysmuotoihin joko binäärilukuja tai muuttujien kokonaislukuarvoja käyttäen. Aluksi muodostetaan lähtöpopulaatio, jota muokataan edelleen algoritmin edetessä sukupolvesta seuraavaan kolmen pääoperaation avulla, jotka ovat valinta (*reproduction*), risteytys (*crossover*) ja mutaatio (*mutation*). Nämä operaatiot voidaan toteuttaa useilla eri tavoilla, joita on esitelty, kuten geneettistä algoritmia yleensäkin, tarkemmin lähteissä [7] ja [8].

Tässä esityksessä 10% edellisen sukupolven parhaista yksilöistä (*elite solutions*) valitaan suoraan seuraavaan sukupolveen ja loput yksilöt muodostetaan risteytyksellä sekä mutaation seurauksena. Mutaation kautta muodostettavaksi yksilömääräksi on valittu noin 1% populaation koosta. Suunnitteluvektori koodataan listana, jossa on alkioita sijoiteltavien massojen lukumäärä eli n kappaletta. Kukin listan alkion i (i = 1,...,n) arvo esittää suoraan paikkaa j (j = 1,...,n), johon kyseinen massa i on sijoitettu. Lisäksi lista ei voi sisältää kahta yhtä suurta alkiota. Esitetyssä koodaustavassa on se merkittävä etu, että rajoitusyhtälöt (2), (3) ja (4) tulevat kaikki huomioitua, jolloin kukin generoituva ratkaisu on käypä. Tällöin erityisesti risteytysoperaatiolta vaaditaan, että risteytettäessä edellisen sukupolven käypiä ratkaisuja tulee myös syntyvän ratkaisun olla käypä.

Risteytysoperaation toteutus on lähimpänä kirjallisuudesta löytyvää PMX-operaatiota (*partially matched crossover*), jossa kahdesta edellisen sukupolven yksilöstä eli vanhemmista muodostetaan uudet yksilöt eli jälkeläiset vanhempien satunnaisesti valitut osat vaihtamalla [7]. Tässä risteytys toteutetaan generoimalla satunnainen "risteytyskoodi" eli binäärilukulista, joka on muuttujien lukumäärän n pituinen. Satunnaisesti valituista kahdesta vanhemmasta P_1 ja P_2 muodostetaan jälkeläiset O_1 ja O_2 risteytyskoodin mukaan. Jälkeläisiksi kopioidaan aluksi vanhemmat ja sovelletaan sitten risteytyskoodia. Kun risteytyskoodin alkio i on 1, poimitaan vanhempien alkioiden i muodostama lukupari ja vaihdetaan sitä vastaavat luvut keskenään kummassakin jälkeläisessä. Risteytyskoodin alkion arvolla 0 ei vaihdeta mitään. Näin käydään läpi jokainen risteytyskoodin alkio. Jälkeläisistä valitaan seuraavaan sukupolveen vain elinkelpoisuusarvoltaan parempi.

1.3.1 Risteytysesimerkki

Olkoon n = 6 ja edellisen sukupolven risteytettävät yksilöt

$$P_1 = (1\ 2\ 3\ 4\ 5\ 6) \text{ ja } P_2 = (2\ 1\ 4\ 5\ 6\ 3).$$

Ensimmäinen yksilö edustaa ratkaisua $x_{11} = x_{22} = x_{33} = x_{44} = x_{55} = x_{66} = 1$, muiden muuttujien ollessa nollia ja jälkimmäinen yksilö edustaa ratkaisua $x_{12} = x_{21} = x_{34} = x_{45} = x_{56} = x_{63} = 1$, muiden muuttujien ollessa nollia. Olkoon sitten generoitu risteytyskoodi (0 1 0 0 0 0). Tällöin risteytyvät jälkeläiset ovat

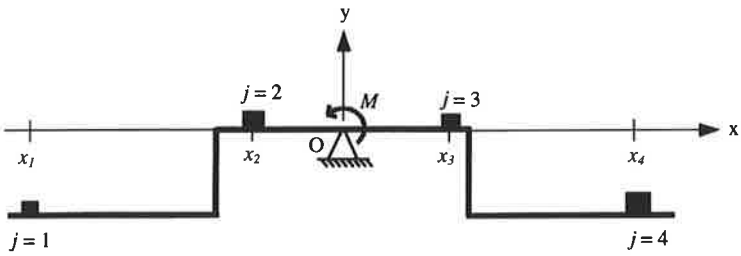
$$O_1 = (2\ 1\ 3\ 4\ 5\ 6) \text{ ja } O_2 = (1\ 2\ 4\ 5\ 6\ 3).$$

Edelleen, jos risteytyskoodi on esimerkiksi (0 1 1 0 1 0), ovat jälkeläiset

$$O_1 = (2\ 1\ 4\ 3\ 6\ 5) \text{ ja } O_2 = (1\ 2\ 3\ 6\ 5\ 4).$$

2. TUTKITTAVAT RAKENTEET JA NIIDEN LASKENTAMALLIT

2.1 Yksiulotteinen tasapainottavan momentin minimointi



Kuva 1. Jäykkä ja symmetrinen niveltuettu palkkirakenne

Kuvan 1 jäykkä ja symmetrinen niveltuettu palkkirakenne on esimerkki staattisesta tasapainotuksesta yhdessä tasossa. Neljä erisuurta massaa m_i (i = 1,...,4) tulee kukin sijoittaa yhteen neljästä mahdollisesta paikasta j (j = 1,...,4) palkilla siten, että tasapainottavan ulkoisen momentin M itseisarvo $|M|$ momenttipisteen O suhteen on mahdollisimman pieni. Sijoittelupaikkojen j koordinaatit x-akselin suunnassa momenttipisteestä ovat x_j. Kuvattu sijoitteluongelma optimointitehtävänä on muotoa

$$\min \quad \left| M(a_{ij}) \right| = g \left| \sum_{i=1}^{n} \sum_{j=1}^{n} m_i x_j a_{ij} \right| \tag{6}$$

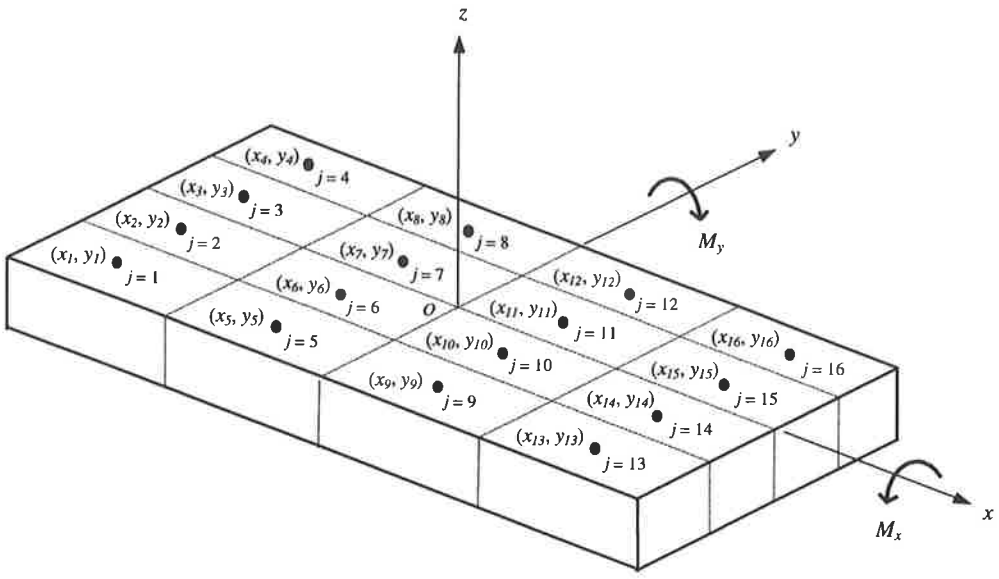
$$\text{s.e.} \quad \sum_{j=1}^{n} a_{ij} = 1, \qquad i = 1,2,\ldots,n \tag{7}$$

$$\sum_{i=1}^{n} a_{ij} = 1, \qquad j = 1,2,\ldots,n \tag{8}$$

$$a_{ij} \in \{0, 1\} \tag{9}$$

Jos vapaa putoamiskiihtyvyys g jätetään pois kohdefunktiosta (6), ei optimiratkaisu muutu. Toisin kuin tavallisesti kohdefunktion kertoimet $c_{ij} = m_i x_j$ saavat sekä positiivisia että negatiivisia arvoja.

2.2 Momenttien tasapainotus kahden akselin suhteen



Kuva 2. Osastoidun säiliölaivan yksinkertaistettu malli

Staattista tasapainotusta kahdessa tasossa tarkastellaan kuvan 2 yksinkertaistetun osastoidun säiliölaivan (*combination carrier*) mallin avulla. Laiva on ositettu yhtä suuriin osatilavuuksiin V, joiden tilavuuskeskiöiden koordinaatit ovat x_j ja y_j (j = 1,...,16) tasossa xy. Annetut eritiheyksiset ρ_i (i = 1,...,16) kemikaalit tulee kukin sijoittaa yhteen kustakin mahdollisesta osatilasta siten, että kohdefunktio minimoituu. Kohdefunktiona on maksimi kemikaalien massavoimista aiheutuvien, akseleiden x ja y suhteen laskettujen momenttien M_x ja M_y painotetuista itseisarvoista. Sijoitteluongelma matemaattisena optimointitehtävänä on siten muotoa

$$\min \quad \max\left\{ \omega_1 \left| M_x(a_{ij}) \right|, \ \omega_2 \left| M_y(a_{ij}) \right| \right\} \tag{10}$$

$$\text{s.e.} \quad \sum_{j=1}^{n} a_{ij} = 1, \qquad i = 1,2,\ldots,n \tag{11}$$

$$\sum_{i=1}^{n} a_{ij} = 1, \qquad j = 1,2,\ldots,n \tag{12}$$

$$a_{ij} \in \{0, 1\} \tag{13}$$

jossa esiintyvät momenttien lausekkeet ovat

$$M_x(a_{ij}) = g\left(\sum_{i=1}^{n}\sum_{j=1}^{n} m_i x_j a_{ij}\right) = gV\left(\sum_{i=1}^{n}\sum_{j=1}^{n} \rho_i x_j a_{ij}\right) \tag{14}$$

$$M_y(a_{ij}) = g\left(\sum_{i=1}^{n}\sum_{j=1}^{n} m_i y_j a_{ij}\right) = gV\left(\sum_{i=1}^{n}\sum_{j=1}^{n} \rho_i y_j a_{ij}\right) \tag{15}$$

ja kertoimet ω_1 ja ω_2 ovat edellisiä lausekkeita kohdefunktiossa painottavia kertoimia.

Optimiratkaisu ei muutu, jos vapaa putoamiskiihtyvyys g ja vakio V jätetään pois molemmista kohdefunktion lausekkeista (14) ja (15). Lausekkeissa esiintyvät kertoimet $c_{xij} = \rho_i x_j$ ja $c_{yij} = \rho_i y_j$ saavat sekä positiivisia että negatiivisia arvoja. Esitetyn optimointitehtävän ratkaisuna saadaan yksi Pareto-optimi kullakin painokertoimien ω_1 ja ω_2 yhdistelmällä taustalla olevalle kaksitavoitteiselle tehtävälle. Kohdefunktioina kaksitavoitteisessa tehtävässä ovat molemmat momentit eli juuri lausekkeet (14) ja (15). Monitavoitteisia optimointitehtäviä ja niiden ratkaisemista eri tavoin on käsitelty yksityiskohtaisesti lähteessä [9].

3. TULOKSET JA NIIDEN TARKASTELU

Laskenta-algoritmi implementoitiin Matlab-ohjelman avulla. Populaatiokoolle ja sukupolvien lukumäärälle muodostettiin empiiriset lausekkeet siten, ettei ratkaisuun tarvittava laskentatyö kasva liian voimakkaasti muuttujien n lukumäärän mukana. Populaatiokoko laskettiin lausekkeesta

$$N = 10 + \text{round}(14\sqrt{n}) \tag{16}$$

ja sukupolvien lukumäärä vastaavasti lausekkeesta

$$n_{gen} = \dim\left[20 + \text{round}(3\sqrt{n})\right] \tag{17}$$

joissa n on muuttujien lukumäärä, dim on kohdefunktion sisältämien osafunktioiden lukumäärä ja round on operaattori, joka pyöristää argumenttinsa lähimpään kokonaislukuun. Esitetylle yksiulotteiselle tehtävälle dim = 1 ja kaksiulotteiselle tehtävälle dim = 2.

Aluksi yksiulotteinen tehtävä ratkaistiin pienellä muuttujamäärällä n = 4. Laskennassa käytetyt tiedot sekä saavutettu minimiratkaisu ja minimiarvo on esitetty taulukossa 1. Data on valittu siten, että tehtävällä on vain yksi minimiratkaisu. Optimitulos on ilmoitettu 10:n ajon keskiarvona. Koska tehtävä on pieni, antaa käytetty algoritmi yleensä aina tarkan optimiratkaisun. Optimiratkaisu on saavutettavissa helposti näin pienelle tehtävälle myös

permutoimalla ja laskemalla läpi kaikki vaihtoehdot. Erilaisten ratkaisuvaihtoehtojen lukumäärä yleisesti on $n!$ kappaletta eli 24 tapaukselle $n = 4$.

Taulukko 1. Sovellettu data sekä laskettu minimiratkaisu ja minimiarvo yksiulotteiselle tehtävälle, kun $n = 4$

Dimensiot / m				Massatiedot / kg				Minimi-ratkaisu	$\|M\|$ / Nm	Tarkka arvo
x_1	x_2	x_3	x_4	m_1	m_2	m_3	m_4			
-5	-2	2.1	5	20	31	35	40	(2 3 4 1)	0,981	0,981

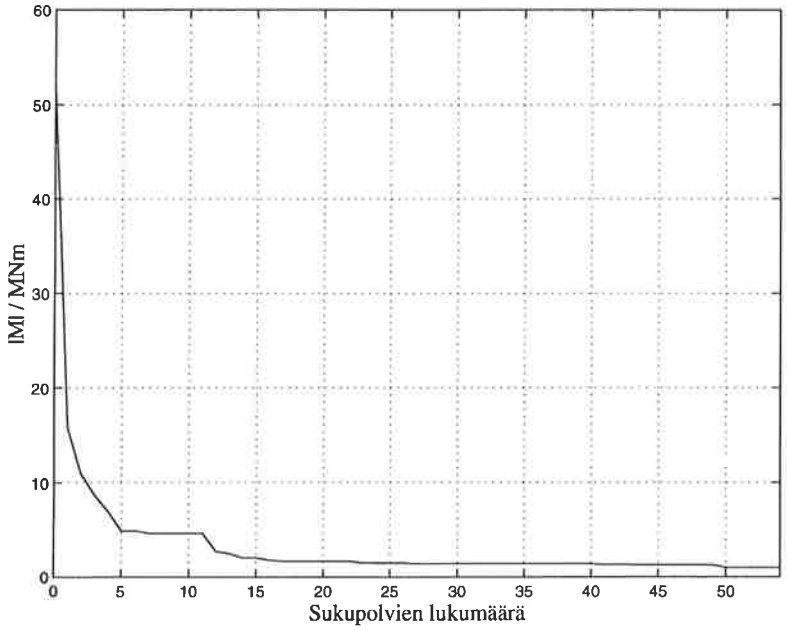
Yksiulotteista tehtävää ratkaistiin edelleen kasvattamalla muuttujien lukumäärää n taulukon 2 mukaisesti. Kullakin muuttujamäärällä ja vastaavalla datalla suoritettiin 10 riippumatonta ajoa. Ajojen optimiarvot kirjattiin talteen ja taulukossa 2 on esitetty optimiarvoista paras, huonoin sekä keskiarvo. Taulukossa on lisäksi keskiarvo yhteen ajoon tarvittavasta laskentatyöstä eri muuttujamäärillä n. Laskentatyö on esitetty suoritettujen analyysien eli funktion laskentakertojen lukumääränä ja toisaalta liukulukuoperaatioiden (*Floating point operations* eli FLOPS) lukumääränä. Tällöin jokainen suoritettu yhteen-, vähennys-, kerto- ja jakolaskuoperaatio vastasi yhtä flopsia.

Taulukko 2. Lasketut 10 ajon tulokset kasvavilla muuttujamäärillä n yksiulotteiselle tehtävälle, jolloin kohdefunktiona oli $|M|$.

Muuttujien lukumäärä (n)	**Populaation koko** (N)	**Sukupolvien lukumäärä** (n_{gen})	**Minimiarvo** / Nm				**Laskenta työ** / ajo	
			Paras arvo	*Keski-arvo*	*Huonoin arvo*	*Tarkka arvo*	*Analyysien lukumäärä*	*Flopsia* / 10^5
8	50	28	0	0,235	0,981	0	1400	0,499
16	66	32	0,029	0,742	2,119	0	2112	1,222
32	89	37	0,009	0,598	1,463	-	3291	3,393
64	122	44	0,023	0,634	3,685	-	5368	10,404
128	168	54	0,254	0,995	2,150	-	9072	33,854

Tuloksista nähdään, että muuttujamäärän n kaksinkertaistuessa kasvaa yhteen ajoon tarvittava laskentatyö käytetyllä geneettisellä algoritmilla noin kolminkertaiseksi, jos resursseja mitataan flopsien avulla. Funktion laskentakertojen lukumäärä, joka saadaan populaation koon ja sukupolvien lukumäärän tulona, kasvaa vastaavasti alle kaksinkertaiseksi muuttujamäärän kaksinkertaistuessa. Täten paremman arvion lisääntyvästä laskentatyöstä antaa nimenomaan mittaus liukulukuoperaatioiden avulla, sillä tällöin vertailussa ovat mukana myös muuttujien lisäämisestä pitenevien suunnitteluvektorien käsittelyt algoritmin eri osissa.

Kuvassa 3 on optimiarvon käyttäytyminen 10 ajon keskiarvona yksiulotteiselle tehtävälle sukupolvien n_{gen} funktiona, kun muuttujien lukumäärä n = 128. Tällöin siis kunkin sukupolven paras ratkaisu on mukana kuvaajassa. Todetaan, että suurin parannus kohdefunktion arvoon tulee muutaman ensimmäisen sukupolven aikana, jonka jälkeen sukupolvien lisäyksellä saatu parannus on pienimuotoista. Kohdefunktion käyttäytyminen oli vastaavaa muillakin muuttujamäärillä yksiulotteiselle tehtävälle, jonka kohdefunktio sisälsi yhden itseisarvolausekkeen. Sukupolvien järkevä lukumäärä tuleekin sovittaa tapauskohtaisesti geneettisen algoritmin sovelluskohteesta riippuen.



Kuva 3. Optimiarvon käyttäytyminen 10 ajon keskiarvona yksiulotteiselle tehtävälle sukupolvien funktiona, kun muuttujien lukumäärä n = 128

Kaksiulotteista tehtävää ratkaistiin kasvattamalla muuttujien lukumäärää n taulukon 3 mukaisesti. Kuten yksiulotteiselle tehtävälle, suoritettiin kullakin muuttujien lukumäärällä ja vastaavalla datalla 10 erillistä ajoa ja taulukossa on esitetty paras, huonoin sekä keskiarvo saaduista optimiarvoista. Taulukossa 3 on myös keskiarvo yhteen ajoon tarvittavasta laskentatyöstä eri muuttujamäärillä sekä vaadittujen analyysien että liukulukuoperaatioiden lukumääränä. Kohdefunktiona oli

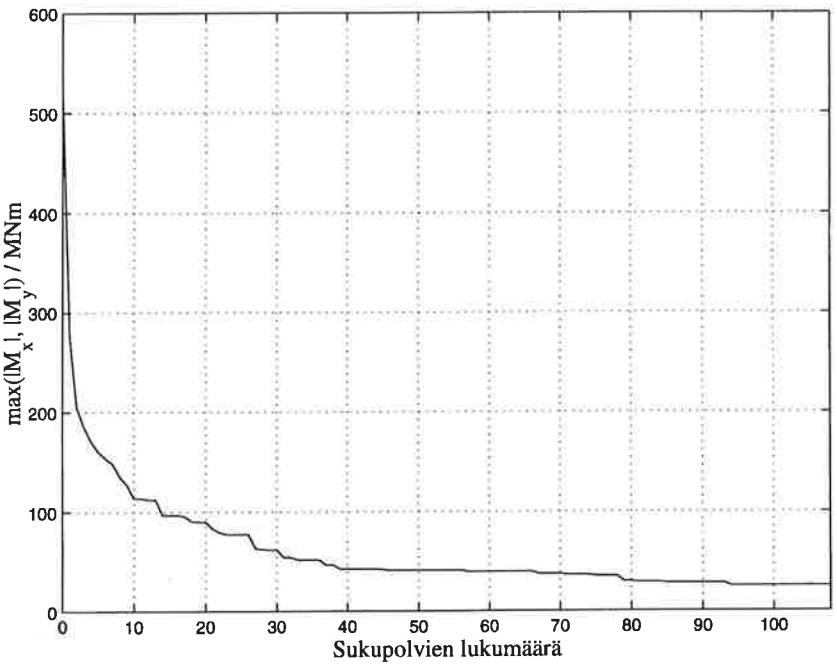
$$F(a_{ij}) = \max\left\{ \left| M_x(a_{ij}) \right|, \left| M_y(a_{ij}) \right| \right\} \tag{18}$$

joka saatiin lausekkeesta (10) painokertoimien arvoilla $\omega_1 = 1$ ja $\omega_2 = 1$.

Taulukko 3. Lasketut 10 ajon tulokset kasvavilla muuttujamäärillä n kaksiulotteiselle tehtävälle, jolloin kohdefunktiona oli $\max\{|M_x|, |M_y|\}$.

Muuttujien lukumäärä (n)	**Populaation koko** (N)	**Sukupolvien lukumäärä** (n_{gen})	**Minimiarvo** / Nm				**Laskenta työ** / ajo	
			Paras arvo	*Keski-arvo*	*Huonoin arvo*	*Tarkka arvo*	*Analyysien lukumäärä*	*Flopsia* / 10^5
16	66	64	1,501	6,162	13,008	0	4224	4,321
32	89	74	3,552	13,645	28,151	-	6582	12,636
64	122	88	7,094	22,616	39,556	-	10736	39,910
128	168	108	0,786	25,091	55,717	-	18144	132,32

Kuten yksiulotteiselle tehtävälle edellä kasvaa yhteen ajoon tarvittava laskentatyö muuttujamäärän kaksinkertaistuessa noin kolminkertaiseksi, jos resursseja mitataan flopsien avulla. Funktion laskentakertojen lukumäärä kasvaa sen sijaan alle kaksinkertaiseksi muuttujamäärän kaksinkertaistuessa. Kuvassa 4 on optimiarvon käyttäytyminen 10 ajon keskiarvona kaksiulotteiselle tehtävälle sukupolvien n_{gen} funktiona, kun muuttujien lukumäärä n = 128.



Kuva 4. Optimiarvon käyttäytyminen 10 ajon keskiarvona kaksiulotteiselle tehtävälle sukupolvien funktiona, kun muuttujien lukumäärä n = 128

Vertaamalla kuvien 3 ja 4 kohdefunktioiden minimiarvojen käyttäytymistä toisiinsa sukupolvien funktiona havaitaan, että kaksiulotteinen tehtävä tarvitsee enemmän sukupolvia eli iteraatiokierroksia optimiarvon konvergenssin tasaantumiseen. Tämä huomioitiin sukupolvien lukumäärän empiirisessä lausekkeessa (17).

4. YHTEENVETO

Työssä sovellettiin geneettistä algoritmia sijoitteluongelmien ratkaisemiseen. Algoritmin sisältämät operaatiot ohjelmoitiin siten, että kaikki generoitavat ratkaisut ovat käypiä. Kyseinen geneettinen algoritmi, kuten heuristiset algoritmit yleensäkin, pääsevät oikeuksiinsa vasta suuremmilla muuttujien lukumäärillä, sillä yleensä pienet tehtävät voidaan ja kannattaa ratkaista tarkasti. Lasketut esimerkit osoittavat, että algoritmi parantaa ratkaisua voimakkaimmin muutaman ensimmäisen sukupolven aikana. Koska tehtävälle saadaan aina ratkaisu myös suurilla muuttujien lukumäärillä, soveltuu kyseinen algoritmi parhaiten tilanteisiin, joissa kohtuullisen ratkaisun löytyminen on tärkeää ja riittävää. Varmoja takeita globaalin optimin saavuttamiseksi ei voida antaa.

LÄHTEET

[1] Nemhauser, G.L. and Wolsey, L. A., *Integer and Combinatorial Optimization*, John Wiley & Sons, Inc., New York, 1988.

[2] Rardin, R.L., *Optimization in Operations Research*, Prentice-Hall, 1998.

[3] Taha, H.A., *Operations Recearch: An Introduction*, 4th ed., Collier Macmillan, London, 1987.

[4] Haataja, J., *Optimointitehtävien ratkaiseminen*, Yliopistopaino, Helsinki, 1993.

[5] Floudas C.A., *Nonlinear and Mixed-Integer Optimization: Fundamentals and Applications*, Oxford University Press, New York, 1995.

[6] Laarhoven P. J. M. van and Aarts, E., *Simulated Annealing: Theory and Applications*, Reidel cop., Dordrecht, 1987.

[7] Goldberg, D.E., *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, Reading, MA,1989.

[8] Michalewitz, Z., *Genetic Algorithms + Data Structures = Evolution Programs*, 2nd ed., Berlin, Springer-Verlag, 1994.

[9] Koski, J., *Multicriterion Structural Optimization*. In: Adeli, H., Advances in Design Optimization, London, Chapman & Hall, pp. 194-224, 1994.

OPTIMIZATION OF CONDUCTING STRUCTURES USING THE HOMOGENIZATION METHOD

A. HILLEBRAND [1] T. KÄRKKÄINEN [1] and M. MIETTINEN [2]

University of Jyväskylä
[1] Department of Mathematical Information Technology
[2] Department of Mathematics
P.O. Box 35, FIN-40351 Jyväskylä, FINLAND

ABSTRACT

Optimization problems of conducting structures are considered. The homogenization method with rotated 2-laminated structures is used for relaxation of the original material layout problems. Numerical experiments are presented and some conclusions that can be drawn from the computations are given.

1. INTRODUCTION

In this paper we consider optimal layout problems of different conducting materials:

$$\inf_{a \in A_{ad}} J(u(a)) \text{ such that } F(a,u) = 0, \tag{1}$$

where F is a linear conductivity equation or a system of them and A_{ad} the set of admissible structures (described by means of the characteristic functions of different materials). The cost function J could be, e.g., the energy of the structure or the mean temperature within some subregion of the studied body. It is well known that these problems do not usually have classical solutions: fine scale structures yield better and better solutions. Therefore, it is necessary to relax (1) by extending the original set of structures. In the case of weakly continuous cost functions the right relaxation is given by adding homogenized conductivity tensors of all possible periodic microstructures to the control set A_{ad} (see [3]). This extended control set is called the G-closure of A_{ad}. The main difficulty is that the explicit characterization of the G-closure is known only in some rare cases. However, in practise, it is often sufficient to find the optimal structures from some subclass of all possible ones, e.g., from the class of single or two-rank laminates.

The aim of our work is to develop efficient and robust numerical methods for the structural optimization problems when optimization is performed over some class of

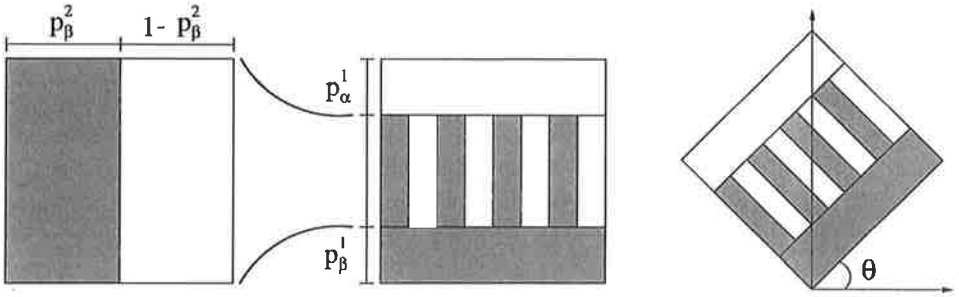


Figure 1: 2-laminated structure and its parametrization. The parametrization of the laminated structure can be described using the three volume fractions $p_\beta^2, p_\beta^1, p_\alpha^1$. The unknowns p_β^2, p_β^1 correspond to the amount of better material used on the two scales, and p_α^1 refers to the amount of the material α on the coarser scale. Finally, the relaxation is completed with an extra unknown θ to describe the rotation $R(\theta)$ of the laminate.

composite structures. In Section 2, we formulate a model problem, an optimal layout problem of two isotropic conducting materials. We consider two different objective functions: a minimal temperature and a given temperature. In Section 3 some numerical results are presented.

2. PARAMETRIZATION OF MODEL PROBLEMS

We have two linearly conducting isotropic materials with the conductivity tensors αI and βI with $\alpha < \beta$. The set of admissible materials A_{ad} for the optimal layout problems is defined by:

$$A_{ad} = \{a \mid a = \big((1-\chi(x))\alpha + \chi(x)\beta\big)I \ \text{ and } \ \int_\Omega \chi(x)\,dx = p\,|\Omega|\,\}, \qquad (2)$$

where p is the volume fraction of the better conducting material βI and χ its characteristic function.

Homogenization based relaxation of the original set A_{ad} is used since it allows us to determine the macroscopic (or effective) properties of the composites from their microgeometry when the scale parameter of the microstructure tends to zero. For layered materials effective properties are determined analytically by the width and orientation of the layers in the periodic cell [6].

In numerical calculations we perform optimization over the set of rotated 2-laminated structures (cf. Figure 1). In the case of two isotropic materials this actually gives the G-closure of A_{ad} (see [4, 5]) and, therefore, we obtain the full relaxation in Examples 1 and 2, i.e., we are able to optimize over all possible microstructures.

In this paper we consider two types of examples associated with (2): one state equation and a system of (two) state equations; both with weakly continuous cost functions.

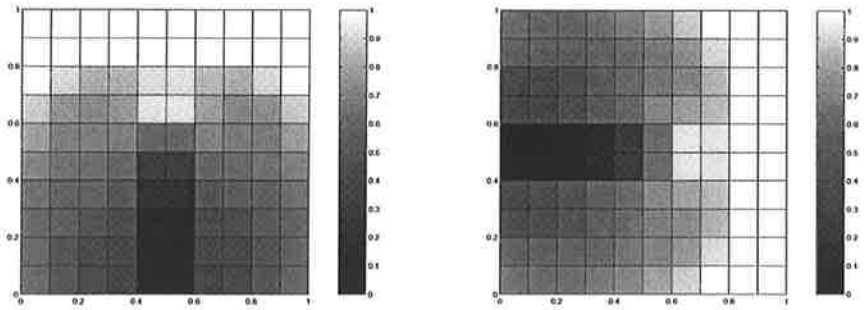


Figure 2: Computed optimal material density distributions in Example 1 with 10×10 material elements for $p = 0.3$ when using state equations separately: left figure corresponds to the first state equation and right figure to the second state equation, respectively. The black color represents better conductor βI and the white color worse conductor αI.

Example 1 Minimal temperature problem, two state equations:

$$\inf_{a \in A_{ad}} \frac{1}{2} \int_{\Omega_0} u_1(x)^2 + u_2(x)^2 \, dx,$$

where u_1 and u_2 solve

$$\begin{cases} -\mathrm{div}(a(x)\,\nabla u_1) &= 0 & \text{in } \Omega, \\ u_1 &= 0 & \text{on } \overline{\Gamma}_1, \\ u_1 &= 1 & \text{on } \overline{\Gamma}_3, \\ a\,\nabla u_1 \cdot n &= 0 & \text{on } \Gamma_2 \cup \Gamma_4, \end{cases} \qquad \begin{cases} -\mathrm{div}(a(x)\,\nabla u_2) &= 0 & \text{in } \Omega, \\ u_2 &= 1 & \text{on } \overline{\Gamma}_2, \\ u_2 &= 0 & \text{on } \overline{\Gamma}_4, \\ a\,\nabla u_2 \cdot n &= 0 & \text{on } \Gamma_1 \cup \Gamma_3. \end{cases}$$

The conductivities are given by $\alpha = 10^{-2}$ and $\beta = 1$. The computational domain is chosen as $\Omega = (0,1) \times (0,1)$ and the control region as $\Omega_0 = (0.4, 0.6) \times (0.4, 0.6)$. The boundary $\partial\Omega$ is split into four parts: $\Gamma_1 = (0,1) \times \{0\}$, $\Gamma_2 = \{1\} \times (0,1)$, $\Gamma_3 = (0,1) \times \{1\}$, and $\Gamma_4 = \{0\} \times (0,1)$.

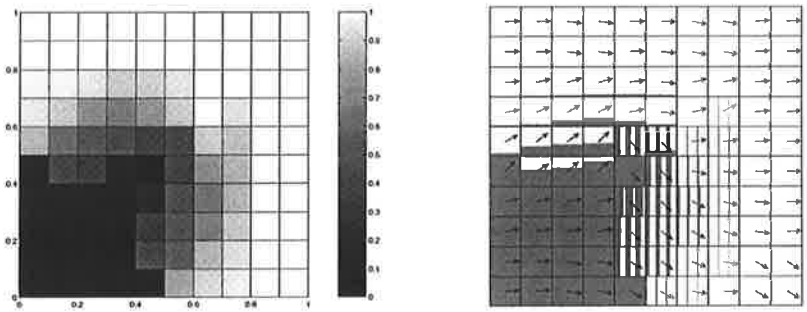


Figure 3: Computed optimal structure in Example 1 with 10×10 material elements for $p = 0.3$. Material density distribution (left) and 2-laminated structures (right) in each material element. The arrows indicate the rotation $R(\theta)$ of the 2-laminated structure (see Figure 1).

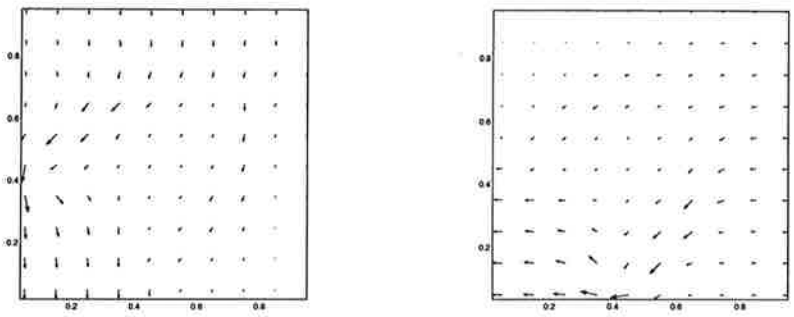


Figure 4: Heat fluxes of u_1 (left) and u_2 (right) for the obtained structure in Figure 3.

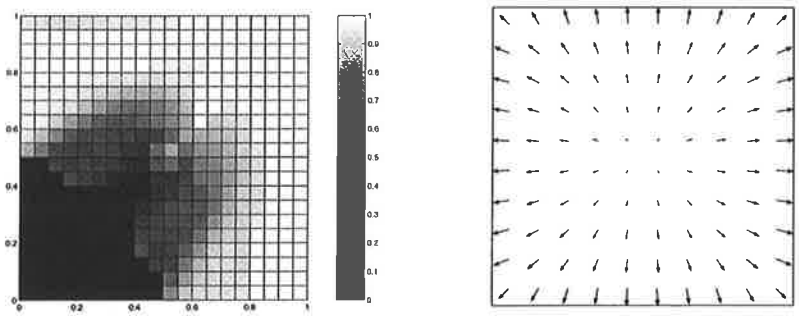


Figure 5: Computed material density distribution in Example 1 with 20×20 material elements for $p = 0.3$ (left); heat flux for the obtained structure in Example 2 for $p = 0.5$ (right).

Example 2 Given temperature:

$$\inf_{a \in A_{ad}} \frac{1}{2} \int_\Omega |u - z|^2 \, dx,$$

where u solves

$$\begin{cases} -\mathrm{div}(a(x) \nabla u) &= 10 \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega. \end{cases}$$

Here the conductivities are given by $\alpha = 1$ and $\beta = 10$, and the computational domain is $\Omega = (0,1) \times (0,1)$. The desired temperature is taken as $z = 4(x - x^2)(y - y^2)$.

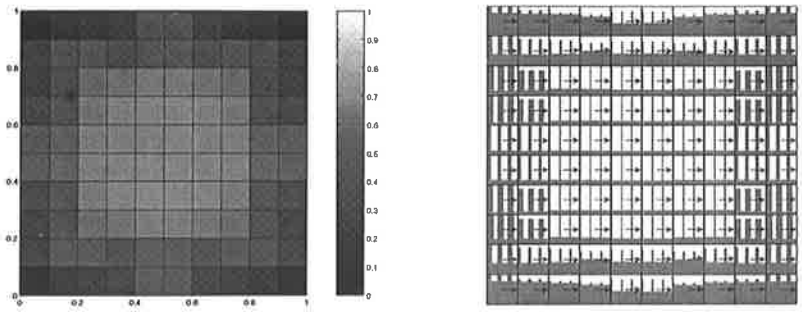


Figure 6: Computed material density distribution and 2-laminated structures in Example 2 with 10×10 material elements for $p = 0.5$.

3. NUMERICAL EXPERIMENTS

Obtained numerical results for Examples 1 and 2 are illustrated in Figures 2-6 and in Table 1. The more detailed results will be presented in [2] (see also [1] where a different class of composites is used).

3.1. Numerical realization

The numerical realization is based on the finite element method (FEM). The original domain Ω is divided into the so-called material elements Ω_i, where a piecewise constant P^0-approximation for the set of unknowns $((p_\beta^2)_i, (p_\beta^1)_i, (p_\alpha^1)_i, \theta_i)$ arising from the periodic cells is applied. The state and adjoint variables in material elements are approximated using biquadratic polynomials as basis functions. The actual optimization is performed with an SQP-method. The sensitivity analysis is carried out through the adjoint state equation and linear systems are solved using preconditioned CG-method.

3.2. Computational observations

Some general observations that can be drawn from the numerical experiments include:

Mesh-dependency: In Example 1 we get qualitatively the same solution for different mesh-sizes (cf. Figure 3 (left) and Figure 5 (left)).

Effect of discretization: We have performed a comparison of biquadratic (P^0/Q^2-discretization for the whole problem) and linear (P^0/P^1-discretization) finite elements for state and adjoint variables in Example 1 with only one state equation. The obtained solutions coincide if the the mesh ratio h_a/h_u for the P^0/P^1-discretization is large enough, i.e., $h_a/h_u \geq 4$. Otherwise, we get a wrong solution when using the P^0/P^1-discretization. This shows that by means of the number of unknowns in the linear problems the P^0/Q^2-discretization seems to be optimal.

		$p = 0.9$		$p = 0.3$	
x^0	algorithm	J^*	CPU	J^*	CPU
1	O	$1.171821 \cdot 10^{-4}$	114.97	$1.322683 \cdot 10^{-5}$	327.07
2	O	$3.266105 \cdot 10^{-4}$	124.44	$1.458877 \cdot 10^{-5}$	390.19
3	O	$2.948157 \cdot 10^{-4}$	345.22	$3.448160 \cdot 10^{-4}$	448.78
1	N	$1.171821 \cdot 10^{-4}$	43.98	$1.305134 \cdot 10^{-5}$	72.07

Table 1: Comparison of results using different initial guesses x^0 and different algorithms in Example 1 with 10×10 material elements. Letter O denotes ordinary algorithm on one grid and N nested iterations. J^* is final cost function value. Starting point $x^0 = 1$ is feasible point, $x^0 = 2, 3$ are not feasible points. CPU is the elapsed CPU-time in seconds.

Problem with discretization: Discrete maximum principle (M-matrix property) is not valid for linear systems with highly oscillating coefficients when either P^0/Q^2- or P^0/P^1-discretization is used. This means that the temperature is not necessarily non-negative and this must be taken into account in problem formulations.

Local minima: We have a lot of local minima in the optimization problems (cf. Table 1). This means that optimization either needs to be started using several initial configurations or some globalization strategy should be applied.

Problem with optimization software: Reliable results require fine discretization which means that the full Hessian approximation in an SQP-method is problematic due to its memory requirements.

Efficiency: Using nested iteration technique decreases the elapsed CPU-time considerably compared to the algorithm on one grid (see Table 1).

3.3. Physical interpretation of obtained structures

Example 1: According to Figures 2 and 3 the obtained structures direct the heat flow around the control region Ω_0. The worse conducting material αI forms an isolation zone around the "warm boundary". On the contrary, the better conductor βI is placed as far as possible from the "warm boundaries". The behaviour is the same for one or more state equations.

Example 2: Heat flux of obtained result is presented in Figure 5 (right). We can see that the obtained material distribution in Figure 6 is composed according to the desired temperature distribution. We notice also that the obtained microstructures are constructed from 2-laminated structures.

REFERENCES

[1] J. Dvořák, J. Haslinger, and M. Miettinen, *Homogenization & optimal shape design-based approach in optimal material distribution problems. Part I. The scalar case*, Advances in Math. Sciences and Appl., 29 (1999), pp. 665–694.

[2] J. Haslinger, A. Hillebrand, T. Kärkkäinen, and M. Miettinen. In preparation.

[3] R. Kohn and G. Strang, *Optimal design and relaxation of variational problems I-III*, Comm. Pure Appl. Math., 39 (1996), pp. 113–137,139–182,353–377.

[4] K. Lurie and A. Cherkaev, *Exact estimates of conductivity of composites formed by two isotropically conducting media taken in prescribed proportion*, Proc. Roy. Soc. Edinburgh, 99A (1984), pp. 71–87.

[5] F. Murat and L. Tartar, *Calcul des variations et homogénéisation*, in Les Méthodes de l'Homogénéisation: Théorie et Applications en Physique, Eyrolles, 1985, pp. 319–369.

[6] V. Zhikov, S. Kozlov, and O. Oleinik, *Homogenization of Differential Operators and Integral Functionals*, Springer Verlag, 1994.

STRUCTURAL ANALYSIS OF INNOVATIVE SOLUTIONS OF A LARGE COMPOSITE PRESSURE VESSEL FOR PROCESS INDUSTRY

H. MARTIKKA* and E. TAITOKARI†
*Faculty of Construction Design, Department of Mechanical Engineering
Lappeenranta University of Technology, P.O.Box 20,
FIN-53851 Lappeenranta, Finland
† Plastilon Oy , Muovikuja 7, FIN-55120 Imatra, Finland

ABSTRACT

Background for this research is the recognised need in industry for utilizing large vessels for processing hot and hazardous liquids with minimal ecological problems. The vessels are required to serve reliably a lifetime of several years with minimal initial cost, maintenance and still stand costs. The goal of this paper is to describe the design process of structural innovation and analysis of the global geometry and local details of a general class of vertical industrial composite vessels. The structure is made of glass fiber reinforced plastic, various steel structures and sandwich structures for the roof. The methods of analytic basic theory of mechanics and composite materials and FE method are used. An orthotropic failure criterion of Hill for three dimensional materials is used also as programmed to commercial FEM programs like MSC Nastran and Cosmos M to explore and pinpoint critical locations which need redesign. Optimal dimensions and material selections are found which satisfy criteria of good manufacturability, transportability, ease of assembly and the needs of the end user. The optimality of critical details is verified by materials testings.

1. INTRODUCTION

Background for this research is the recognised need for processing hazardous liquids with minimal ecological problems. A means to satisfy this need is use of processing vessels. The vessels are required to serve reliably a lifetime of several years with minimal initial cost, maintenance and still stand costs. The goal of this paper is to describe how a modern industrial design process is used to obtain innovative and optimal solutions for the global structure and for local critical details of a composite vessel.Cost effective material alternatives are glass fiber reinforced plastic and metal. The goal of the producer is to maximise the customer satisfaction within the technical and economic constraints like e.g economic manufacturability and strength. First global solutions are innovated using creative heuristics and experience. Then these ideas are analysed and the optimal choice is designed in detail. Final designs are optimised with FEM using suitable failure criteria programmed into FEM macros. In this study the innovative constructions and design calculations were done by E.Taitokari. Further analyses of these ideas were made in cooperation with Plastilon by H. Martikka and teams of FEM specialists.

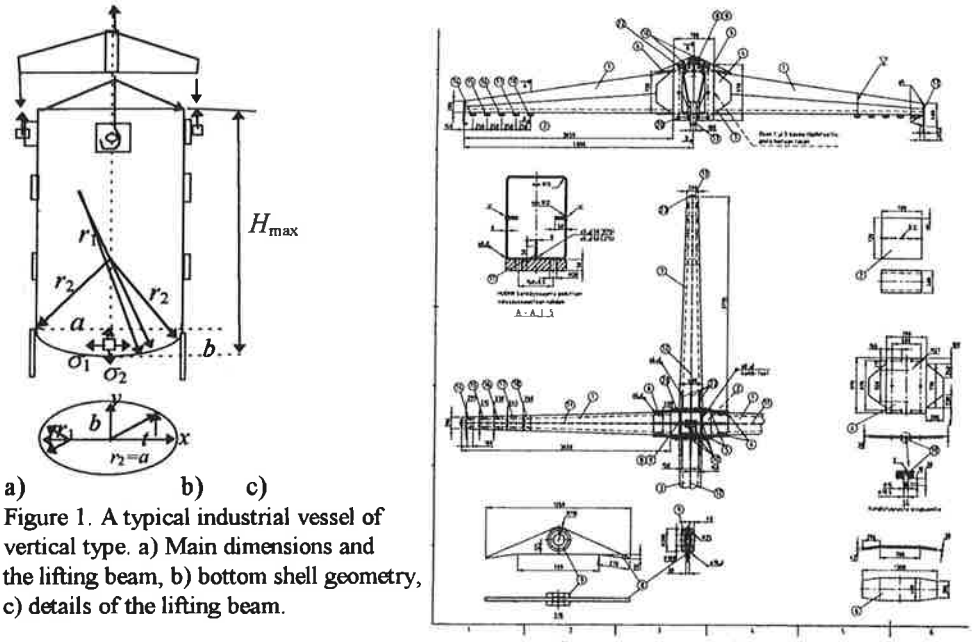
2. MATERIALS AND METHODS

The design tools were innovative idea generation, basic structural mechanics, design of composite material structures, optimal selection of materials and reinforcing alternatives, optimisation of the manufacturing , transportation and assembly.

3. RESULTS

3.1 General geometry and loads of the structure

A typical industrial vessel of vertical type is shown schematically in Figure 1 a.



a) b) c)

Figure 1. A typical industrial vessel of vertical type. a) Main dimensions and the lifting beam, b) bottom shell geometry, c) details of the lifting beam.

The radii of the cylinders of this class of vessels generally range from a = 5 to over 10 m. The wall thickness of the shell can be made to change continuously. Typical loadings are as follows

- Hydrostatic loading. The liquid height may vary in a range 10..30m causing pressure variation in a range $p = \rho g H_{max}$ = 0.1...0.4MPa .
- Thermal loading. Typical surface temperature range for the inner surface is 100...104°C and for the outer surface 40...50°C. Often insulation is used to cover the shell to reduce temperature gradients and heat losses .
- Chemical loading. This is due to aggressive liquids.

3.1.1 The ellipsoidal bottom shell

Common shape alternatives for the bottom shell are cone, sphere and ellipsoid.
For the ellipsoidal shell the tangential σ_1 and the meridian σ_2 stresses at point x, y are,

$$\sigma_1 = \frac{pr_2}{2t} \quad , \quad \sigma_2 = \frac{pr_2}{t}\left[1-\frac{r_2}{2r_2}\right] \quad , r_2 = \frac{1}{b^2}\left[a^4y^2 + b^4x^2\right]^{1/2} \quad , \quad r_1 = r_2{}^3\frac{b^2}{a^4} \quad (1)$$

The shape of the shell is expressed parametrically as $x = a\cos t$, $y = b\sin t$.

At the joint between the bottom shell and the cylinder or at the equator $y = 0$, $x = a$, the stresses for a typical structure may be calculated as

$$r_2 = a = 4 \quad , \quad r_1 = \frac{b^2}{a} = \frac{2.3^2}{4} = 1.32 \quad , \quad \rightarrow \sigma_1 = \frac{pa}{2t} \quad , \quad \sigma_2 = \frac{pa}{t}\left[1 - \frac{a^2}{2b^2}\right] \tag{2}$$

Substituting here typical numerical values, gives, for example

$$\sigma_1 = \frac{pa}{2t} = \frac{0.36\text{MPa} \cdot 8 \cdot 0.5}{2 \cdot 0.1} \approx 7 \quad , \quad \sigma_2 = 14\left[1 - \frac{4^2}{2 \cdot 2.3^2}\right] \approx -5$$

The design goal is to lower the risk of failure to an acceptable level. Now for the shell the criterion of failure can be diminished by choosing $a/b = \sqrt{2}$ since then $\sigma_2 = 0$ at the joint $x = a$. A typical value of ratio may be somewhat higher , or $a/b = 1.63$.

3.2 Optimal wall thickness for the cylindrical part of the vessel

The tangential stress of the cylinder is dominant. The thickness of the shell is mainly determined by the pressure and temperature loads. The thermal stress can be determined depending on the loading options chosen:
a) The temperature difference is constant in the shell wall but gradient may vary
b) The heat flow through an unit area of the shell is constant.

3.2.1 Optimal wall when the temperature difference is constant in the shell wall

The requirement of satisfactory useful life time may be expressed in many ways. One way to express the requirement is as a strength constraint

$$R(1) = \sigma_{\text{all}} - \sigma \geq 0 \tag{3}$$

The allowed stress σ_{all} can be defined in several ways. One simple method is to define it using allowed strain

$$\sigma_{\text{all}} = E_{\text{T}} \varepsilon_{\text{all}} = 25000\text{MPa} \cdot 0.2\% = 50\text{MPa} \rightarrow \tag{4}$$

The total stress is the sum of pressure and thermal stresses

$$\sigma = \sigma_{\text{p}} + \sigma_{\Delta t} = p\frac{r}{t} + \frac{E_{\text{T}}}{1 - \nu_{\text{TL}}}\alpha_{\text{T}}\left(\frac{t_1 - t_2}{2}\right) \quad , \quad p = \rho g(H - x) \tag{5}$$

The optimal wall satisfying this strength constraint may now be written as

$$\rho g r\left(\frac{H - x}{t(x)}\right) \leq \sigma_{\text{all}} - \sigma_{\Delta t} \rightarrow t(x) \geq K(H - x) \quad , \quad K = \frac{\rho g r}{\sigma_{\text{all}} - \sigma_{\Delta t}} \tag{6}$$

Typical numerical input values give for this factor and for the wall

$$K = \frac{\rho g r}{\sigma_{\text{all}} - \sigma_{\Delta t}} = \frac{1500 \cdot 10 \cdot 4\ \text{Pa}}{(50 - 15)\text{MPa}} = \frac{0.06\ \text{MPa}}{35\ \text{MPa}} \approx 1.6 \cdot 10^{-3}$$
$$t(x) \geq K(H - x) = 0.0016 \cdot (25 - x) \tag{7}$$

At the joint location $x = 0$, the wall recommendation is $t(0) = 0.04$m and at mid height $x = \frac{1}{2}H$ the wall recommendation is $t = 0.020$m.

3.2.2 Optimal wall when the heat flow per unit area of the shell is constant.

Now an average heat flow may be assumed through an average wall of 0.025m.

$$q = \lambda\left(\frac{T_1 - T_2}{t}\right) = 0.5\frac{\mathrm{W}}{\mathrm{mK}}\left(\frac{105^\circ\mathrm{K} - 40^\circ\mathrm{K}}{0.025\mathrm{m}}\right) = 1300\frac{\mathrm{W}}{\mathrm{m}^2} \tag{8}$$

The total stress is now

$$\sigma = \sigma_p + \sigma_{\Delta t} = p\frac{r}{t} + \tfrac{1}{2}q\cdot t\frac{E_T}{1-\nu_{TL}}\frac{\alpha_T}{\lambda} = \sigma = p\frac{r}{t} + A\cdot t \tag{9}$$

For steel $\lambda = 47$ W/Km, for composite materials $\lambda = 0.5$.Here the constant is

$$A = qB = q\tfrac{1}{2}\frac{E_T}{1-\nu_{TL}}\frac{\alpha_T}{\lambda} = 1300\tfrac{\mathrm{W}}{\mathrm{m}^2}\tfrac{1}{2}\frac{14000\mathrm{MPa}}{1-0.2}\frac{20\cdot 10^{-6}\,\frac{1}{\mathrm{K}}}{0.5\,\frac{\mathrm{W}}{\mathrm{mK}}} = 450\frac{\mathrm{MPa}}{\mathrm{m}} \tag{10}$$

The optimal wall satisfying this strength constraint may now be written as

$$R(2) = \sigma_{all} - \sigma \geq 0 \;\rightarrow\; p\frac{r}{t_H} + At_H \geq \sigma_{all} \rightarrow At_H^{\,2} - \sigma_{all}t_H + pr \geq 0 \tag{11}$$

The recommended wall thickness at $x = 0$ is $t_H = 0.05$, and at $x = \frac{1}{2}H$ it is $t_H = 0.02$m.

3.2.3 Some ideas for closing the opening in the bottom shell by a plate

The bottom plate may be joined by several well known ways to the shell.

3.2.3.1 A conventional flanged tube for joining the plate to shell

This construction is shown in Figure 2a as a segment cut by symmetry planes.

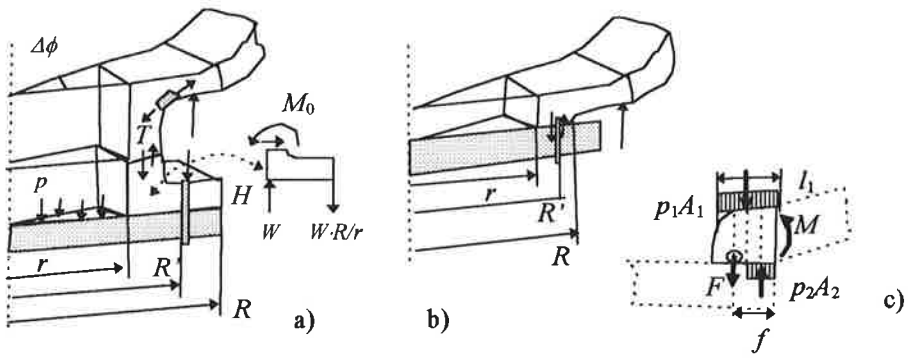


Figure 2. Some typical methods of joining the plate to the bottom shell.a) a conventional flange model, b) an improved joint model, c) force resultants.

The pressure p acts on the circular plate of radius r producing a force F. It is supported by the screws of the flange joint. The line force per unit length is W

$$F = p\pi r^2 \quad , \quad W = \frac{F}{2\pi r} \tag{12}$$

The moment balance gives the moment and the bending stress at the root of the flange

$$M_0 = W\Delta R = pr(R-r) \quad , \quad \sigma = \frac{M_0 l}{lT^2/6} \rightarrow \sigma = 6\left(\frac{r}{T}\right)^2\left(\frac{R}{r} - 1\right)p \tag{13}$$

This formula always overestimates the stresses according to Blake [1]. It can be seen that conventional design produces a stress which is not desired. It can be minimized using the guidelines suggested by the design formula. By decreasing the moment arm or by setting R close to radius r the bending stress at the shell is minimized.

3.2.3.2 An improved joint for joining the plate to shell

An improved plate to shell bolt joint is shown in Figure 2b. The equilibria give

$$\Sigma F = p_1 A_1 - p_2 A_2 + F = 0 \quad , \quad \Sigma M = M - p_1 A_1 (f - l_2) + p_2 A_2 (f - \tfrac{1}{2} l_2) = 0 \qquad (14)$$

Thus it is possible to obtain minimal bending moment $M = 0$.

3.2.4 Lifting beam

The total load to be lifted by 4 trunnions is $F_t = 4\,F' = 320$ kN. Thus the force per one trunnion is $F' = 8$ kN. The dynamic factor $K_{dyn} = 1.1$. According to safety norm SFS 200 the applied load is $F = 1.1\,F' = 8.8$kN. The material is Fe52-D6 with UTS $R_m = 490$MPa and yield strength $R_e = 355$MPa. The allowed stresses are obtained by dividing the strength by required factors of safety

$$\sigma_{Rm} = R_m / 4 = 490/4 = 123\text{MPa} \quad , \quad \sigma_{Re} = R_e / 1.5 = 355/1.5 = 236\text{MPa}$$

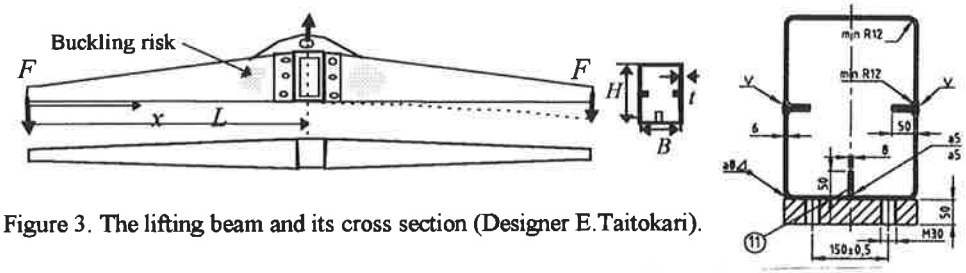


Figure 3. The lifting beam and its cross section (Designer E.Taitokari).

The half length of the beam is $L = 3.85$ m and the wall thickness is constant $t = 0.006$m. The dimensions of the variable cross section are : At tip $x = 0$ is $H_1 = 0.2$, $B_1 = 0.2$ and at $x = L$, $H_2 = .7$, $B_2 = .4$ m.

3.2.4.1 Deflection and strength constraints on the beam

Deflections and stresses along the beam were obtained by numerical integration

$$M = Fx \quad , \quad f = \frac{dC}{dF} = \frac{d}{dF}\left[\int_0^L \frac{M^2}{2EI}dx\right] = \int_0^L \frac{M}{EI}\frac{dM}{dF}dx = \int_0^L \frac{Fx^2}{EI(x)}dx$$

$$\sigma = \frac{M\frac{1}{2}H}{I(x)} = \frac{Fx}{2EI(x)} \quad , \quad I(x) = \frac{1}{12}\left(B \cdot H^3 - b \cdot h^3\right) \quad , \quad b = B - 2t, h = H - 2t \qquad (15)$$

$$H = H_1 + \frac{H_2 - H_1}{L}x, B = B_1 + \frac{B_2 - B_1}{L}x$$

Using the input data gave for the deflection 18 mm. Now the standard gives a constraint for deflections $v < L/200 = 3850/200 = 19$mm. The applied stress ranges between 130..146..132 MPa. The allowed stress for UTS is $\sigma_{Rm} = 123$MPa.

3.2.4.2 Buckling strength estimates

A buckling model is shown in Figure 4. The model is applied to the beam side wall.

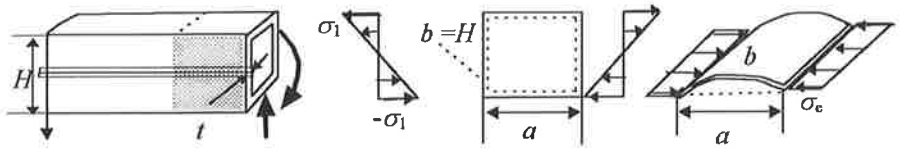


Figure 4. Buckling model for a side plate of the beam.

The buckling strength can be estimated using the following considerations [2], [3]
a) A rectangular plate of physical length $L_{phys} = a$ is loaded in a direction along edge b by a linear bending stress distribution. The aspect ratio $\alpha = a/b = a/H$ is now close to 1. The ratio of plate thickness to load line is $t/b = t/H = .006/.7 = 8.6\cdot10^{-3}$.
b) The plate is freely supported on all edges and the buckling factor is $k = 26$ by equation (16). c) The buckling stress is k times the basic Euler stress σ_e when a plate is compressed along b side in direction of the freely supported a-edge, Figure 4.

$$\sigma_{el} = k\cdot\sigma_e = k\cdot\frac{\pi^2}{12(1-\nu^2)}E\left(\frac{t}{b}\right)^2 = 336\text{MPa}$$
$$k = 15.97 + \frac{1.87}{\alpha^2} + 8.6\alpha^2 \quad ,\alpha = \frac{a}{b} \quad , \quad \frac{t}{b} \to \frac{t}{H} = \frac{.006}{.7} = 8.6\cdot10^{-3} \tag{16}$$

This means that buckling may start at stress 336MPa or close the yield strength of 355 MPa. But there is an additional stiffener at the side wall which may double the ratio t/b causing the buckling strength to increase by a factor of about 4 .

3.2.5 Trunnions for lifting the shells

The vessel is assembled as components. These have to be lifted and also rotated. The forces for lifting the shell modules are applied to trunnions designed by E.Taitokari.

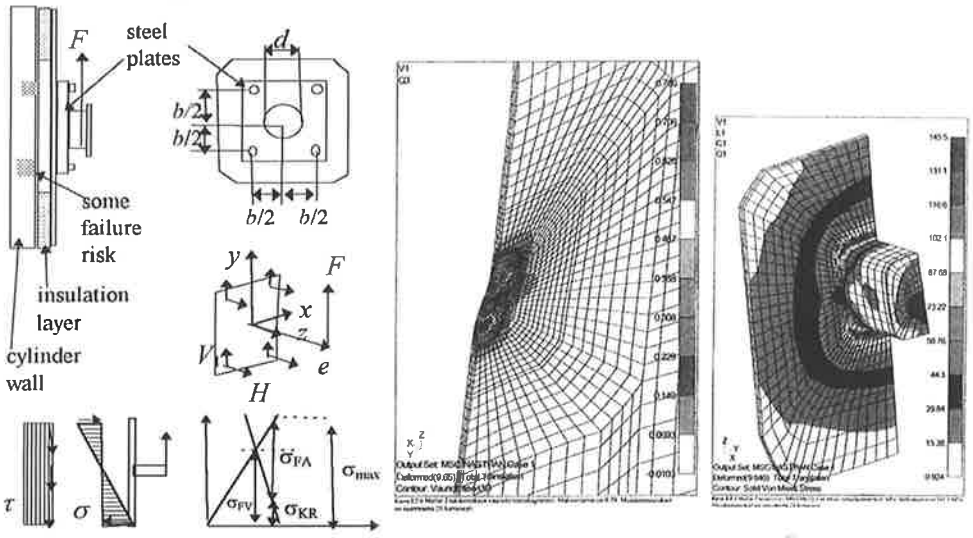


Figure 5. The trunnion joint .a) Schematic view, b) Hill [4] failure criterion, max =0.76.

3.2.5.1 Screw joints of the trunnion

The trunnion is welded to a plate which is joined with four screws to a large inner plate. The equilibrium of forces and moments gives the loads on the bolts

$$\begin{aligned}
&\Sigma F_x = +2H_1 + 2H_2 = 0 \;\rightarrow\; H_1 = -H_2 = H \\
&\Sigma F_y = +2V_1 + 2V_2 + F = 0 \;\rightarrow\; V_1 = V_2 = \tfrac{1}{2}F \\
&\Sigma M_x = +b\cdot H_1 + b\cdot H_2 + eF = 0 \;\rightarrow\; H = \tfrac{e}{2b}F = \tfrac{0.08\text{m}}{2\text{-}0.4\text{m}}\,88000 = 8800
\end{aligned} \tag{17}$$

The maximum stress on the bolts is acceptable if the tightest constraint is satisfied.

1) the maximum allowable stress is constrained and 2) joint loosening is not allowed

$$\begin{aligned}
&\sigma_{\max} = \sigma_{FV} + \phi_n \sigma_{FA} < R_e \;\rightarrow 100 + 0.3\sigma_{FA} < 200\ \text{MPa} \rightarrow \\
&H < \sigma_{FA} A_3 < 330\ \text{MPa}\cdot A_3 = 330\ \text{MPa}\cdot 225\cdot 10^{-6}\,\text{m}^2 = 74000 \\
&\sigma_{FA} < \sigma_{KR} = \frac{\sigma_{FV}}{1-\phi_n} = \frac{100}{1-0.3} = 140 \;\rightarrow H < 31500\ \text{N}
\end{aligned} \tag{18}$$

Here σ_{FV} is the initial prestressing stress on a screw, A_3 is stress area of the screw, σ_{FA} is the applied load stress, ϕ_n is joint stiffness factor.The stress to seal the joint is σ_{KR}.

3.2.5.2 Bending and shearing stress at the trunnion joint

Constraint on the endurance of the glue joint can be derived from equations of static equilibria. For the glue joint allowed tensile stress is 5 and shear stress is 1MPa.

$$\Sigma M = Fe - \sigma W = 0 \quad , \quad W = \tfrac{1}{6}bh^2 \qquad h = b$$

$$R(1) = \sigma_{sall} - \sigma \geq 0 \quad ,\rightarrow b^3 \geq \frac{6F\cdot e}{\sigma_{sall}} = \frac{6\cdot 88000\text{N}\cdot 0.08\text{m}}{5\text{MPa}} \rightarrow \quad b > .02\text{m} \tag{19}$$

$$\Sigma F = F - \tau b^2 = 0 \rightarrow R(2) = \tau_{sall} - \tau = 0 \rightarrow b^2 \geq \frac{F}{\pi\tau_{sall}} = \frac{88000\ \text{N}}{3\cdot 1\ \text{MPa}} \rightarrow \quad b > .17\text{m}$$

The planned width $b = 0.4$m is satisfactory.The Hill failure criterion [4],[5] gave max values of 0.76 at two small locations where the bending curvature is biggest. The allowable value of the criterion is 1.

3.2.5.3 Local bending at the trunnion joint

The bending moment is $M_o = F\cdot e$. The total deflection at the vicinity of the load area is

$$w(x) = e^{-\beta x}\left[C_1 \cos\beta x + C_2 \sin\beta x\right] + e^{\beta x}\left[C_3 \cos\beta x + C_4 \sin\beta x\right] + w_1(x) \tag{20}$$

The boundary conditions are

$$\begin{aligned}
&w_1(\infty) = w_0(\infty) + w_1(\infty) = 0 \rightarrow w_1(x) = 0, Z = 0 \quad , \quad w(0) = 0 \\
&\rightarrow C_3 = 0 \quad , \quad C_4 = 0 \quad , \quad C_1 = 0 \quad , \quad M_{xx} = -Kw_{xx} = M_0, \quad at \quad x = 0
\end{aligned} \tag{21}$$

The deformation is

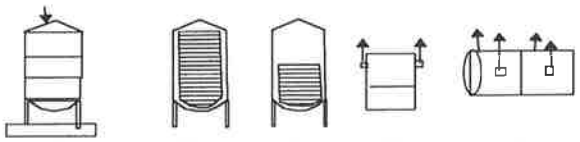
$$w(x) = e^{-\beta x}\left[C_2 \sin\beta x\right] \quad , \quad C_2 = \frac{-1}{2K\beta^2} M_0 \tag{22}$$

Stress estimate due to local bending is 42 MPa with maximal lever $e = 0.08$m.

$$\sigma = \frac{M_0 b}{W} \quad , \quad W = \tfrac{1}{6}bh^2 \quad , \quad \sigma = \frac{F\cdot e}{h^2/6} = \frac{88000N\cdot 0.08m}{0.03^2 m^2/6} \approx 42MPa \tag{23}$$

3.2.6 Joints betweeen modules of the cylindrical shell

These are shown schematically in Figure 6.



a) b) c) d) e)

Figure 6. Loadings on a vessel.a) Structure with joints and load on the roof. b) full hydrostatic loading. c) discontiuity of loading on a joint, d) lift up loading , e) bending loading.

3.2.6.1 Failure criteria analysis of the joint

The cross section of the joint and some results of the failure criterion F of Hill at some locations are shown. The structure was modified to decrease the criterion values.

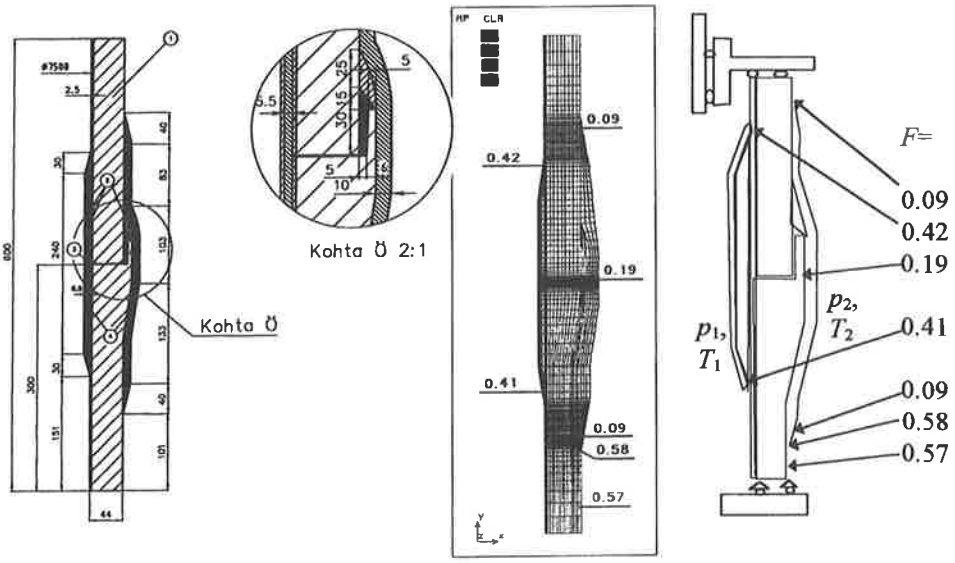


Figure 7. The joint construction. a) A CAD model,a) a FEM model showing failure criterion results c) boundary conditions for the load case 1, $p_1 = 0.3$MPa, $T_1 = 104°C$, $p_2 = 0$, $T_2 = 40°C$.

Several empirical modifications have been suggested so that octahedral shear stress criterion can be used for anisotropic or uneven materials. Anisotropic materials have different properties in different directions, and uneven materials have different properties in tension versus compression. Now in the present case the locations studied are within a three dimensional orthotropic material. In thinner laminate sections the Tsai-Hill and Tsai-Wu [5] criterions are applicable. Now the criterion described by Hill [4] is used. Let us consider an element of material whose X-Y-Z axes are aligned with planes of material symmetry. The material is constructed by building up the body by laminate layers. The uniaxial strengths in main directions are

$$\sigma_{LU} = 250 \quad , \quad \sigma_{TU} = 250 \quad , \quad \tau_{LTU} = 10$$

Elastic moduli are for a laminate layer are typically

$E_L = 14000\text{MPa}$, $E_T = 14000\text{MPa}$, $E_z = 10000\text{MPa}$

At first the Poisson's ratio were assumed to be equal for simplicity

$\nu_{LT} = 0.3$, $\nu_{TL} = 0.3$, $\nu_{Lz} = 0.3$

Table 1. Material data for the FRP structure.

	Stress	Strength	Elastic modulus
X-direction	$\sigma_X = \sigma_L$	$X = \sigma_{XU} = 250$	$E_X = 14000$
Y-direction	$\sigma_Y = \sigma_T$	$Y = \sigma_{YU} = 250$	$E_Y = 14000$
Z-direction	σ_Z	$Z = \sigma_{ZU} = 80$ compr..25 tens.	$E_Z = 10000$
XZ-plane Y-direction	$\tau_{XY}=\tau_{LT}$	$\tau_{XYU} = S_{XY} = 10$	$G_{XY} = 2000$
XY-plane Z-direction	τ_{YZ}	$\tau_{YZU} = S_{YZ} = 10$	$G_{YZ} = 2000$
YZ-plane X-direction	τ_{ZX}	$\tau_{ZXU} = S_{ZX} = 10$	$G_{ZX} = 2000$

The Hill criterion is [4]

$$H(\sigma_X - \sigma_Y)^2 + F(\sigma_Y - \sigma_Z)^2 + G(\sigma_Z - \sigma_X)^2 + 2N\tau_{XY}^2 + 2L\tau_{YZ}^2 + 2M\tau_{ZX}^2 = 1 \quad (24)$$

The parameters are

$$2N = \frac{1}{S_{XY}^2} \quad . \quad 2L = \frac{1}{S_{YZ}^2} \quad . \quad 2M = \frac{1}{S_{ZX}^2} \quad (25)$$

$$F = \frac{1}{2}\left[-\frac{1}{X^2} + \frac{1}{Y^2} + \frac{1}{Z^2}\right], \quad G = \frac{1}{2}\left[+\frac{1}{X^2} - \frac{1}{Y^2} + \frac{1}{Z^2}\right], \quad H = \frac{1}{2}\left[+\frac{1}{X^2} + \frac{1}{Y^2} - \frac{1}{Z^2}\right] \quad (26)$$

This criterion highlights the weak points. These occur especially at those places where there are at the same time a) tensile stress acting in the low strength thickness *z*-direction, b) in places where there is high shear stress between the fiber layers and c) in places where there are tensile and compressive normal stresses.

3.2.6.2 The joints considered as a stiffener ring

The stiffener ring can be considered as a short cylinder under pressure

$$P = pH_r \quad , \quad A_r = H_r T_r \quad (27)$$

, where P is line load (N/m) due to pressure p (N/m^2), H_r is axial height and T_r is thickness. Displacement of a short cylindrical stiffener ring is positive outwards

$$w_P = \frac{pa}{E_r T_r} \rightarrow P = p \cdot H_r \rightarrow w_P = \frac{Pa^2}{E_r A_r} \quad (28)$$

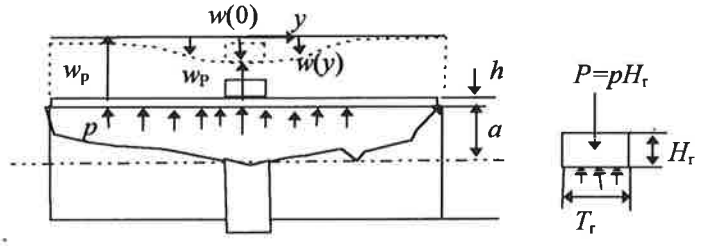


Figure 8. A stiffener ring in a cylinder.

The bending stress due to the perturbation caused by the ring may be written in a dimensionless form

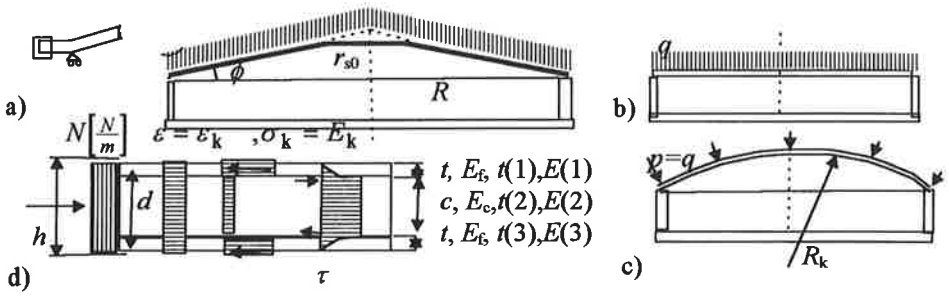
$$\sigma_{max} = \frac{M}{W} = \frac{M_0}{h^2/6} = \frac{3}{h\beta\left[2\frac{E}{E_r}\cdot\frac{h^2}{A_r}+h\beta\right]}\cdot p = \frac{3}{C\sqrt{x_3}\left[2\cdot x_1\cdot x_2 + C\sqrt{x_3}\right]}\cdot p \qquad (29)$$

$$x_1 = \frac{E}{E_r} \quad , \quad x_2 = \frac{h^2}{A_r} \quad , \quad x_3 = \frac{h}{a} \quad , \quad C = \left[3\left(1-\nu^2\right)\right]^{1/4}$$

Substituting here typical values , $x_1 = 1$, $x_2 = 0.2$, $x_3 = .01$ and $p = 0.3$MPa gives for the bending stress about 16 MPa.

3.2.7 The conical sandwich roof

This roof was designed using simplified models, Figure 9, and using a detailed FEM model, Figure 10. It is loaded with vertical pressure $q = 4000\text{N/m}^2$ and thermal stresses.



Fiure 9. Models for the conical sandwich roof. a) The cone model, b) circular plate, c) a shallow sphere model and d) sandwich structure.

3.2.7.1 FEM model results for the sandwich roof

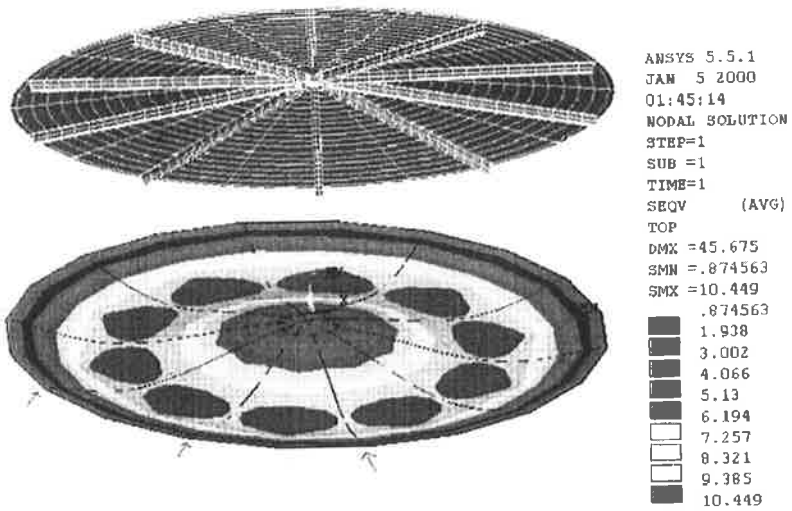


Figure 10. Results of ANSYS 5.5.1 [6] FEM program models.

3.2.7.2 Analytical models for the sandwich roof

Now the stiffeners were not considered in analytical models.

a) The circular plate model

The bending moment per unit length is according to Szilard [7].

$$m_r = -\frac{1}{16} q \cdot R^2 (3+\nu) C_1 \quad , \quad C_1 = 1-\rho^2 \quad , \quad \rho = \frac{r}{R} \quad , \quad C_0 = 1-\rho^4 \qquad (30)$$

For a sandwich beam model the bending stress in a layer number k is

$$\sigma_{k,b} = \frac{M}{\bar{I}_k} y \quad , \quad \bar{I}_k = \frac{\Sigma E \cdot I}{E_k} = \frac{D'}{E_k} \quad , \quad D' = \tfrac{1}{2} E_f \cdot t d^2 \cdot b = D \cdot b \quad , \quad M = m_r \cdot b \qquad (31)$$

the shear stress is

$$\tau = \frac{Q}{bd} = \frac{q_r}{d} \quad , \quad q_r = \frac{Q}{b} = -\tfrac{1}{2} qR\rho = -q\frac{r}{2d} \quad , \quad \rho = \frac{r}{R} \qquad (32)$$

b) The cone model

In the cone now only in plane non bending stresses are considered using line loads

$$\sigma_k = \frac{N_s}{\bar{A}_k / b} \left[\frac{\mathrm{N/m}}{\mathrm{m^2/m}}\right] = \frac{N_s}{\bar{t}_k} \left[\frac{\mathrm{N/m}}{\mathrm{m}}\right] \quad , \quad \bar{A}_k / b = \bar{t}_k = \frac{\Sigma E_i \cdot t_i}{E_k} \qquad (33)$$

Two constraints of endurance are applied: a) prevention of the crushing of the core balsa b) prevention of the wrinkling of the compressed surface layers using Allen's [8] models

$$R(2) = \sigma_{c,pur} - \sigma_2 \geq 0 \quad , \quad 12.9\mathrm{MPa} - \frac{N_s}{\bar{t}_2} \geq 0 \qquad (34)$$

$$R(3) = \sigma_{wrink} - \sigma_1 \geq 0 \quad , \quad \sigma_{wrink} - \frac{N_s}{\bar{t}_1} \geq 0 \quad , \quad \sigma_{wrink} = 0.5\left[E_f \cdot E_c^{\,2}\right]^{1/3} \qquad (35)$$

c) The shallow spherical shell model

By this method an effective radius of curvature was used and the vertical load stress q was equated to hydrostatic pressure, p, $q = p$. Then the model of Dast Richtlinie 13 [9] was used. In this model the material is isotropic.The ideal buckling strength is decreased using a strength reduction factor depending on boundary conditions.

Case 1 : shell edges are fixed giving $\alpha = 0.15$, Case 2: shell edges are free to rotate giving $\alpha = 0.13$ and Case 3: shell edges are supported by rollers giving $\alpha = 0.06$. Now this latter is chosen as the best approximation.

An equivalent isotropic model is used. The elastic moduli are equated and the thicknesses, too. In a case study the following values are chosen. Material class IM = 1 with 30 weight % glass fiber at the faces of the sandwich and geometry class IG = 1 of the thinnest thickness choice for the core $c(1) = .050$ m are typical choices.

$$\bar{E} \cdot \Sigma t = \Sigma(E \cdot t) = E_1 \cdot t_1 + E_2 \cdot t_2 + E_3 \cdot t_3 = 8000 \cdot 0.014 + 4070 \cdot 0.05 + 8000 \cdot 0.014 = 430$$

$$\rightarrow \quad \Sigma t = d(IG) = 0.0141 + 0.050 = 0.064 m$$

$$\rightarrow \quad \bar{E} = \frac{\Sigma(E \cdot t)}{\Sigma t} = \frac{430\ \mathrm{MPa}}{0.064\ m} = 6700\ \mathrm{MPa} \qquad (36)$$

The stresses of the simple equivalent shell model is

$$\sigma_\theta = \sigma_\phi = \sigma = q\frac{R}{2t} \rightarrow q\frac{R_{\text{kugel}}}{2\cdot\Sigma t} = 4000\frac{\text{N}}{\text{m}^2}\cdot\frac{13\text{ m}}{0.064\text{ m}} = 0.4\text{ MPa} \tag{37}$$

The real buckling strength and factor of safety are

$$\sigma_{\text{Ki}} = \frac{E}{\sqrt{3(1-\nu^2)}}\frac{t}{R} \equiv 0.605E\frac{t}{R} = 0.605\cdot\frac{430\text{MPa}\cdot\text{m}}{13\text{m}} \quad , \quad \nu = 0.3$$

$$\sigma_{\text{Be}} = \alpha\sigma_{\text{Ki}} = \alpha \bullet 0.605 \bullet \frac{Et}{R_{\text{kugel}}} = 0.06 \bullet 20 = 1.7\text{MPa} \quad , N_{\text{kugel}} = \frac{\sigma_{\text{Be}}}{\sigma} = 2.9 \tag{38}$$

4. DISCUSSION

The total process used by manufacturers of fiber reinforced composite products for manufacturing important and safety critical industrial process is challenging. Although main dimensions and loads may be specified much innovative design ideas need to be generated and then optimally selected. At the next stage the best ideas designed optimally and then manufactured to satisfy the increasingly high reliability, quality, cost-effectiveness and manufacturability requirements.

5. CONCLUSIONS

The following conclusions may be drawn:

- The construction of advanced fiber reinforced composite vessels requires innovative and optimal design which should be done concurrently with manufacturing.
- Highly satisfying solutions are obtained by using efficient design methods, basic and advanced structural mechanics, FEM programs and materials science. Good team work and cooperation with end users are essential for success in high tech work.

Acknowledgements

The assistance in FEM modelling of the roof by Toni Heikkilä, Jarkko Bohm and Aleksi Mäkelä are gratefully acknowledged and also the FEM modelling of the cylinder joint by Arto Heikkinen, Reijo Hämäläinen and Juhani Laine.The lower structure and lifting joint were modelled by Stressfield Oy and by Juhani Laakso, Plastilon Oy.

REFERENCES

[1] Blake, A., Practical stress analysis in engineering design, Marce Dekker, Inc, New York, 1990.

[2] Steel structures , The Finnish structural design code, part B7, Helsinki 1987.

[3] SFS-ENV 1993-1-1 .Eurocode 3: Design of steel structures.Part 1-1: General rules and rules for buildings (in Finnish) SFS,Helsinki 1993.

[4] Hill, R,The mathematical theory of plasticity, Oxford University Press, London, 1983

[5] Dowling, N. E; Mechanical Behaviour of Materials, Prentice Hall, 1999

[6] ANSYS 5.5.1 FEM program.

[7] Szilard, R., Theory and analysis of plates.Classical and Numerical Methods. Prentice Hall-Inc , 1974.

[8] Allen H G, Analysis and Design of Structural Sandwich Panels, Pergamon press Ltd, Oxford, 1969.

[9] Dast Richtlinie 13 Juli 1980, Beulsicherkeitsnachweis fur Schalen,4. Unversteifte Kugelschalen4.2 Nicht abgestufte Kugelschalen unter Aussendruck4.2.1 Konstante Aussendruck.

CROSS-SECTION PROPERTY CALCULATION OF GENERALISED BEAM THEORY IN PLATE STRUCTURES USING BEAM EIGENFREQUENCY ANALYSIS

Tapani Halme
Researcher
Lappeenranta University of Technology
Department of Mechanical Engineering

ABSTRACT

Generalised Beam Theory (GBT) is an extension of the conventional engineering beam theory, which in addition to the basic deformations, i.e. axial deformation, bending about two main axes and torsion, tackles the higher order modes or distortional modes of prismatic structural elements. In this paper a simple eigenvalue extraction method is presented which allows for obtaining the cross-sectional properties of plate structures in the framework of GBT. The proposed method is based on the standard eigenfrequency analysis of a plane frame beam structure using the finite element method (FEM).

1. INTRODUCTION

Generalised Beam Theory (GBT) developed by Prof. Schardt [1] and his associates in Darmstadt is a tool for the analysis of prismatic thin-walled structures. The theory includes the Bernoulli-Euler-Vlasov beam theory and extends the analysis to the distortion of the cross-section. The basis lies in the theory of Vlasov [2] who used warping in defining the torsional cross-section properties. GBT unifies the concept of warping to all modes of displacement from axial deflection and beam bending to torsion and distortion. GBT handles the different degrees of freedom, i.e. extension, bending about principal axes, torsion and distortion as modes which in the first order theory are orthogonal. This means that they are uncoupled and can be analysed separately before their effects are combined with the simple procedure of superposition. In the second order theory the modes are coupled but their orthogonality ensures that individual modes can be separated from the analysis results. The first order GBT can be used in the linear static analysis of prismatic structures, and the second order GBT can be used in the stability or vibration analysis of prismatic structures. In this paper a simple eigenvalue extraction method is presented which allows for obtaining the cross-sectional properties of plate structures in the framework of GBT. The method is based on the eigenfrequency analysis of plane frame beams with the standard finite element method (FEM).

2. THEORY

The total deformation of a prismatic structure can be expressed as sum of the product functions $F(s) \cdot V(x)$ where $F(s)$ expresses the relative cross-sectional deformation, which includes also distorsion, and $V(x)$ (generalised displacement) which is the amplitude function along the x-axis [1]. In Fig. 2.1 a prismatic structure shows the division of displacement functions in the two main directions:

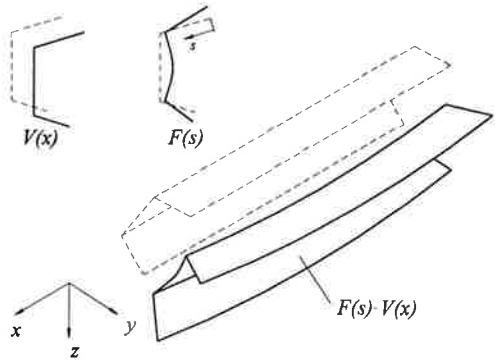


Figure 2.1 *Product statement of displacement functions.*

Both functions must fulfil the boundary conditions of the structure, $F(s)$ transversally and $V(x)$ longitudinally. In the notation of GBT two important points must be defined :

- a forward subscript is used to denote the deformation mode number, e.g. ${}^{k}u(s)$ is the axial displacement function of mode k
- ~ over a symbol denotes a unit value of quantity, e.g. the warping function or a related quantity derived from it.

The three basic displacements of GBT are the axial displacement u, in-plane displacement f_s and out-of-plane displacement f as shown in Fig 2.2:

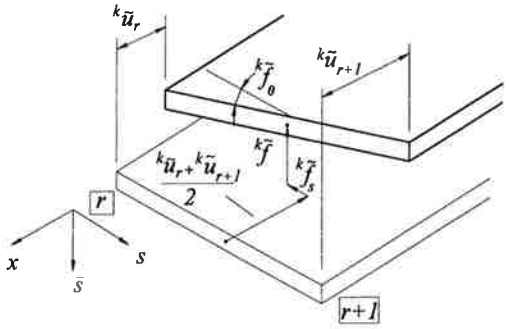


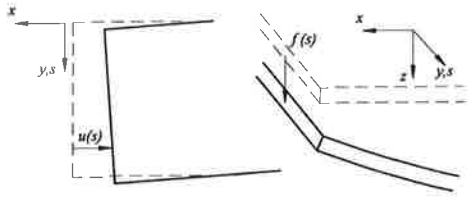
Figure 2.2 *The unit displacement functions* $\tilde{u}, \tilde{f}$ *and* $\tilde{f}_s$ *of a plate strip and the local coordinate system.*

Only the static terms are included in the first phase. At a later stage the second order terms are included in the expression to cover the stability and dynamic analysis. The total derivation of the basic equation of GBT is presented in Ref. [1]. After the procedure, the basic equation of GBT, which is a fourth order differential equation, emerges:

$$E \cdot {}^{k}C \cdot {}^{k}V'''' - G \cdot {}^{k}D \cdot {}^{k}V'' + {}^{k}B \cdot {}^{k}V = {}^{k}q \,. \tag{2.1}$$

All the deformation modes are orthogonal, i.e. independent of each other. This orthogonality can be expressed with two conditions, first for primary warping and second for secondary warping, i.e. plate bending, see Fig.2.3:

Primary warping *Secondary warping*



$$\int_s {}^{i}\tilde{u} \cdot {}^{k}\tilde{u}\, \mathrm{d}s = 0 \quad , i \neq k \qquad\qquad \int_s {}^{i}\ddot{\tilde{f}} \cdot {}^{k}\ddot{\tilde{f}}\, \mathrm{d}s = 0 \quad , i \neq k$$

Figure 2.3 *Primary and secondary warping and their orthogonality conditions.*

When deriving the basic equations the following cross-sectional constants are defined:

$${}^{k}C = \int_A {}^{k}\tilde{u}^2 \mathrm{d}A + \frac{K}{E}\int_s {}^{k}\tilde{f}^2 \mathrm{d}s \tag{2.2a}$$

$$\begin{aligned} {}^{k}D = {}^{k}D_1 - \frac{\nu \cdot E}{G}({}^{k}D_2 + {}^{k}D_2) = \\ \frac{t^3}{3}\int_s {}^{k}\dot{\tilde{f}}^2 \mathrm{d}s - \frac{2 \cdot \nu \cdot K}{G}\int_s {}^{k}\ddot{\tilde{f}} \cdot {}^{k}\tilde{f}\, \mathrm{d}s \end{aligned} \tag{2.2b}$$

$${}^{k}B = K\int_s {}^{k}\ddot{\tilde{f}}^2 \mathrm{d}s, \qquad K = \frac{E \cdot t^3}{12(1-\nu^2)}. \tag{2.2c}$$

Notice the ($\dot{}$) symbol relating to differentiation with respect to the s coordinate. The first two cross-sectional properties above are analogous to the well known section properties of the beam theory. In GBT the first constant ${}^{k}C$ is the warping constant, ${}^{k}D$ is the torsion constant and ${}^{k}B$ is the transverse bending stiffness, which defines the flexibility of the cross-section in distorsion.

3. CROSS SECTIONAL PROPERTIES IN PLATE BENDING ANALYSIS

In plate bending analysis only out-of-plane displacement function f appears as shown in Fig.3.1:

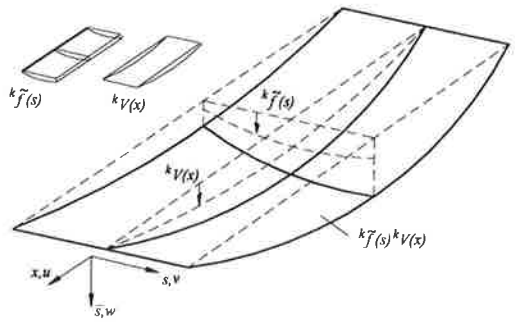


Figure 3.1 *Total deformation of mode k of a plate structure.*

The cross-sectional properties are obtained in the orthogonalisation process which leads to the basic equation, Eq. 2.1. This process can be effectively simplified in the plate bending analysis by using the finite element method (FEM) in obtaining the displacement functions ^kf for each deformation mode and, consequently, for all the required cross section properties. The idea is to use a two-dimensional beam element to describe the cross section displacements. The required stiffness and mass matrices are readily available from standard FEM-textbooks. Mathematically, the analogy is based on the similarity of the equation of motion of a Euler-Bernoulli beam theory

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(EI\frac{\mathrm{d}^2W}{\mathrm{d}x^2}\right)-\omega^2\rho\cdot A\cdot W=0, \tag{3.1}$$

where W is the mode shape, to the non-orthogonal basic equation of GBT, when shear is neglected:

$$E\cdot\overline{C}\cdot\overline{V}''''+\overline{B}\cdot\overline{V}=0. \tag{3.2}$$

The orthogonal cross-section deformation modes U are obtained from Eq. 3.2 by setting an eigenvalue problem

$$\left(E\cdot\overline{C}+\lambda\cdot\overline{B}\right)\cdot U=0. \tag{3.2}$$

Similarly, the mode shapes of Eq. 3.1 are obtained from the generalised eigenvalue problem involving stiffness and mass matrices. These matrices are derived from the interpolation functions between the element nodes which, in the case of a two-node beam element of length b, are cubic polynomials

$$\begin{aligned}
\psi_1(s)&=1-3\left(\frac{s}{b}\right)^2+2\left(\frac{s}{b}\right)^3 & \psi_2(s)&=s\left(1-\frac{s}{b}\right)^2\\
\psi_3(s)&=3\left(\frac{s}{b}\right)^2-2\left(\frac{s}{b}\right)^3 & \psi_4(x)&=\frac{s^2}{b}\left(\frac{s}{b}-1\right)
\end{aligned} \tag{3.3}$$

The stiffness and corresponding so-called consistent mass matrix of a four-degree-of-freedom (two translations and two rotations) plane frame beam element is then

$$K_e = \frac{2EI}{b^3}\begin{bmatrix} 6 & 3b & -6 & 3b \\ 3b & 2b^2 & -3b & b^2 \\ -6 & -3b & 6 & -3b \\ 3b & b^2 & -3b & 2b^2 \end{bmatrix} \qquad M_e = \frac{\rho Ab}{420}\begin{bmatrix} 156 & -22b & 54 & 13b \\ -22b & 4b^2 & -13b & -3b^2 \\ 54 & -13b & 156 & 22b \\ 13b & -3b^2 & 22b & 4b^2 \end{bmatrix}. \quad (3.4)$$

The stiffness and mass properties can be set as $E = \rho = 1$ for the eigenfrequency analysis since the eigenforms are of interest and the eigenvalues only sort the deformation modes into hierarchically ascending order. The cross-section is a slice of the plate with unit width which means that $I = t^3 / 12$ and $A = b$.

The global stiffness and mass matrices are obtained by joining the element stiffness matrices together by adding the nodal co-ordinate stiffness in the corresponding degree of freedom. In a plate structure the resulting matrices are banded since the elements are adjacent to each other. The eigenfrequency analysis or normal mode analysis is based on the generalised eigenvalue problem as (compare Eq. 3.2)

$$\left(\lfloor K_g \rfloor - \lambda \cdot \lfloor M_g \rfloor\right) \cdot \{\delta\} = 0, \quad (3.5)$$

where $\lambda = \omega^2$. The solution of Eq. 3.5 produces the eigenvectors or frequency modes k for each eigenfrequency in the element vectors $\{{}^k\delta_i\} = \{{}^k w_{i,1}\ {}^k\varphi_{i,1}\ {}^k w_{i,2}\ {}^k\varphi_{i,2}\}$. These vectors are the required node displacements for the deformation modes as the total displacement function between two nodes is given by the superposition of the interpolation functions ψ_i (Eq. 3.3) as

$${}^k f_i(s) = \psi_1(s)\,{}^k w_{i,1} + \psi_2(s)\,{}^k\varphi_{i,1} + \psi_3(s)\,{}^k w_{i,2} + \psi_4(s)\,{}^k\varphi_{i,2}. \quad (3.6)$$

The displacement function is used to develop closed form solutions for the cross-sectional properties expressed in Eq. 2.2. The interpolation functions are polynomials which are easily differentiated and integrated. The summing over elements N gives

$${}^kC = \sum_{i=1}^{N} \frac{K_i}{E} \int_0^{b_i} {}^k\tilde{f}_i^{\,2}\,\mathrm{d}s \quad (3.7a)$$

$${}^kD = \sum_{i=1}^{N} \left[\frac{t_i^3}{3} \int_0^{b_i} {}^k\tilde{\tilde{f}}_i^{\,2}\,\mathrm{d}s - \frac{2 \cdot \nu \cdot K_i}{G} \int_0^{b_i} {}^k\tilde{\tilde{f}}_i \cdot {}^k\tilde{f}_i\,\mathrm{d}s \right] \quad (3.7b)$$

$${}^kB = \sum_{i=1}^{N} K_i \int_0^{b_i} {}^k\tilde{\tilde{f}}_i^{\,2}\,\mathrm{d}s \qquad K_i = \frac{E \cdot t_i^3}{12(1-\nu^2)}. \quad (3.7c)$$

4. BASIC PROBLEMS IN PLATE BENDING ANALYSIS

4.1 First Order Generalised Beam Theory

The first order GBT offers a solution method for obtaining the deformations and stresses of prismatic structures with different loading and boundary conditions. Two numerical solution methods are available for solving the basic equation. The first one is the difference method [1] and the other one is the FE-method [3]. In this paper no examples are given for solving problems relating to the first order GBT.

4.2 Second Order Generalised Beam Theory in Stability Analysis

The second-order effect in the context of the GBT means interaction between the distribution of the longitudinal stresses σ_x in the cross-section and the elastic curvature $f''(s)$ of the elements of the cross-section, see [4],[5]. Basically, the 'second-order effect' can be described as the geometrically non-linear behaviour of the structure. The second order theory requires the fulfillment of equilibrium conditions with respect to the deformations of the structure. The virtual work done by the deviating forces can then be directly added to the basic equation of GBT thus resulting the basic equation of the second-order GBT as

$$E \cdot {}^{k}C \cdot {}^{k}V'''' - G \cdot {}^{k}D \cdot {}^{k}V'' + {}^{k}B \cdot {}^{k}V + \sum_{i=1}^{m} \sum_{j=2}^{n} {}^{ijk}\kappa_{\sigma} ({}^{i}W \cdot {}^{j}V')' = {}^{k}q \quad 2 \le k \le n \tag{4.1}$$

where the interaction coefficient is

$${}^{ijk}\kappa_{\sigma} = \frac{1}{{}^{i}C} \int_{s} {}^{i}\tilde{u} ({}^{j}\tilde{f}_{s} \cdot {}^{k}\tilde{f}_{s} + {}^{j}\tilde{f} \cdot {}^{k}\tilde{f})\, t \, ds \tag{4.2}$$

which represents the deviation forces of membrane stresses in the equilibrium. In many bifurcation problems a load is applied which is constant over the length of the member. The derivatives of ${}^{i}W$ are then zero. If the axial stresses are caused by normal force $N \equiv {}^{1}W$ the equation is simplified to the form

$$E \cdot {}^{k}C \cdot {}^{k}V'''' - G \cdot {}^{k}D \cdot {}^{k}V'' + {}^{k}B \cdot {}^{k}V + {}^{1}W \sum_{j=2}^{n} {}^{1jk}\kappa_{\sigma} \cdot {}^{j}V'' = 0 . \tag{4.3}$$

Sinus function is assumed for the displacement function $V(x)$ with a simply supported structure. When inserting ${}^{k}V$ and its derivatives in Eq. 4.3 the following formula is derived

$${}^{1}W(L) = \frac{1}{\sum_{j=2}^{n} {}^{1jk}\kappa_{\sigma}} \left[E \cdot {}^{k}C \cdot \left(\frac{\pi}{L} \right)^{2} + G \cdot {}^{k}D + {}^{k}B \cdot \left(\frac{L}{\pi} \right)^{2} \right] . \tag{4.4}$$

4.3 Second Order Generalised Beam Theory in Frequency Analysis

Dynamic terms are added to the basic equation the same way as in the previous chapter. The model of D'Alembert's forces of inertia can be used in deriving the virtual work done by the dynamic forces, whether they are acting transversely or longitudinally. The deduction in formulas is given in Ref. [6], in this paper only the final additional term is presented. The masses of the structure and the non-structural masses without stiffness are handled equivalently producing the dynamic virtual work as

$$\delta W_{dyn} = \sum_j \left(- {}^{jk}\kappa_{dyn,u} \cdot {}^{j}\ddot{V}'' + {}^{jk}\kappa_{dyn,f} \cdot {}^{j}\ddot{V} \right) \cdot \delta\, {}^{k}V \tag{4.5}$$

where

$${}^{jk}\kappa_{dyn,u} = \int_s \rho \cdot t \cdot {}^{j}\tilde{u} \cdot {}^{k}\tilde{u}\, \mathrm{d}s + \sum_l m_l \cdot {}^{j}\tilde{u}(s_l) \cdot {}^{k}\tilde{u}(s_l) \tag{4.6a}$$

$${}^{jk}\kappa_{dyn,f} = \int_s \rho \cdot t \cdot \left({}^{j}\tilde{f} \cdot {}^{k}\tilde{f} + {}^{j}\tilde{f}_s \cdot {}^{k}\tilde{f}_s \right) \mathrm{d}s + \sum_l m_l \cdot \left({}^{j}\tilde{f}(s_l) \cdot {}^{k}\tilde{f}(s_l) + {}^{j}\tilde{f}_s(s_l) \cdot {}^{k}\tilde{f}_s(s_l) \right). \tag{4.6b}$$

Notice the symbol ($\dot{}$) in this case being derivation with respect to time. The masses m_l are the concentrated masses along the x-axis situated in s_l. There are also deviating forces in shear as well as transverse membrane forces but they are neglected in this context. Now the total system of differential equations can be written in the form

$$E \cdot {}^{k}C \cdot {}^{k}V'''' - G \cdot {}^{k}D \cdot {}^{k}V'' + {}^{k}B \cdot {}^{k}V + \sum_{i=1}^{m}\sum_{j=2}^{n} {}^{ijk}\kappa_{\sigma} ({}^{i}W \cdot {}^{j}V')' + \sum_{j=1}^{n} \left(- {}^{jk}\kappa_{dyn,u} \cdot {}^{j}\ddot{V}'' + {}^{jk}\kappa_{dyn,f} \cdot {}^{j}\ddot{V} \right) = {}^{k}q \quad 2 \le k \le n. \tag{4.7}$$

This equation allows for the determination of the eigenfrequencies of a prismatic structure whether static loads are present or not. In this paper only membrane forces are included in the analysis and, therefore, the right hand side of the equation is set to zero. The stiffness of the structure decreases or increases by the deviating forces due to the membrane loading. When constant axial loading is set on the structure and the analysis is limited to plate bending problems, i.e. $f_s = 0$, Eq. 4.7 can be written in the form

$$E \cdot {}^{k}C \cdot {}^{k}V'''' - G \cdot {}^{k}D \cdot {}^{k}V'' + {}^{k}B \cdot {}^{k}V + {}^{1}W \cdot \sum_{i=1}^{m}\sum_{j=2}^{n} {}^{ijk}\kappa_{\sigma} \cdot {}^{j}V'' + {}^{k}\omega^2 \cdot \sum_{j=1}^{n} {}^{jk}\kappa_{dyn,f} \cdot {}^{j}V = 0 \quad 2 \le k \le n. \tag{4.8}$$

The uncoupled eigenfrequencies of a prismatic plate under membrane static load, tension or compression, can be calculated from this equation. A simple example is calculated and the results are compared to the analytic results from the plate theory in the next chapter.

5. NUMERICAL EXAMPLE

The fundamental or lowest eigenfrequency of a simply supported square isotropic plate structure is analysed. The eigenfrequency is dependent on the direction and magnitude of the membrane end load N:

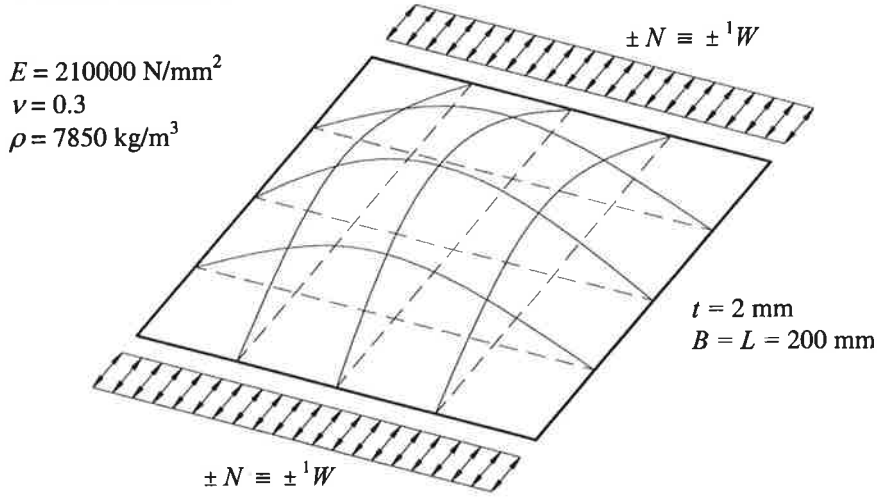


Figure 2.3. *Fundamental or lowest eigenmode of example plate structure.*

Sinus function is assumed for the generalised displacement ${}^kV$ due to the simple supports at the ends of the plate. Then the eigenfrequencies for all modes can be derived from Eq. 4.8 producing the expression

$$ {}^k\omega^2 = \frac{E \cdot {}^kC \cdot \frac{\pi^4}{L^4} + G \cdot {}^kD \cdot \frac{\pi^2}{L^2} + {}^kB + {}^1W \cdot {}^{1kk}K_\sigma \cdot \frac{\pi^2}{L^2}}{{}^{kk}K_{dyn,f}} \quad . \tag{5.1} $$

The natural frequencies in the free vibration of a uniform rectangular plate are, of course, available in numerous books about the dynamic analysis of plate structures. However, much less is written about plate dynamics under membrane stresses. According to Ref. [7] the natural frequencies of a uniform rectangular plate subjected to in-plane loading can be related to the natural frequencies of the same plate without loads by the following expression

$$ \omega^2_{N\neq 0} = \omega^2_{N=0} + \frac{N \cdot J \cdot \pi^2}{\gamma \cdot L^2} \tag{5.2} $$

where γ is the mass per unit area of the plate. The dimensionless coefficient J is a function of the mode number (in this case the first mode) and the boundary conditions applied on the length L. With a square plate $J = 1$.

The comparison between GBT and Eq. 5.2 is presented in Fig. 5.1:

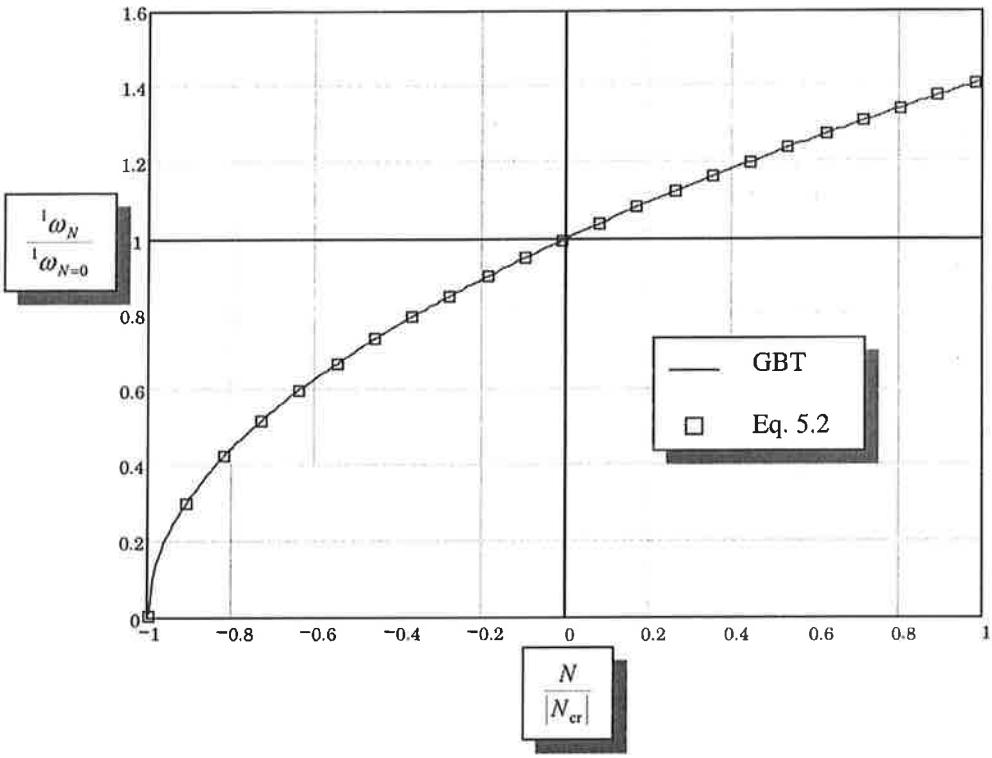


Figure 5.1 *Comparison between GBT and Ref. [7].*

When the positive membrane load increases, the fundamental eigenfrequency increases monotonically. On the other hand, a sufficiently large compressive load N will cause the plate to buckle and the fundamental eigenfrequency goes to zero. Critical compressive loading N_{cr} is the buckling load of the plate without vibration, which can be calculated from Eq. 4.4.

6. CONCLUSIONS

The generalised beam theory offers a transparent and systematic way of analysis for prismatic structures. The generalised eigenvalue problem is used in obtaining the required cross-sectional properties. After this procedure uncoupled ordinary differential equations emerge. In this paper a simplified method based on beam eigenfrequency analysis was presented for the plate bending problems and the proposed method was verified with a numerical example. Further research will be carried out to expand the proposed way of analysis to the general prismatic cross-section as well as to thick plates when shear deformation can not be neglected.

REFERENCES

[1] Schardt, R., Verallgemeinerte Technische Biegetheorie. Springer-Verlag, Berlin, 1989. ISBN 3-540-51339-6.

[2] Vlasov, V.Z., Thin-walled Elastic Beams. Israel program for scientific calculations. Translation of: Tonkostennye uprugie sterzhni, Jerusalem 1961, 493 s.

[3] Davies, J.M. & Leach, P., First-Order Generalised Beam Theory, Thin-Walled Structures 31 (1994) s. 187-220.

[4] Schardt, R., Generalized Beam Theory - An Adequate Method for Coupled Stability Problems, Thin-Walled Structures 19 (1994) s. 161-180.

[5] Davies, J.M. & Leach, P., Second-Order Generalised Beam Theory, Thin-Walled Structures 31 (1994) s. 221-241.

[6] Schardt, R. & Heinz, D., Vibrations of Thin-walled Prismatic Structures Under Simultaneous Static Load Using Generalized Beam Theory, In: Krätzig et al. (eds.), Structural Dynamics 1990 Balkema, Rotterdam. ISBN 90-6191-168-0.

[7] Dickinson, S.M., The Buckling and Frequency of Flexural Vibration of Rectangular Isotropic and Orthotropic Plates Using Rayleigh's Method, Journal of Sound Vibrations 61, 1-8 (1978).

Stability Analysis of Belt Transmission System

Markku Keskiniva
Research Engineer
Sandvik Tamrock Corp./ Technology Center
P.O.Box 100
FIN-33311, Tampere, Finland

ABSTRACT

Belts are classified as "gyroscopic" dynamic systems. The mathematical structure of such a system is entirely analogous to that of companion rotating structures. In this paper the governing equations of motion of the belt is deduced and based on these equations the critical frequencies are determined. This study is applied to the diesel engine/screw compressor-system where the power is transmitted with a toothed belt. The system under consideration has met severe vibration problems, which have eventually lead to premature failures of the belt.

1. INTRODUCTION

One of the Tamrock's drill rig models has had vibration problems with the belt transmission of the screw compressor. The life times of the belts have been considerably short. The belt undergoes excessive vibrations that take place within the wide range of operating speed of the diesel engine. The vibration amplitudes may easily exceed 20 - 30 mm. The vibration is characterised by non-stationary modes, which travel back- and forward within the belt span. These vibrations are considered as a primary reason for the belt failures.

In this paper a theoretical study of the belt vibration is carried out. The governing equation of motion of the belt system will be deduced. Based on these equations the critical operating speeds at where the belt becomes unstable are determined.

The belt forms a complex vibrating system, which is mathematically analogous to a gyroscopic system. The critical frequencies of the belt are dependent on the span length, the belt mass (mass per unit length), the belt tension force and on the belt transport speed.

The torsional vibrations are typical excitations encountered with a combustion engine. The excitation frequencies are equal to the operating speed of the engine and different orders of harmonics. The torsional vibration measurements of the diesel engine have revealed that all the harmonics starting from 0.5 to 6 (at least) are present.

2. THEORY OF BELT VIBRATION

2.1. Belt kinematics

Displacement of a material point of the belt can be given by two uncoupled sets of variables as

$$u(t,x) = \sum_j U_j(t) f_j(x) \quad , \qquad j = 1,2,3,...,N \quad , \tag{1}$$

where $U_j(t)$ are generalized coordinates relating to shape functions $f_j(x)$ and N is the number of shape functions in the basis. In this case natural choice for the shape functions is a sinusoidal basis, since it is kinematically feasible and with a zero belt speed they represent directly the eigenmode solution of the system. In this basis the displacement u takes the form

$$u(t,x) = \sum_j U_j(t) \sin \frac{j\pi}{L} x \quad , \tag{2}$$

where x is a fixed axial coordinate (inertial coordinate system) and L is the span length of the belt shown in Fig. 1. Displacement $u(t,x)$ is specified in Eulerian sense and it measures the instantaneous configuration of the belt. The displacement of the individual material particle u_p is given by

$$u_p(t,x) = \sum_j U_j(t) \sin \frac{j\pi}{L} (x + vt) \quad , \tag{3}$$

where vt-term comprises the transport speed effect of the belt.

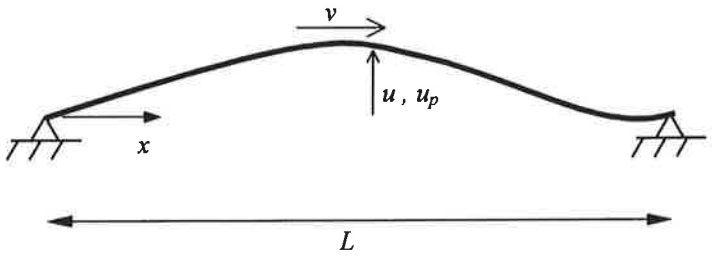


Figure 1. Belt displacement.

The particle velocity and acceleration are obtained by time differentiation to produce

$$\dot{u}_p = \sum_j \dot{U}_j \sin\frac{j\pi}{L}(x+vt) + U_j \frac{j\pi v}{L}\cos\frac{j\pi}{L}(x+vt) \quad , \tag{4a}$$

$$\ddot{u}_p = \sum_j \ddot{U}_j \sin\frac{j\pi}{L}(x+vt) + 2\dot{U}_j \frac{j\pi v}{L}\cos\frac{j\pi}{L}(x+vt) + \\ - U_j\left(\frac{j\pi v}{L}\right)^2 \sin\frac{j\pi}{L}(x+vt) \quad . \tag{4b}$$

2.2. Equations of motion

The derivation of equations of motion is carried out by employing Lagrangean method which stands

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{U}_i}\right) - \frac{\partial T}{\partial U_i} = Q_i \qquad , \qquad i = 1,2,3,...,N \quad , \tag{5}$$

where T and Q_i are the kinetic energy and the generalized force component related to a generalized coordinate U_i, respectively. The kinetic energy is defined as

$$T = \frac{1}{2} m \int_L \dot{u}_p^2 + v^2 dx \quad , \tag{6}$$

where the integral is carried out over the span length of the belt. m is a mass per unit length. The left hand-side of the Lagrangean equation of motion (5) is written as

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{U}_i}\right) - \frac{\partial T}{\partial U_i} = \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{u}_p}\frac{\partial \dot{u}_p}{\partial \dot{U}_i}\right) - \frac{\partial T}{\partial \dot{u}_p}\frac{\partial \dot{u}_p}{\partial U_i} = m\int_L \frac{d}{dt}\left(\dot{u}_p \frac{\partial \dot{u}_p}{\partial \dot{U}_i}\right) - \dot{u}_p \frac{\partial \dot{u}_p}{\partial U_i} dx \quad . \tag{7}$$

The differentiations of $\dot{u}_p$ (4a) with respect to the $\dot{U}_i$ and further with respect to time and $\dot{u}_p$ with respect to the U_i yields

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{U}_i}\right) - \frac{\partial T}{\partial U_i} = m\int_L \ddot{u}_p \sin\frac{i\pi}{L}(x+vt)\, dx \quad . \tag{8}$$

Substituting (4b) for the acceleration gives

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{U}_i}\right)-\frac{\partial T}{\partial U_i}=m\int_L\left[\sum_j \ddot{U}_j \sin\frac{j\pi}{L}(x+vt)+2\dot{U}_j\frac{j\pi v}{L}\cos\frac{j\pi}{L}(x+vt)+\right.$$
$$\left.-U_j\left(\frac{j\pi v}{L}\right)^2\sin\frac{j\pi}{L}(x+vt)\right]\sin\frac{i\pi}{L}(x+vt)\,dx \quad . \tag{9}$$

Since the displacement is fixed with the material particle, the integration must be carried out over the belt span having integration limits as *-vt* and *L-vt* (instead of *0* and *L*). Employing the following results

$$\int_{-vt}^{L-vt}\sin\frac{j\pi}{L}(x+vt)\sin\frac{i\pi}{L}(x+vt)\,dx=$$

$$\int_{-vt}^{L-vt}\cos\frac{j\pi}{L}(x+vt)\cos\frac{i\pi}{L}(x+vt)\,dx=\begin{cases}\dfrac{L}{2} & ,i=j\\ 0 & ,i\neq j\end{cases} \quad , \tag{10a}$$

$$\int_{-vt}^{L-vt}\cos\frac{j\pi}{L}(x+vt)\sin\frac{i\pi}{L}(x+vt)\,dx=\begin{cases}0 & ,i+j \ \ is\ even\\ \dfrac{L}{\pi}\dfrac{2i}{i^2-j^2} & ,i+j \ \ is\ odd\end{cases} \tag{10b}$$

we obtain for Eq. (5)

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{\mathbf{U}}}\right)-\frac{\partial T}{\partial \mathbf{U}}=\mathbf{M}\ddot{\mathbf{U}}+v\mathbf{G}\dot{\mathbf{U}}-v^2\mathbf{K}_G\mathbf{U} \quad . \tag{11}$$

M stands for the generalized mass matrix and is given by

$$\mathbf{M}=\frac{mL}{2}\mathbf{I} \quad , \tag{12}$$

where **I** is an identity matrix. The generalized gyroscopic matrix **G** in (11) is

$$\mathbf{G} = 4m\begin{bmatrix} 0 & -\frac{2}{3} & 0 & & 0 \\ \frac{2}{3} & 0 & -\frac{6}{5} & \cdots & \frac{N}{2^2 - N^2} \\ 0 & \frac{6}{5} & 0 & & 0 \\ & \vdots & & \ddots & \\ 0 & \frac{N}{N^2 - 2^2} & 0 & & 0 \end{bmatrix} \tag{13}$$

and it has a skew-symmetric structure. $\mathbf{K}_G$ is the generalized centrifugal matrix as

$$\mathbf{K}_G = \frac{m\pi^2}{2L}\begin{bmatrix} 1 & 0 & 0 & & 0 \\ 0 & 4 & 0 & \cdots & 0 \\ 0 & 0 & 9 & & 0 \\ & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & N^2 \end{bmatrix} . \tag{14}$$

The generalized force component in the right hand-side in (5) incorporates the internal forces due to the belt tension force and the external forces. The internal generalized forces are derived by employing the principle of virtual work. When the system is given variation in U_i the corresponding increment in work is given

$$\delta W = \sum_i Q_i^{\text{int}} \delta U_i \quad , \tag{15}$$

where Q_i^{int} is the internal generalized force component. On the other hand by giving a variation in u for the differential element (Fig. 2) the same virtual work can be written

$$\delta W = \int_L F \frac{\partial^2 u}{\partial x^2} \delta u \, dx \quad , \tag{16}$$

where F is the belt tension force. In (16) it is assumed that any variation in the tension due to the displacement of the belt is negligible. This is fairly true if the displacement amplitudes are reasonable. When the variation in u is stated in terms of U

$$\delta u = \sum_i \frac{\partial u}{\partial U_i} \delta U_i \quad , \tag{17}$$

the virtual work becomes

$$\delta W = \sum_i \int_L F \frac{\partial^2 u}{\partial x^2} \frac{\partial u}{\partial U_i} \, dx \, \delta U_i \quad . \tag{18}$$

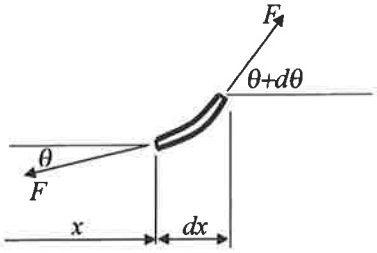


Figure 2. Differential element of the belt under tension force F.

From (15) and (18) it is seen that the internal generalized force component is

$$Q_i^{\text{int}} = \int_L F \frac{\partial^2 u}{\partial x^2} \frac{\partial u}{\partial U_i} dx = F \int_L \left[\sum_j -U_j \left(\frac{j\pi}{L} \right)^2 \sin \frac{j\pi}{L} x \right] \sin \frac{i\pi}{L} x \, dx \quad . \tag{19}$$

Carrying out the integration over the span length one obtains

$$Q_i^{\text{int}} = \begin{cases} \dfrac{Fj^2\pi^2}{2L} U_i \;, & i = j \\ 0 \qquad\quad , & i \neq j \end{cases} \tag{20}$$

or in the matrix form

$$\mathbf{Q}^{\text{int}} = -\frac{F\pi^2}{2L} \begin{bmatrix} 1 & 0 & 0 & & 0 \\ 0 & 4 & 0 & \cdots & 0 \\ 0 & 0 & 9 & & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & N^2 \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \\ U_3 \\ \vdots \\ U_N \end{bmatrix} = -\mathbf{KU} \quad , \tag{21}$$

where **K** is the generalized stiffness matrix. The generalized external forces would be deduced in the same manner but in this paper this phase is omitted because there is no need for them.

The governing equations of motion of the belt are composed from (11) and (21)

$$\mathbf{M}\ddot{\mathbf{U}} + v\mathbf{G}\dot{\mathbf{U}} + (\mathbf{K} - v^2\mathbf{K}_G)\mathbf{U} = \mathbf{Q} \quad , \tag{22}$$

where $\mathbf{Q}$ stands for the generalized external forces. When the belt transport speed is zero, the set of equation is reduced to a form

$$\mathbf{M}\ddot{\mathbf{U}} + \mathbf{K}\mathbf{U} = \mathbf{Q} \ . \tag{23}$$

This set of equations is applicable for systems like a vibrating guitar string where the vibration is characterized by standing waves. However, in the belt vibration the gyroscopic nature of the system induces non-stationary modes. Also the critical frequencies differ from that of stationary systems.

2.3. Critical speeds of the belt

Let there be an external harmonic excitation force acting on the belt. The physical source of excitation is unknown but the excitation frequency is ω. This excitation is denoted as

$$\mathbf{Q} = \mathbf{q}e^{i\omega t} \quad . \tag{24}$$

The solution for (22) is given by

$$\left(-\omega^2\mathbf{M} + i\omega v\mathbf{G} + \mathbf{K} - v^2\mathbf{K}_G\right)\mathbf{U}_q e^{i\omega t} = \mathbf{q}e^{i\omega t} \quad . \tag{25}$$

This system is said to be unstable when the values of $\mathbf{U}$ increases without bounds for any values of $\mathbf{q}$. A sufficient condition is that the coefficient matrix in the left hand-side is singular i.e. its determinant is zero

$$\det\left(-\omega^2\mathbf{M} + i\omega v\mathbf{G} + \mathbf{K} - v^2\mathbf{K}_G\right) = 0 \quad . \tag{26}$$

This eigenvalue problem as it now stands is not directly solvable by standard mathematical procedures. More convenient is to bring it into the following form

$$\begin{bmatrix} \mathbf{K} - v^2\mathbf{K}_g & \mathbf{0} \\ \mathbf{0} & -\mathbf{M} \end{bmatrix}\begin{bmatrix} \mathbf{U} \\ i\omega\mathbf{U} \end{bmatrix}e^{i\omega t} - \omega\begin{bmatrix} -iv\mathbf{G} & -i\mathbf{M} \\ -i\mathbf{M} & \mathbf{0} \end{bmatrix}\begin{bmatrix} \mathbf{U} \\ i\omega\mathbf{U} \end{bmatrix}e^{i\omega t} = \begin{bmatrix} \mathbf{q} \\ \mathbf{0} \end{bmatrix}e^{i\omega t} \quad . \tag{27}$$

Although the problem size is now doubled it is still identical with the original one. By introducing following notations

$$\mathbf{A} = \begin{bmatrix} \mathbf{K} - v^2\mathbf{K}_g & \mathbf{0} \\ \mathbf{0} & -\mathbf{M} \end{bmatrix}, \ \mathbf{B} = \begin{bmatrix} -iv\mathbf{G} & -i\mathbf{M} \\ -i\mathbf{M} & \mathbf{0} \end{bmatrix}, \ \tilde{\mathbf{U}} = \begin{bmatrix} \mathbf{U} \\ i\omega\mathbf{U} \end{bmatrix}, \ \tilde{\mathbf{q}} = \begin{bmatrix} \mathbf{q} \\ \mathbf{0} \end{bmatrix} \tag{28}$$

equation (27) becomes

$$(\mathbf{A}-\omega\mathbf{B})\tilde{\mathbf{U}}e^{i\omega t}=\tilde{\mathbf{q}}e^{i\omega t} \quad . \tag{29}$$

The eigenvalue problem has reached the standard form

$$\det(\mathbf{A}-\omega\mathbf{B})=0 \tag{30}$$

and it is solvable with standard mathematical tools.

3. CRITICAL SPEEDS OF BELT TRANSMISSION SYSTEM

Equation (34) is a function of the excitation frequency ω and of the belt transport speed v. A common approach is to "fix" the belt speed and solve the eigenvalues with respect to ω. By varying the belt speed the critical frequency map over the speed range of interest can be determined. Fig 4. shows an example of a critical frequency map of the belt.

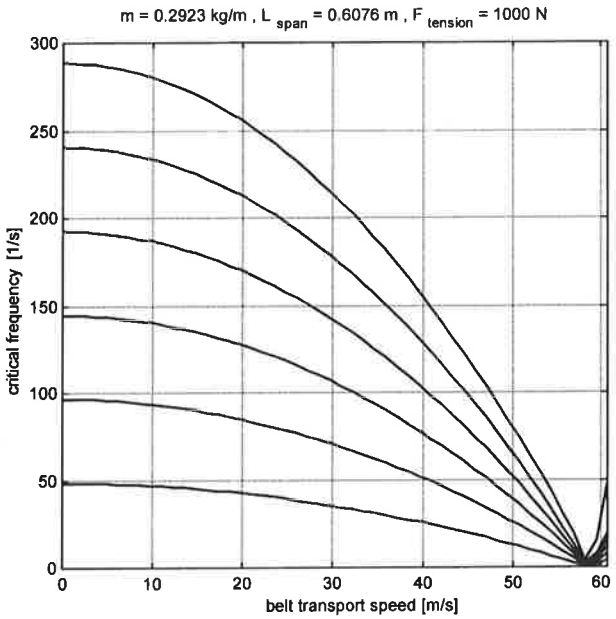


Figure 4. Critical frequency map.

As can be seen, the critical frequencies vary with the belt speed. The belt speed of 58 m/s is a zero-order critical speed. The belt becomes unstable even without any external harmonic excitations. At that speed the centrifugal forces cancel out the stiffness forces and, thus, the net stiffness of the system is zero.

In order to investigate the critical operating speeds of the belt power transmission system it is practical to use a Campbell diagram in where the critical frequencies as well as the excitation frequencies are plotted in the same diagram. For example with the combustion engine one can anticipate excitations having frequencies equal to the operating speed and different orders of harmonics. Strictly speaking, the torsional vibration does not represent a real external excitation. It is rather a parametric excitation since the torsional vibration of the driver pulley shows up as disturbances in the stiffness terms due to variation in the belt tension force. This will give some extra complexity in the stability analysis. However, in this context we will use somewhat simplified approach and treat the torsional vibrations as external excitations.

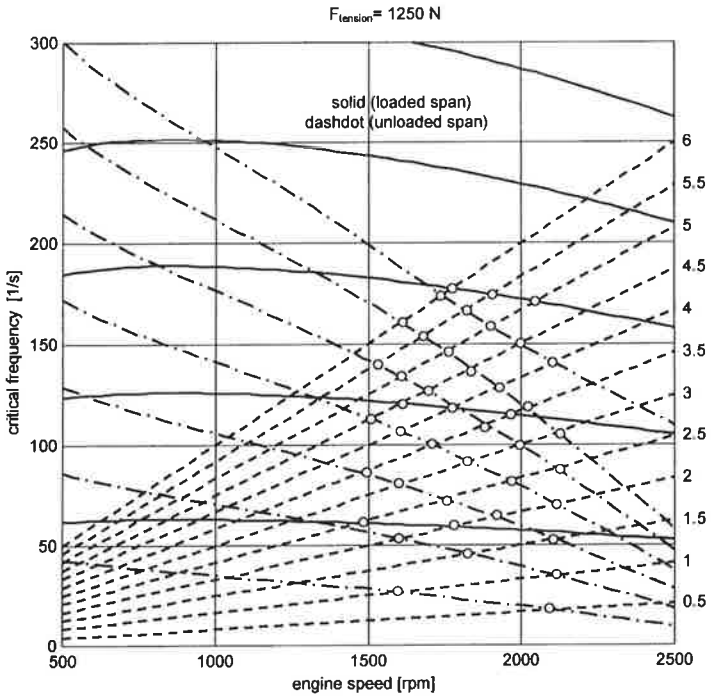


Figure 5. Campbell diagram of the belt transmission system. The output power varies linearly with the engine speed so that 29.1 kW ↔ 1500 rpm and 51.9 kW ↔ 2500 rpm.

In Fig. 5 the Campbell diagram for the belt transmission system is plotted. The diagram includes both belt spans. The static belt tension force is 1250 N and the output power of the screw compressor varies linearly with the engine speed so that 29.1 kW ↔ 1500 rpm and 51.9 kW ↔ 2500 rpm. Along the power variation also the belt tension forces vary. The diameter of the driver pulley is 289.6 mm. The orders of harmonics were verified by the torsional vibration measurements of the crankshaft. The Campbell diagram is plotted with respect to the engine speed. Whenever there is an intersection between critical frequencies and excitation frequency lines (denoted with circles) this point is a critical speed. As seen

in Fig. 5 a great number of critical speeds can be found within the operating range (1500 rpm-->2200 rpm).

From the Fig. 5 it is obvious that no resonance-free operating speed can be found. The main problem is the great number of excitation frequencies. As long as all the harmonics are present the system operates in the vicinity of the critical speeds no matter how the belt tension force or the operating speed are chosen. The only option is to remove the excitation harmonics. Isolating the driver pulley from the crankshaft using a flexible coupling can do this. Were the coupling flexible enough, all the high excitation frequencies could be filtered out. In practice, the flexibility of the coupling is limited by the power transmission capability: a very flexible coupling tends to have poor power transmission capability. The practice has shown that the cut-off frequency is roughly equal to 0.5 harmonic. Still, this is fairly enough, since the belt tension force can then be easily adjusted so that the all the critical frequencies are higher than the remaining excitation frequencies.

4. CONLUSIONS

In this paper the equations of motion for the belt transmission system were derived and the critical frequencies were determined. It is shown that the critical frequencies are dependent on the span length, the belt mass (mass per unit length), the belt tension force and on the belt speed. From the Campbell diagram it was found that the belt transmission system under consideration have large number of critical speeds within the operating range of the diesel engine. This explains the excessive vibration encountered. From the same diagram it was also concluded that no resonance-free system can be found. The only option left is to remove the critical excitation harmonics. This can be done with a flexible coupling between the driver pulley and the crankshaft of the diesel engine. The laboratory tests have confirmed this and the field tests have shown dramatic improvement in life times of the belt with this concept.

REFERENCES

Wickert, J. A., Transient Vibration of Gyroscopic Systems With Unsteady Superposed Motion, Journal of Sound and Vibration, Volume 195(5), 1996, pp. 797-807.

NOMENCLATURE

F	belt tension force	U_i, $\mathbf{U}$	generalized coordinates
$\mathbf{G}$	generalized gyroscopic matrix	T	kinetic energy
$\mathbf{K}$	generalized stiffness matrix	f	shape function
$\mathbf{K}_G$	generalized centrifugal matrix	u	displacement
L	span length	u_p	displacement of material particle
$\mathbf{M}$	generalized mass matrix	v	belt transport speed
Q_i, $\mathbf{Q}$, $\mathbf{q}$	generalized forces	x	axial coordinate
		ω	angular frequency

EPÄLINEAARISEN AKSIAALISESTI LIIKKUVAN LANGAN JA PALKIN DYNAAMISEN ONGELMAN FORMULAATIOSTA

H.KOIVUROVA
Teknillisen mekaniikan laboratorio
Oulun yliopisto, Konetekniikan osasto
PL 4200
90014 Oulun yliopisto

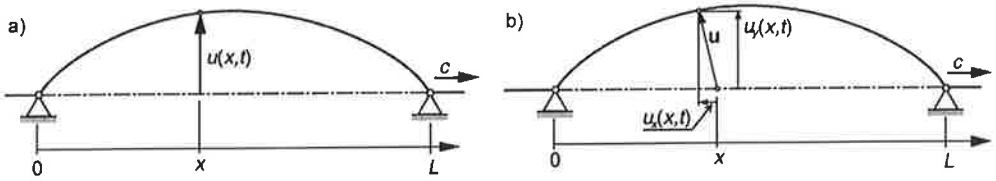
TIIVISTELMÄ

Tässä artikkelissa tarkastellaan ja pyritään selventämään epälineaarisen aksiaalisessa kuljetusliikkeessä olevan rakenteen formulaatiota ja kinemaattista esitystä. Klassisessa aksiaalisesti liikkuvan langan ja palkin tarkasteluissa liikkeen esitys perustuu joko Eulerin tai yhdistettyyn Euler-Lagrangen esitystapaan. Yhdistetty esitys johtaa yksinkertaisempiin yhtälöihin, mutta vaatii laajempia rajoituksia. Näitä rajoituksia ja itse formulaatiota käsitellään ja pyritään selittämään käyttämällä ns. ALE (arbitrary Lagrangian Eulerian) menetelmää.

1. JOHDANTO

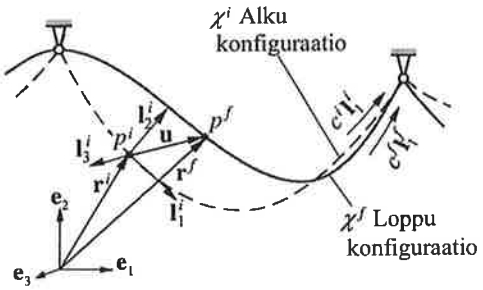
Aksiaalisesti liikkuvan materiaalin ongelmat tarkastelevat sellaisten rakenteiden dynaamista vastetta, värähtelyjä ja stabiilisuutta, jotka ovat jatkuvassa liikkeessä värähtelysuuntaa vastaan kohtisuorassa suunnassa. Ongelmakenttä itsessään on laaja ja koskettaa hyvin erilaisia mekaanisia systeemiä kuten paperikoneet, painokoneet, tekstiilikuidun kelaus, kalvon ja paperin rullaus, magneettinauha-asemat, ketju- ja hihnakäytöt jne. Katsauksia aksiaalisesti liikkuvan materiaalin tutkimuksesta on esitelty mm. Wickertin ja Moten [1] sekä Arbaten [2] toimesta.

Epälineaarisessa aksiaalisesti liikkuvan materiaalin tarkasteluissa liikkeen esitystapa jakautuu selkeästi kahteen eri käsittelytapaan; Eulerin formulaatioon [3,4 ja 5] ja yhdistettyyn Euler-Lagrangen formulaatioon [6, 7 ja 8]. Näistä ensimmäisessä liike kuvataan spatiaalisella muuttujalla $u(x,t)$, joka kuvaa asemassa x olevan materiaalipartikkelin poikittaissiirtymän. (kts kuva 1a). Yhdistetyssä esitystavassa määritellään lisäksi pitkittäissiirtymä u_x (kts kuva 1b), joka Moten [6] mukaan on aksiaalinen siirtymä nopeudella c liikkuvan koordinaatiston suhteen kun taas poikittaiskomponentti u_y mitataan kiinteän koordinaatiston suhteen. Myöhemmin Wickert [8] esitti molempien siirtymäkomponenttien niin poikittaisen kuin pitkittäisen olevan esityksiä spatiaalisessa koordinaatistossa. Thurman ja Mote [7] eivät maininneet omassa tarkastelussaan liikkeen esitystavasta yhtään mitään.



Kuva 1. Aksiaalisesti liikkuvan langan formulaatiot. a) Eulerin esitystapa [4]. b) yhdistetty Euler-Lagrange esitystapa [8].

Edellä esitetyt tarkastelut käsittelivät klassisia kuljetusliikkeessä olevia "kireän" langan tai suoran palkin systeemejä, jotka ovat esikiristyksen vuoksi täysin suoria alkutilassaan. Yhdistettyä formulaatiota on käytetty myös alkutilassaan kaarevalle (ns. löysälle) langalle [9]. Tässä esityksessä siirtymä määriteltiin loppukonfiguraation ja alkukonfiguraation pisteiden paikkavektoreiden erotusvektoriksi (kts. kuva 2). Tekijöiden mukaan se kuvaa loppukonfiguraation kolmiulotteista liikettä ja eroaa lankapartikkelin liikkeestä partikkelin kuljetusliikkeen $c^f \mathbf{l}_1^f$ verran.



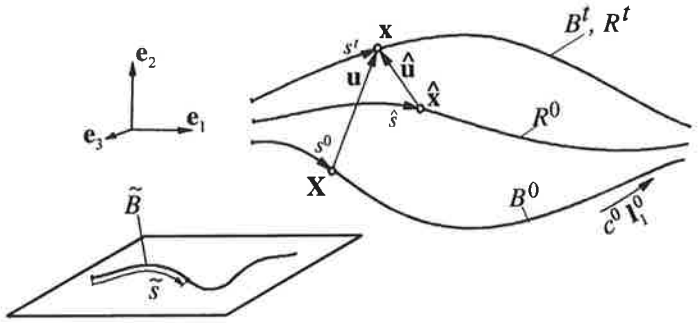
Kuva 2. Alkutilassaan kaarevan aksiaalisesti liikkuvan langan liikkeen esitys Perkinsin ja Moten [9] mukaan.

Varsinaisesti yhdistetty Euler-Lagrangen esitystapa selitettiin vasta lähteessä [10], jossa edellä mainittujen suoran ja kaarevan alkutilan tapausten kinemaattiset suureet johdettiin käyttämällä Vu-Quocin ja Lin [11] ja Behdinan *et al.* [12] tapaa formuloida liukuvan palkin ongelma. Esitys selitti liikesuureet ja perusteli liikeyhtälöiden johdot, mutta se johti myös varsin monimutkaiseen ja monia rajoituksia sisältävään ajatusrakennelmaan, joilla pyrittiin perustelemaan partikkelien ja "konfiguraation" pisteiden välistä assosiaatiota.

Tämän artikkelin tarkoituksena on esittää aksiaalisesti liikkuvan langan ja palkin ongelmien yleinen formulaatio käyttäen ns. "mielivaltaista" Lagrange-Euler menetelmää (arbitrary Lagrangian Eulerian, ALE). ALE esitystapa perustuu mielivaltaiseen referenssitilan liikkeeseen, jonka ei tarvitse riippua materiaalisesta liikkeestä mitenkään. ALEa käytetään erityisesti differenssi- ja elementtimenetelmien yhteydessä virtaus-, nesterakenne vuorovaikutus- sekä muokkausongelmien yhteydessä [13,14,15].

2. AKSIAALISESTI LIIKKUVAN LANGAN ALE FORMULAATIO

ALE:n esitystapaa on käytetty laajalti jo 80-luvun alusta ja sen formuloinnista löytyy kattavia esityksiä. Tarkastellaan ALE esitystavan kinemaattista esitystä yleiselle aksiaalisesti liikkuvalle yksidimensioiselle kontinuumille lähteisiin [13, 14, 15] perustuen ja sovelletaan saatuja yhtälöitä alkutilassaan suoralle ja kaarevalle tapaukselle.



Kuva 3. Kaavamainen esitys domeeneista ja kinemaattista suureista.

Tarkastellaan kuvan 3 mukaista aksiaalisesti liikkuvaa lankasysteemiä. Oletetaan, että
- lanka on lineaarisesti kimmoisaa materiaalia, ja
- on olemassa luonnollinen tila $\tilde{B}$, jossa lanka on jännityksetön.

Oletetaan, että liikkuvan langan materiaalipartikkelit varaavat avaruuden alueen B^0 ajanhetkellä $t = 0$. Merkitään partikkeleiden liikettä kuvauksella φ siten, että B^0:n kuvaa hetkellä $t > 0$ merkitään B^t:llä ja $\mathbf{X} \in B^0$:n kuvaa merkitään $\mathbf{x}$:llä, eli

$$\mathbf{x} = \varphi(\mathbf{X},t) = \mathbf{X} + \mathbf{u}(\mathbf{X},t)\,. \tag{1}$$

Yllä esitetty kuvaus edustaa Lagrangen esitystapaa, jossa kontinuumin liikettä seurataan materiaalisen koordinaatin $\mathbf{X}$ suhteen, kiinnittäen tarkastelu materiaalipartikkeliin. Materiaalinen siirtymä ja nopeus voidaan määritellä tällöin seuraavasti:

$$\mathbf{u} = \mathbf{u}(\mathbf{X},t) = \mathbf{x} - \mathbf{X} \tag{2a}$$

$$\mathbf{v} = \frac{\partial}{\partial t}\varphi(\mathbf{X},t)\Big|_{\mathbf{X}} = \frac{\partial}{\partial t}\mathbf{u}\Big|_{\mathbf{X}} = \dot{\mathbf{u}}\,. \tag{2b}$$

Yhtälössä (2) symbolilla $|_{\mathbf{X}}$ tarkoitetaan osittaisderivoinnin suorittamista pitäen $\mathbf{X}$ vakiona. Kun yhtälön (1) esittämä kuvaus oletetaan jatkuvaksi ja yksikäsitteiseksi, voidaan sille määritellä käänteiskuvaus

$$\mathbf{X} = \varphi^{-1}(\mathbf{x},t), \tag{3}$$

joka edustaa Eulerin esitystapaa (spatiaalista esitystapaa). Tässä esitystavassa liikettä seurataan spatiaalikoordinaatin $\mathbf{x}$ suhteen, kiinnittäen tarkastelu avaruuden pisteisiin.

ALE esitystavassa tulee vielä määritellä kolmas ns. referentiaalinen domeeni R, joka ajanhetkellä $t = 0$ varaa alueen R^0. Tällä referentiaalisella "kappaleella" on oma riippumaton liikkeensä, joka yhdistyy materiaaliseen kappaleeseen B', jollakin hetkellä $t > 0$, jolloin $R' \equiv B'$. Referenssidomeenin liike voidaan esittää (1):sen tavoin kuvauksena $\hat{\varphi}$ siten, että $\mathbf{x}$ on $\hat{\mathbf{x}} \in R^0$ kuva, eli

$$\mathbf{x} = \hat{\varphi}(\hat{\mathbf{x}},t) = \hat{\mathbf{x}} + \hat{\mathbf{u}}(\hat{\mathbf{x}},t), \tag{4}$$

ALE esitystavassa liikettä seurataan koordinaatin $\hat{\mathbf{x}}$ suhteen ja tarkastelu kiinnitetään domeenin R. Referentiaalinen siirtymä ja nopeus voidaan määritellä seuraavasti:

$$\begin{aligned} \hat{\mathbf{u}}(\hat{\mathbf{x}},t) &= \mathbf{x} - \hat{\mathbf{x}} \\ \hat{\mathbf{v}}(\hat{\mathbf{x}},t) &= \frac{\partial}{\partial t}\hat{\varphi}(\hat{\mathbf{x}},t)\Big|_{\hat{\mathbf{x}}} = \frac{\partial}{\partial t}\hat{\mathbf{x}}\Big|_{\hat{\mathbf{x}}} = \hat{\mathbf{x}}_{,t} = \hat{\mathbf{u}}_{,t} \quad , \end{aligned} \tag{5}$$

missä symbolilla $|_{\hat{\mathbf{x}}}$ tarkoitetaan osittaisderivoinnin suorittamista pitäen $\hat{\mathbf{x}}$ vakiona. Yleisesti ajateltuna liike $\hat{\varphi}$ voi olla mielivaltaista ja referenssiesitystavan hyödyllisyys riippuu sitä, miten tämä liike valitaan. Kiinnittämällä $\hat{\varphi}$ sellaiseksi, että kuvaus $\mathbf{x}$:stä $\hat{\mathbf{x}}$:ään on yksikäsitteinen voidaan kuvaukset yhdistämällä saadaan kolmas kuvaus

$$\hat{\mathbf{x}} = \hat{\varphi}^{-1}(\mathbf{x},t) = \hat{\varphi}^{-1}(\varphi(\mathbf{X},t),t) = \psi(\mathbf{X},t). \tag{6}$$

2.1 Materiaalinen aikaderivaatta ja konvektiivinen nopeus

ALE esitystavassa sekä materiaalinen että spatiaalinen domeeni ovat yleisessä tapauksessa liikkeessä ja tarkastelu kiinnittyy referenssidomeenin R, joten materiaaliset aikaderivaatat on mukavinta esittää referentiaalisessa muodossa. Tarkastellaan fysikaalisen ominaisuuden, f, eri kuvauksilla saatuja esityksiä

$$f = f(\mathbf{x},t) = f(\mathbf{X},t) = f(\hat{\mathbf{x}},t). \tag{7}$$

f:n materiaaliderivaatat voidaan kirjoittaa

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \dot{f} \qquad \text{(muuttujat } \mathbf{X},t\text{)}, \tag{8a}$$

$$= \frac{\partial f}{\partial t}\Big|_{\mathbf{x}} + \frac{\partial f}{\partial x_i}\frac{\partial x_i}{\partial t}\Big|_{\mathbf{x}} = \frac{\partial f}{\partial t}\Big|_{\mathbf{x}} + \mathbf{v}\cdot\nabla_{\mathbf{x}} f \qquad \text{(muuttujat } \mathbf{x},t\text{)}, \tag{8b}$$

$$= \frac{\partial f}{\partial t}\Big|_{\hat{\mathbf{x}}} + \frac{\partial f}{\partial \hat{x}_i}\frac{\partial \hat{x}_i}{\partial t}\Big|_{\mathbf{x}} = f_{,t} + \hat{\mathbf{c}}\cdot\nabla_{\hat{\mathbf{x}}} f \qquad \text{(muuttujat } \hat{\mathbf{x}},t\text{)}, \tag{8c}$$

missä $\hat{\mathbf{c}} = (\partial\hat{\mathbf{x}}/\partial t)|_{\mathbf{X}}$ on materiaalipartikkelin nopeus referentiaalikoordinaatin suhteen ja alaindeksi i on standardi suorakulmaisen karteesisen koordinaatiston indeksimerkintä. Muutoinkin tässä esityksessä alaindeksit tarkoittavat tensorikomponentin indeksimerkintää noudattaen summaussääntöä. Jos nyt tarkasteltavaksi fysikaaliseksi ominaisuudeksi valitaan

spatiaalikoordinaatti **x**, niin yhtälöt (7) ja (8c) johtavat

$$\mathbf{x} = \boldsymbol{\varphi}(\mathbf{X}, t) = \hat{\boldsymbol{\varphi}}(\hat{\mathbf{x}}, t) \tag{9}$$

ja

$$\mathbf{v} = \frac{\partial \hat{\boldsymbol{\varphi}}}{\partial t}\bigg|_{\hat{\mathbf{x}}} + \frac{\partial \hat{\boldsymbol{\varphi}}}{\partial \hat{x}_i}\frac{\partial \hat{x}_i}{\partial t}\bigg|_{\mathbf{X}} = \hat{\mathbf{v}} + \frac{\partial \mathbf{x}}{\partial \hat{x}_i}\frac{\partial \hat{x}_i}{\partial t}\bigg|_{\mathbf{X}} = \hat{\mathbf{v}} + \hat{\mathbf{c}} \cdot \nabla_{\hat{\mathbf{x}}}\, \mathbf{x}\,. \tag{10}$$

Jälkimmäisestä yhtälöstä voidaan vielä määritellä ns. konvektiivinen nopeus

$$\mathbf{c}' = \frac{\partial \mathbf{x}}{\partial \hat{x}_i}\frac{\partial \hat{x}_i}{\partial t}\bigg|_{\mathbf{X}} = \hat{\mathbf{c}} \cdot \nabla_{\hat{\mathbf{x}}}\, \mathbf{x} = \mathbf{v} - \hat{\mathbf{v}} \tag{11}$$

2.2 Massan säilyminen

Referenssidomeenin R^0 kokonaismassa M voidaan muuntaa eri koordinaattien suhteen klassisella muunnoksella

$$M = \int_L \hat{\rho}\hat{A}\,\mathrm{d}\hat{s} = \int_{L'} \rho' A'\,\mathrm{d}s' = \int_{L^0} \rho^0 A^0\,\mathrm{d}s^0\,, \tag{12}$$

missä

$$\hat{\rho}(\hat{\mathbf{x}}, t) = \hat{J}\rho'(\mathbf{x}, t)\,, \qquad \rho^0(\mathbf{X}, t) = J\rho'(\mathbf{x}, t)\,, \qquad \hat{\rho}(\hat{\mathbf{x}}, t) = J^*\rho^0(\mathbf{X}, t)\,, \tag{13a}$$

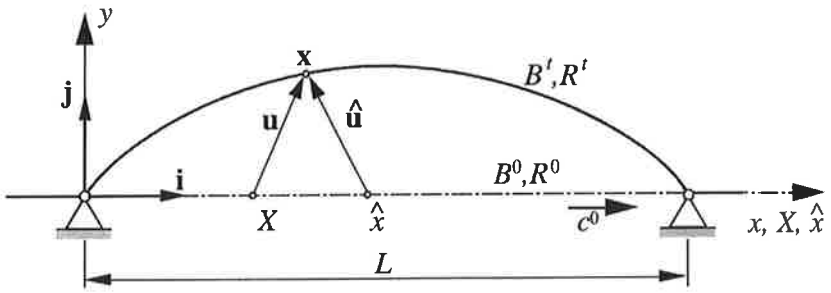
$$\hat{J} = \det\left[\frac{\partial x_i}{\partial \hat{x}_j}\right] = \frac{\mathrm{d}s'}{\mathrm{d}\hat{s}}\,, \qquad \hat{J} = \det\left[\frac{\partial x_i}{\partial X_j}\right] = \frac{\mathrm{d}s'}{\mathrm{d}s^0}\,, \qquad J^* = \det\left[\frac{\partial X_i}{\partial \hat{x}_j}\right] = \frac{\mathrm{d}s^0}{\mathrm{d}\hat{s}}\,. \tag{13b}$$

3. ALKUTILASSAAN SUORA AKSIAALISESTI LIIKKUVA LANKA

Klassisilla aksiaalisesti liikkuvien rakenteiden probleemoilla tarkoitetaan yleensä "kireän" langan tai suoran palkin systeemejä, jotka ovat jännitetty suoriksi kahden silmukan väliin esikiristyksellä T^0 alkutilassaan B^0 (kts. kuva 4). Lisäksi oletetaan, että materiaalipartikkelit liikkuvat suorassa alkutilassa vakiokuljetusnopeudella c^0, jolloin alkutila on myös vakiotila (dynaaminen tasapainotila).

Materiaalidomeeniksi B^0 valitaan 2. kappaleen esityksen tapaan alkutila ajanhetkellä $t = 0$, jossa materiaalipartikkelit identifioidaan. Referenssidomeeniksi R^0 valitaan sama suora alkutila, joten referenssidomeeni R^0 ja materiaalidomeeni B^0 täyttävät saman avaruuden ajanhetkellä $t = 0$. Koska tila on suora voidaan molemmat sekä B^0 että R^0 esittää yhdellä koordinaatilla X ja $\hat{x}$. Spatiaalikoordinaatti esitetään kaksidimenioisessa avaruudessa $\mathbf{x} = (x, y)$. Kuvaus (1) yksinkertaistuu nyt muotoon

$$\mathbf{x} = X\mathbf{i} + \mathbf{u}(X, t) = \hat{x}\mathbf{i} + \hat{\mathbf{u}}(\hat{x}, t)\,. \tag{14}$$



Kuva 4. Kaavamainen esitys domeeneista ja kinemaattisista suureista "kireän" langan tapauksessa.

Referenssi- ja materiaalidomeenin välinen kuvaus (6) on nyt sama kuin Galileomuunnoksessa eli

$$\hat{x} = X + c^0 t\,. \tag{15}$$

Partikkelin nopeus referenssikoordinaatin $\hat{x}$ suhteen on

$$\hat{c} = \left.\frac{\partial \hat{x}}{\partial t}\right|_X = c^0 \tag{16}$$

ja konvektiivinen nopeus $\mathbf{c}^t$ on yhtälöiden (11), (14) ja (16) perusteella

$$\mathbf{c}^t = \hat{\mathbf{c}} \cdot \nabla_{\hat{\mathbf{x}}}\, \mathbf{x} = \hat{\mathbf{c}} \cdot \frac{\partial \mathbf{x}}{\partial \hat{x}} = c^0 \left(\mathbf{i} + \hat{\mathbf{u}}_{,\hat{x}}\right) = \mathbf{v} - \hat{\mathbf{v}}\,. \tag{17}$$

$\mathbf{c}^t$ edustaa partikkelin kuljetusnopeutta referenssidomeenin R^t pisteen $\mathbf{x} = \hat{\varphi}(\hat{\mathbf{x}}, t)$ suhteen. Materiaalipartikkelin nopeus voidaan edellisen ja yhtälön (5) perusteella kirjoittaa muotoon

$$\mathbf{v} = \hat{\mathbf{u}}_{,t} + c^0 \left(\mathbf{i} + \hat{\mathbf{u}}_{,\hat{x}}\right) \tag{18}$$

ja materiaalinen aikaderivaatta (8c) saadaan muotoon

$$\dot{f} = f_{,t} + c^0 f_{,\hat{x}}\,. \tag{19}$$

Liikeyhtälön johtamiseen voidaan lähteä joko Eulerin tai Lagrangen esityksestä. Leissa ja Saad [16] tarkastelivat ilman kuljetusliikettä värähtelevän langan suurten siirtymien liikeyhtälöä kuvan 5a mukaiselle lankaelementille ja kirjoittivat sen muotoon

$$\begin{aligned} \left(T^t \cos\theta\right)_{,X} &= \rho^0 A^0 \ddot{u}_x \\ \left(T^t \sin\theta\right)_{,X} &= \rho^0 A^0 \ddot{u}_y\,, \end{aligned} \tag{20}$$

missä kulman θ kosiniksi ja siniksi saadaan

$$\cos\theta = \frac{1 + u_{x,X}}{\sqrt{\left(1 + u_{x,X}\right)^2 + u_{y,X}^2}}, \qquad \sin\theta = \frac{u_{y,X}}{\sqrt{\left(1 + u_{x,X}\right)^2 + u_{y,X}^2}} \tag{21}$$

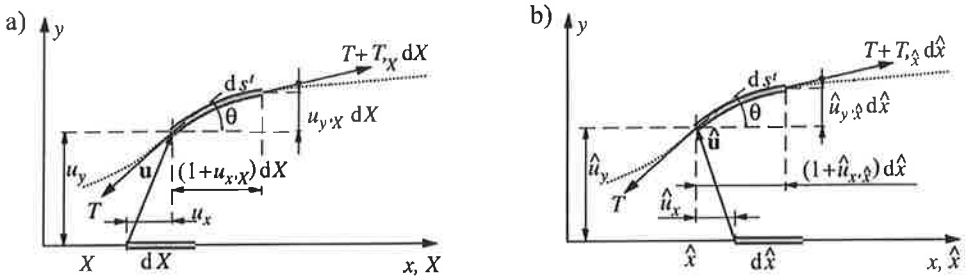
ja lankavoimaksi

$$T' = T^0 + EA^0(\mathrm{d}s - \mathrm{d}X)/\mathrm{d}X = T^0 + EA^0\left[\sqrt{(1 + u_{x',X})^2 + u_{y',X}^2} - 1\right]. \tag{22}$$

Näitä Leissa ja Saad [16] Lagrangen esitystä edustavia yhtälöitä voidaan käyttää lähtökohtana referentiaalisen liikeyhtälön muodostamiselle. Kun nyt huomioidaan, että yhtälön (15) seurauksena materiaalidomeenin B^0 ja referenssidomeenin R^0 välinen Jakobiani $J^* = \det[\partial\hat{x}/\partial X] = 1$, saadaan materiaalidomeenissa vaikuttava lankavoima T^0 yhtäsuureksi kuin referenssidomeenissa R^0 vaikuttava $\hat{T}$. Lisäksi yhtälön (13b) perusteella $\rho^0 = \hat{\rho}$ ja massan säilymisen yhtälön (12) perusteella $A^0 = \hat{A}$. Muodostamalla sitten kuvan 5b mukaiselle referentiaalisen siirtymän kokevalle lankaelementille liikeyhtälö yhtälön (20) tapaan ja huomioimalla materiaalinen aikaderivaatta (19) sekä materiaalisen ja referentiaalisen domeenin suureiden esitetyt yhtäsuuruudet saadaan liikeyhtälöksi

$$\begin{aligned}
&EA^0\hat{u}_{y,\hat{x}\hat{x}} - (EA^0 - T^0)\frac{(1+\hat{u}_{x,\hat{x}})^2\hat{u}_{y,\hat{x}\hat{x}} - \hat{u}_{y,\hat{x}}(1+\hat{u}_{x,\hat{x}})\hat{u}_{x,\hat{x}\hat{x}}}{\left[(1+\hat{u}_{x,\hat{x}})^2 + \hat{u}_{y,\hat{x}}^2\right]^{3/2}} = \rho^0 A^0\left(\hat{u}_{y,tt} + 2c^0\hat{u}_{y,\hat{x}t} + (c^0)^2\hat{u}_{y,\hat{x}\hat{x}}\right) \\
&EA^0\hat{u}_{x,\hat{x}\hat{x}} - (EA^0 - T^0)\frac{\hat{u}_{y,\hat{x}}^2\hat{u}_{x,\hat{x}\hat{x}} - \hat{u}_{y,\hat{x}}(1+\hat{u}_{x,\hat{x}})\hat{u}_{y,\hat{x}\hat{x}}}{\left[(1+\hat{u}_{x,\hat{x}})^2 + \hat{u}_{y,\hat{x}}^2\right]^{3/2}} = \rho^0 A^0\left(\hat{u}_{x,tt} + 2c^0\hat{u}_{x,\hat{x}t} + (c^0)^2\hat{u}_{x,\hat{x}\hat{x}}\right).
\end{aligned} \tag{23}$$

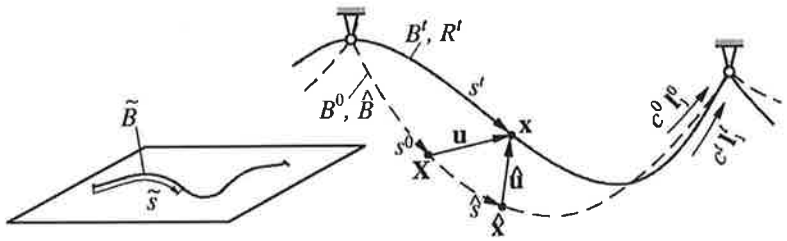
Liikeyhtälö on taivutusjäykkyydestä johtuvaa termiä lukuun ottamatta sama kuin Thurmanin ja Moten [7] esittämä.



Kuva 5. Lankaelementin a) materiaalisiirtymä ja b) referentiaalinen siirtymä.

3. ALKUTILASSAAN KAAREVAN AKSIAALISESTI LIIKKUVAN LANGAN ONGELMA

Perkins ja Mote [9] tarkastelivat alkutilassaan kaarevan eli "löysän" kuljetusliikkeessä olevan langan värähtelyjä yhdistetyllä Euler-Lagrange formulaatiolla. Heidän tarkastelussaan lanka kulki kahden eritasossa olevan silmukan kautta siten, että kaareva alkutila on vakiotila (dynaaminen tasapainotila), jonka muoto riippuu ulkoisista voimista ja kuljetusnopeudesta. Liike määriteltiin formulaatiossa konfiguraation siirtymällä, joka eroaa heidän mukaansa materiaalipartikkelin liikkeestä konfiguraation tangentin suuntaisen kuljetusliikkeen verran.



Kuva 6. Kaavamainen esitys domeeneista ja kinemaattisista suureista "löysän" langan tapauksessa.

Materiaalidomeeniksi B^0 valitaan nytkin alkutila ajanhetkellä $t = 0$, jossa materiaalipartikkelit identifioidaan. Referenssidomeeniksi R^0 valitaan sama alkutila, joten referenssidomeni R^0 ja materiaalidomeeni B^0 täyttävät saman avaruuden ajanhetkellä $t = 0$. Alkutilassa partikkeleiden nopeus on $\mathbf{c}^0 = c^0 \mathbf{l}_1^0$ mikä on yhtäsuuri kuin partikkelin nopeus referenssikoordinaatin $\hat{x}$ suhteen eli nyt

$$\hat{\mathbf{c}} = \left.\frac{\partial \hat{\mathbf{x}}}{\partial t}\right|_{\mathbf{x}} = \mathbf{c}^0(\hat{\mathbf{x}}) = c^0 \hat{\mathbf{l}}_1 , \tag{24}$$

missä $\hat{\mathbf{l}}_1 = \partial \hat{\mathbf{x}} / \partial \hat{s}$ on tangentin suuntainen yksikkövektori. Konvektiivinen nopeus eli materiaalipartikkelin nopeus referenssidomeenin R^t pisteen $\mathbf{x} = \hat{\varphi}(\hat{\mathbf{x}}, t)$ suhteen on yhtälöiden (11) ja (24) perusteella

$$\mathbf{c}^t = \hat{\mathbf{c}} \cdot \nabla_{\hat{\mathbf{x}}} \mathbf{x} = \hat{c}\, \hat{\mathbf{l}}_1 \cdot \frac{\partial \mathbf{x}}{\partial \hat{s}} \hat{\mathbf{l}}_1 = \hat{c} \frac{\partial \mathbf{x}}{\partial \hat{s}} = \hat{c} \frac{\mathrm{d}s^t}{\mathrm{d}\hat{s}} \mathbf{l}_1^t , \tag{25}$$

missä $\mathbf{l}_1^t$ on tangentin suuntainen yksikkövektori spatiaalidomeenissa ja se on

$$\mathbf{l}_1^t(s^t) = \frac{\partial \mathbf{x}}{\partial s^t} = \frac{\partial}{\partial \hat{s}} [\hat{\mathbf{x}}(\hat{s}) + \hat{\mathbf{u}}(\hat{s})] \frac{\mathrm{d}\hat{s}}{\mathrm{d}s^t} . \tag{26}$$

Siten konvektiivinen nopeus saadaan Perkinsin ja Moten [9] esittämään muotoon

$$\mathbf{c}^t = \hat{c} \left(\hat{\mathbf{l}}_1 + \hat{\mathbf{u}}_{,s} \right) . \tag{27}$$

Materiaalinen nopeus voidaan kirjoittaa (5), (11), ja (27) avulla

$$\mathbf{v} = \hat{\mathbf{u}}_{,t} + \hat{c} \left(\hat{\mathbf{l}}_1 + \hat{\mathbf{u}}_{,\hat{s}} \right) \tag{28}$$

ja materiaaliseksi aikaderivaattaksi (8c) referenssidomeenin suhteen saadaan

$$\dot{f} = f_{,t} + \hat{\mathbf{c}} \cdot \nabla_{\hat{\mathbf{x}}} f = f_{,t} + \hat{c}\, f_{,\hat{s}} . \tag{29}$$

Lugano *et al.* [17] ovat esittäneet liikeyhtälön johdon vastaavanlaiselle löysän langan tapaukselle ilman kuljetusliikettä ja tätä Lagrangen esitystä edustavaa tarkastelua käytetään lähtökohtana referentiaaliselle formulaatiolle. Lugano et al. [17] käyttivät Hamiltonin periaatetta kirjoittaen ensin kineettisen energian ja kimmoenergian yhtälöt. Kineettiseksi

energiaksi tulee alkutilan B^0 suhteen esitettynä

$$\Pi_k^t = \rho^0 A^0 \int_{L^0} \mathbf{v} \cdot \mathbf{v} \, \mathrm{d}s^0 = \hat{\rho} \hat{A} \int_{\hat{L}} \mathbf{v} \cdot \mathbf{v} \, \mathrm{d}\hat{s}, \tag{30}$$

joka voidaan esitettää myös referenssitilan R^0 suhteen massan säilymisen yhtälön (12) avulla. Kimmoenergian Π_s^t Lagrangen esitys lasketaan summaamalla esijännitetyn alkutilan B^0 kimmoenergiaan Π_s^0 siirtymän $\mathbf{u}$ tuoma lisä. Tämä tehdään hajoittamalla langan Lagrangen venymä seuraavasti. Langan Lagrangen venymä lopputilassa B^t on

$$e^t = \frac{1}{2}\left(\frac{\partial \mathbf{x}}{\partial \tilde{s}} \cdot \frac{\partial \mathbf{x}}{\partial \tilde{s}} - 1 \right), \tag{31}$$

missä $\tilde{s}$ on luonnollisen jännityksettömän tilan $\tilde{B}$ kaarenpituus. Kun nyt huomioidaan yhtälö (1) voidaan venymä kirjoittaa

$$e^t = e^0 + \left[\mathrm{d}s^0 / \mathrm{d}\tilde{s} \right]^2 \varepsilon, \tag{32}$$

missä e^0 on alkutilan B^0 venymä ja

$$\varepsilon = \frac{\partial \mathbf{u}}{\partial s^0} \cdot \frac{\partial \mathbf{x}}{\partial s^0} + \frac{1}{2} \frac{\partial \mathbf{u}}{\partial s^0} \cdot \frac{\partial \mathbf{u}}{\partial s^0}. \tag{33}$$

Langan kimmoenergiaksi Lagrangen esitytavalla tulee siten

$$\Pi_s^t = \Pi_s^0 + \int_{L^0} \left(T^0 \varepsilon + \frac{1}{2} E A^0 \varepsilon^2 \right) \mathrm{d}s^0, \tag{34}$$

missä Π_s^0 on alkutilan B^0 kimmoenergia. Venymän e^t hajoitelma (32) voidaan tehdä myös referenssitilan R^0 suhteen käyttämällä yhtälöä (4), jolloin venymäksi saadaan

$$e^t = \hat{e} + \left[\mathrm{d}\hat{s} / \mathrm{d}\tilde{s} \right]^2 \hat{\varepsilon}, \tag{35}$$

missä $\hat{e}$ on referenssitilan R^0 venymä ja

$$\hat{\varepsilon} = \frac{\partial \hat{\mathbf{u}}}{\partial \hat{s}} \cdot \frac{\partial \hat{\mathbf{x}}}{\partial \hat{s}} + \frac{1}{2} \frac{\partial \hat{\mathbf{u}}}{\partial \hat{s}} \cdot \frac{\partial \hat{\mathbf{u}}}{\partial \hat{s}}. \tag{36}$$

ja potentiaalienergian esitys saadaan muotoon

$$\Pi_s^t = \hat{\Pi}_s + \int_{\hat{L}} \left(\hat{T} \hat{\varepsilon} + \frac{1}{2} E \hat{A} \hat{\varepsilon}^2 \right) \mathrm{d}\hat{s}, \tag{37}$$

Missä $\hat{\Pi}_s$ on referenssitilan potentiaalienergia. Kineettisen energian ja kimmoenergian referentiaaliesitykset (30) ja (37) sekä materiaalipartikkelin nopeus (28) vastaavat Perkinsin ja Moten [9] esittämiä, joten ne johtavat Hamiltonin periaatetta

$$\delta \left\{ \int_{t_1}^{t_2} \left(\Pi_k^t - \Pi_s^t \right) \mathrm{d}t \right\} = 0 \tag{38}$$

käyttämällä Perkinsin ja Moten [9] johtamaan liikeyhtälöön. Yhtälöissä esiintyvät parametrit $\hat{\rho}$ ja $\hat{A}$ sekä lankavoima $\hat{T}$ voidaan esittää alkutilan vastaavien parametrien ρ^0 ja A^0 sekä lankavoiman T^0 avulla kun huomioidaan kuvaus $\mathbf{x} = \hat{\varphi}(\hat{\mathbf{x}}, t)$, sillä ajanhetkellä t=0

$$\hat{\rho}(\hat{\mathbf{x}}) = \rho^t(\varphi(\hat{\mathbf{x}}, 0), 0) = \rho^t(\mathbf{x}, 0) = \rho^0(\hat{\mathbf{x}}). \tag{39}$$

6. PÄÄTELMÄT

Aksiaalisesti liikkuvan rakenteen ongelmien formulointeja on tarkasteltu. Tarkoituksena oli selventää aikaisemmin esitettyjä yhdistetyn Euler-Lagrange esitystavan formulaatioita, jotka eivät määritele ja selitä aihetta kovinkaan paljon. Käyttämällä tunnettua ALE esitystapaa kinemaattiset suureet ja materiaalinen aikaderivaatta niin alkutilassaan suoralle kuin kaarevalle tapaukselle voitiin esittää. Lisäksi liikeyhtälöiden johtaminen lähtien Lagrangen esitystavasta onnistui suoraviivaisesti ilman hankalia olettamuksia, joilla partikkelin ja referenssidomeenin välistä assosiaatiota tulisi perustella.

LÄHTEET

1. Wickert, J. A. ja Mote, C. D. Jr. Current research on the vibration and stability of axially-moving materials. Journal of Shock and Vibration Digest, 20 (1988), 3-13.
2. Arbate, S. Vibrations of belts and belt drives. Mechanisms and Machines Theory, 27 (1992), 645-659.
3. Zaiser, J. N. Nonlinear vibrations of a moving threadline. Ph.D. Dissertation, University of Delaware. 1964
4. Ames W. F., Lee S. Y. ja Zaiser J. N. Nonlinear vibration of a travelling threadline. International Journal of Nonlinear Mechanics, 3 (1968), 449-469.
5. Kim, Y. ja Tabarrok, B. On the nonlinear vibration of travelling string. Journal of The Franklin Institute, 293 (1972), 381-399.
6. Mote, C. D. Jr. On the nonlinear oscillation of an axially moving string. Journal of Applied Mechanics, 33 (1966), 463-464.
7. Thurman, A. L. ja Mote. C. D. Jr. Free, periodic, nonlinear oscillation of an axially moving strip. Journal of Applied Mechanics, 36 (1969), 83-91.
8. Wickert, J. A. Nonlinear vibration of a travelling tensioned beam. International Journal of Nonlinear Mechanics, 27 (1992), 503-517.
9. Perkins, N. C. ja Mote, C. D. Jr. Three-dimensional vibration of travelling elastic cables. Journal of Sound Vibration, 114 (1987), 325-340.
10. Koivurova, H. ja Salonen, E.-M. Comments on non-linear formulations for travelling string and beam problems, Journal of Sound and Vibration, 225 (1999), 845-856.
11. Vu-Quoc, L. ja Li, S. Dynamics of sliding geometrically-exact beams: Large angle maneuver and parametric resonance. Computer Methods in Applied Mechanics and Engineering. 120 (1995), 65-118.
12. Behdinan, K., Stylianou, M. and Tabarrok, B. Dynamics of flexible sliding beams - non-linear analysis, part I: Formulation. Journal of Sound and Vibration 208 (1997), 517-539.
13. Hughes, T. J. R., Liu, W. K. ja Zimmermann, T. K. Lagrangian-Eulerian finite element formulation for incompressible viscous flows, Computer Methods in Applied Mechanics and Engineering, 29 (1981), 329-349.
14. Liu, W. K. ja Gvildys, J. Fluid-structure interaction of tanks with an eccentric corel barrel, Computer Methods in Applied Mechanics and Engineering, 58 (1986), 51-77.
15. Huerta, A. ja Liu, W. K. Viscous Flow with large free surface motion. Computer Methods in Applied Mechanics and Engineering, 69 (1988), 277-324.
16. Leissa, A. W. ja Saad, A. M. Large amplitude vibrations of strings. ASME Journal of Applied Mechanics, 61 (1994), 296-301.
17. Luongo, A., Rega, G. ja Vestroni, F. Planar nonlinear free vibrations of an elastic cable. International Journal of NonLinear Mechanics, 19 (1984), 39-52.

EXPERIMENTAL MODAL ANALYSIS FROM PARTIALLY CORRUPTED AMBIENT VIBRATION DATA

Jyrki Kullaa
Department of Mechanical Engineering
Helsinki University of Technology
P.O. Box 4100
FIN-02015 HUT, Finland
jyrki.kullaa@hut.fi

Timo Tirkkonen
VTT Building Technology
P.O. Box 18071
FIN-02040 VTT, Finland
timo.tirkkonen@vtt.fi

ABSTRACT

An experimental modal analysis of the Hämeensilta Bridge was performed in order to study the effects of the renovation on the vibration performance. The stochastic subspace identification was used using the response data from traffic excitation. The measurement data was corrupted with harmonic components. Fortunately, two additional vibration transducers were used together with the modal accelerometers and these were used to estimate the natural frequencies and damping ratios. The natural modes were estimated from the corrupted data using spectral analysis. The system identification was also performed on the corrupted data resulting in models comprising both structural and spurious modes.

1. INTRODUCTION

The vibrations of the Hämeensilta Bridge in the centre of Tampere have been seen to exceed the recommendations on the sidewalk, causing annoyance to the pedestrians. Vibration measurements were performed in order to study the vibration levels of the bridge deck at different temperatures and during repair. The bridge was renovated in 1997, and the effects of the renovation on the vibration performance were studied with ambient vibration tests. Two types of loading were used: a single heavy vehicle crossing the bridge, and normal traffic. One part of the study was to determine experimentally the modal parameters of the bridge: natural frequencies, damping ratios, and mode shapes.

The dynamic properties of structures are usually measured using forced vibration testing, where both the excitation and the response are measured. However, for large civil engineering structures, this technique may be very difficult and expensive for the following reasons: The natural frequencies may be very low, and to excite them would require a huge impact exciter or shaker. There is no easily applicable practical method available to excite large structures with natural frequencies below 1 Hz [1]. Forced vibration testing is expensive as it involves sizeable equipment and considerable time and work. Moreover,

structures must be closed down for testing, which may not be feasible for many structures such as highway bridges. In addition, ambient excitation is always present, which would cause disturbance to the measurements.

For large civil engineering structures like bridges, high-rise buildings, masts, oil platforms, etc., the excitation comes from wind, traffic, waves, or micro-seismic tremors and even at low amplitude it can buffet the structure and cause measurable vibrations. These ambient excitations are usually random in nature, comprising the frequency range of interest. Therefore, no artificial excitation is needed, although the excitation is not specifically known. Fortunately, identification techniques for output-only systems have been developed and are quite reliable if the excitation can be considered filtered white noise. Such identification techniques include parametric or non-parametric methods in the frequency or time domain.

This study is restricted to the identification of the modal parameters from the response data of the renovated bridge due to normal traffic excitation. The excitation was not known, but it was assumed to include all frequencies of interest exciting the lowest natural modes.

2. DESCRIPTION OF STRUCTURE AND MEASUREMENT DATA

The Hämeensilta Bridge in the centre of Tampere is a reinforced concrete backfilled arch-bridge with a span of 40 m and an effective width of 28.5 m (Figure 1). There is a single hinge at the top of the arch. The bridge was built in 1929 and renovated in 1997.

The data used in this study was ambient response measurements from traffic excitation. The vertical responses of the bridge deck were measured in three set-ups at 13 points (Figure 2). Each set-up consisted of five HBM B12/500 accelerometers including one reference sensor, which was fixed in all set-ups at point 10. The record length of the data was 9 600 samples with a sampling period of 0.0125 s. It was later re-sampled with a sampling period of 0.025 s.

Later, however, it was noticed that the measurement data was corrupted with additional disturbance at certain frequencies. The reason for these spurious harmonics was not clear. Fortunately, two HBM B2 vibration transducers were used in addition to the five used for the modal measurements. The signals from these two sensors were not corrupted and could be used to analyse the natural frequencies and damping ratios. Typical simultaneous time signals from two different sensors at point 11 are shown in Figure 3, and their spectra in Figure 4. The differences in the spectra can be clearly seen. It seems that the spurious frequencies appear at equal intervals.

Figure 1. The Hämeensilta Bridge.

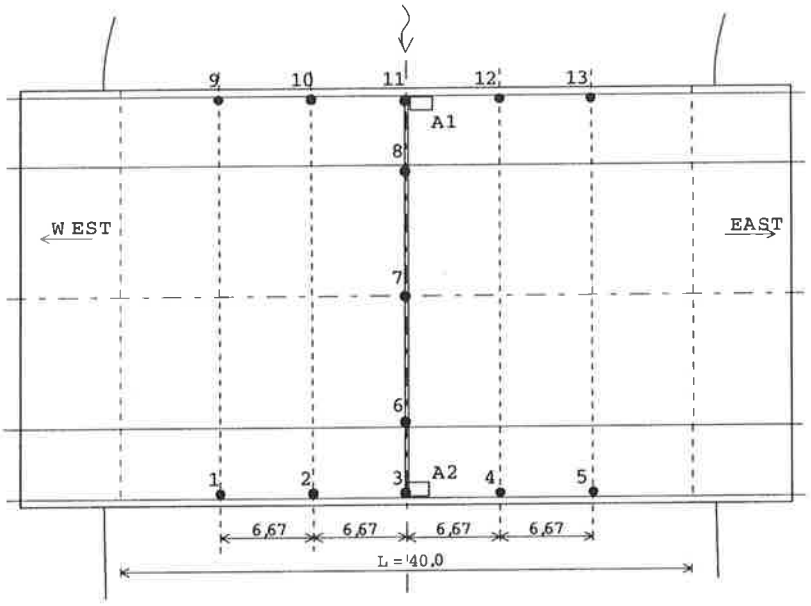


Figure 2. Measurement points. Accelerometers at points 1 – 13, HBM B2 vibration transducers at points 3 and 11 (A1 and A2).

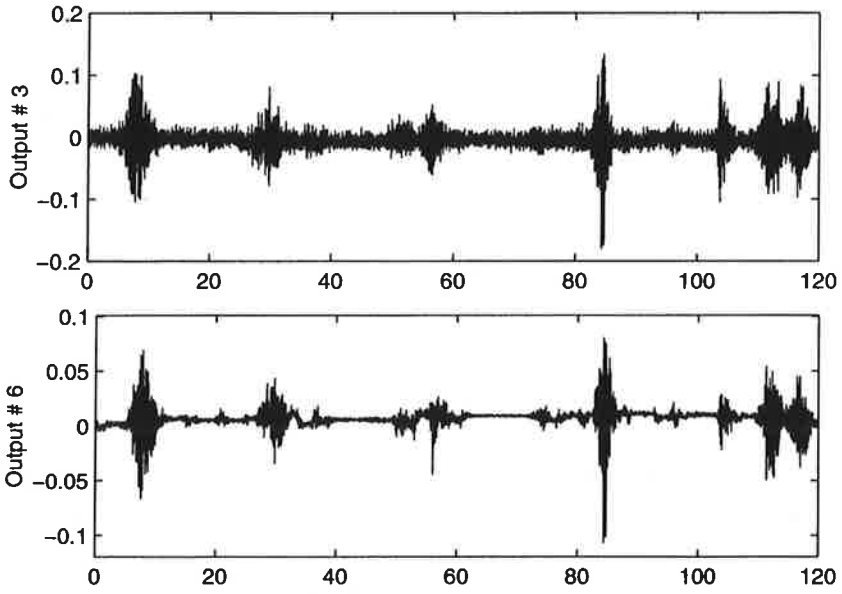


Figure 3. Time signal from accelerometer (top) and HBM B2 vibration transducer (bottom). Set-up 3, point 11.

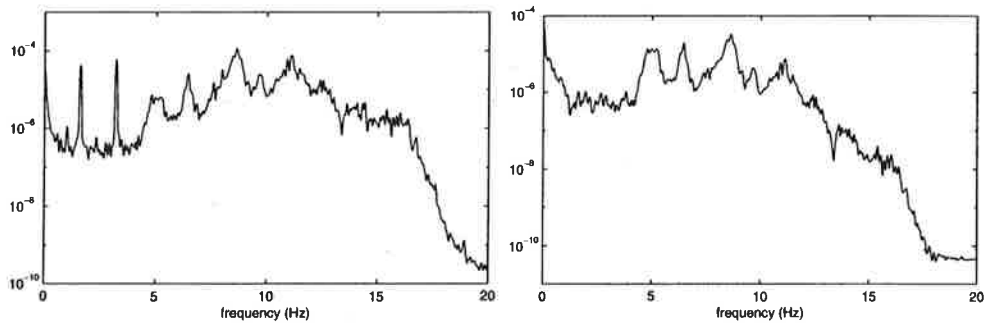


Figure 4. Spectrum from accelerometer (left) and HBM B2 vibration transducer (right). Set-up 3, point 11.

3. SYSTEM IDENTIFICATION

The system identification, also called experimental modal analysis, results in a modal model, or the dynamic properties of the structure, by using experimental data. The data from ambient vibration tests consists of output-only data without knowledge of the

excitation. Several output-only identification techniques are available, for example, spectral methods [1, 2], vector auto-regressive moving average (ARMAV) models [3], polyreference least-square complex exponential method [4], random decrement technique (RD) [5], and stochastic subspace method (SSI) [6]. The stochastic subspace technique has proved to be very reliable and relatively fast. The free parameter of the technique is the model order, which can be chosen by using the stabilisation diagram.

The stochastic subspace identification technique is based on the state-space formulation. Dynamically excited systems can be described by the following matrix equation:

$$M\ddot{U}(t) + D\dot{U}(t) + KU(t) = P(t) \tag{1}$$

where M, D, and K are the mass, damping, and stiffness matrices, respectively; P is the excitation force vector and U is the displacement vector. It can be shown (e.g. [7]) that the dynamic system in Equation 1 can be expressed in a state space formulation. If the excitation is not measured, and stochastic components (noise) are included, Equation 1 can be written in a discrete-time state space formulation

$$\begin{aligned} x_{k+1} &= Ax_k + w_k \\ y_k &= Cx_k + v_k \end{aligned} \tag{2}$$

where x_k is the discrete time state vector, y_k is the response data vector, A is the system matrix, and C is the output matrix. Vectors w_k and v_k are the process noise and measurement noise, respectively. In the case of ambient vibration testing, w_k is the only excitation. The system identification consists of estimating the discrete matrices A and C. The stochastic subspace method identifies the state space matrices from measurements by using robust numerical techniques such as QR factorisation, singular value decomposition, and least squares [4]. Once the matrices are found, the extraction of the modal parameters is quite straightforward.

4. RESULTS

The two vibration transducers only were used to define the natural frequencies and damping ratios, whereas the natural modes were defined using the transfer functions between the measurement point and the reference point at the identified natural frequencies [1]. This method is assumed to work well if there are no closely spaced modes and if the disturbance does not coincide with the natural frequencies. However, some of the modes may have been missed, because the two vibration transducers used for the system identification were located at the mid-span, which is a possible nodal line of asymmetric modes.

In the present study, the stabilisation diagram was used to select the optimal model order. In the stabilisation diagram, the identified poles are plotted against the model order. Five different stabilisation degrees were defined: new pole; stabilised frequency; stabilised

frequency and mode; stabilised frequency and damping; and stabilised frequency, damping, and mode. A typical stabilisation diagram is shown in Figure 5. In this study, the stabilisation conditions were: 0.5% difference for the frequency, 5% for damping ratio, and 2% for natural mode. It can be seen that the model was well stabilised at model orders between 10 and 15. The poles satisfying the following conditions only were plotted:

1. Eigenvalues of matrix *A* must appear in conjugate pairs;
2. Natural frequency must be greater than zero;
3. Damping ratio must be between 0 and 10%.

Another identification was performed using the five accelerometers producing disturbances, and a typical stabilisation diagram is shown in Figure 6. It can be seen that the analysis resulted in more poles, and a higher model order had to be used to extract the structural modes. The harmonic spurious modes appeared at equal intervals and could be distinguished as having a very low damping ratio. The mode shapes were also extracted from the stochastic subspace analysis.

It was seen that the identified natural frequencies differed between each set-up, decreasing from set-up 1 to set-up 3. It can be assumed that all the variations in the frequencies were due to temperature variation. Many studies have shown that temperature variation may cause high variation in the natural frequencies of civil engineering structures [8 – 11]. The identified damping ratios varied considerably between each set-up.

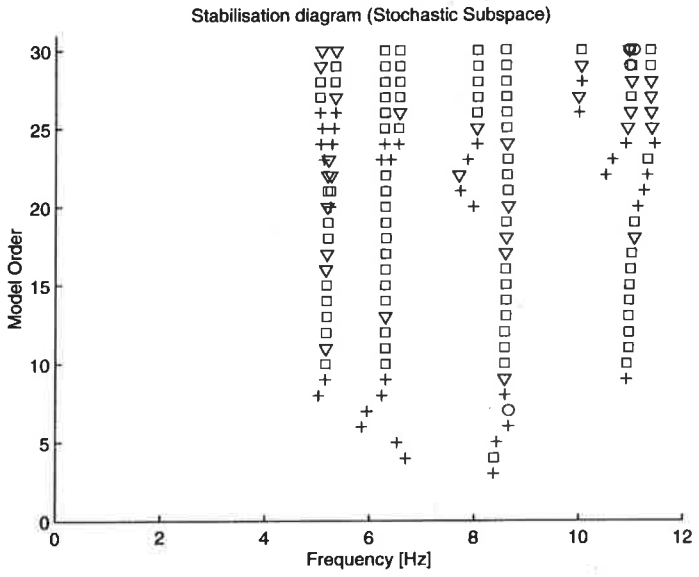


Figure 5. Stabilisation diagram of set-up 2, record 16 using the HBM B2 vibration transducers only. '+' is a new pole; 'o' stabilised frequency; 'x' stabilised frequency and damping ratio; and '□' stabilised frequency, damping ratio and natural mode.

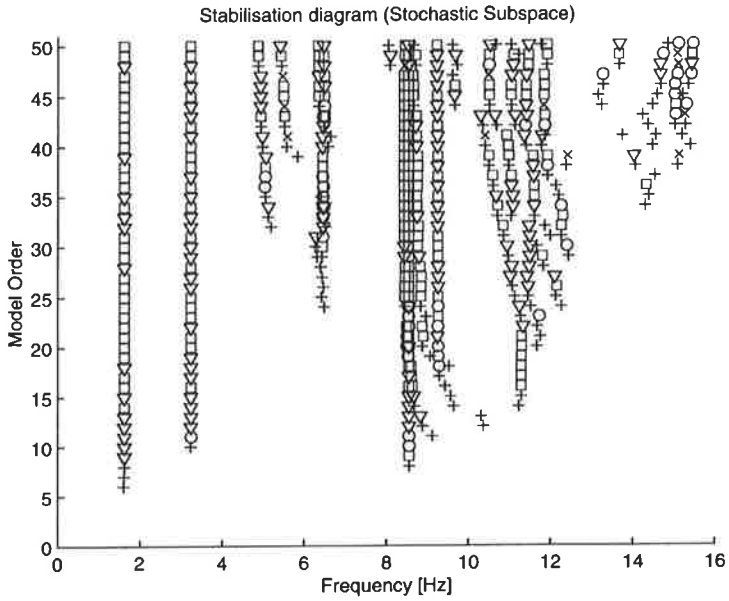


Figure 6. Stabilisation diagram of set-up 2, record 16 using five accelerometers. '+' is a new pole; 'o' stabilised frequency; 'x' stabilised frequency and damping ratio; and '□' stabilised frequency, damping ratio and natural mode.

The comparison between the mode shapes from the two analyses is shown in Figure 7–11. The identified natural frequencies and damping ratios from each set-up are also shown. One additional structural mode was found when using all the accelerometers because the nodal points of the mode were located close to the mid-span. It should be noted that the modes from the spectra led to operational modes rather than natural modes, as can be seen in Figure 10 on the left, where mode 3 also affected the fourth mode shape.

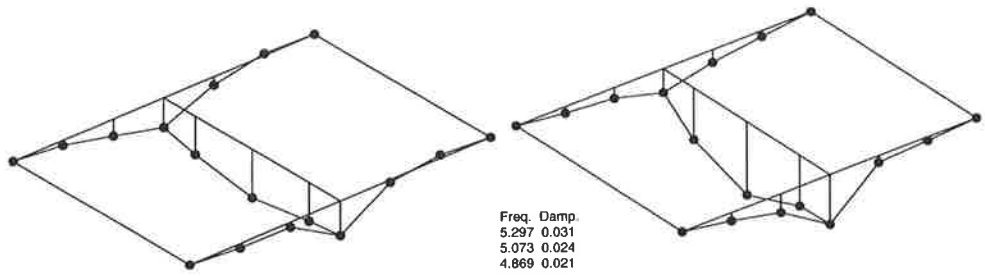


Figure 7. Mode 1 from spectral and ssi analyses. First longitudinal bending mode, 4.9 – 5.3 Hz.

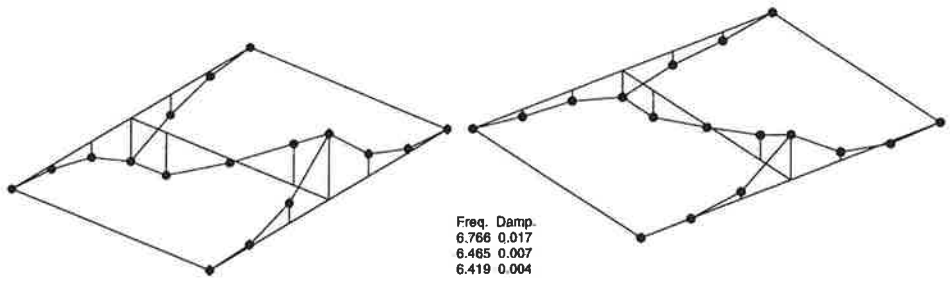


Figure 8. Mode 2 from spectral and ssi analyses. First torsion mode, 6.4 – 6.8 Hz.

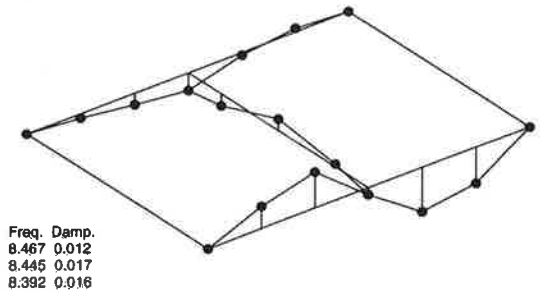


Figure 9. Mode 3 from ssi analysis. Asymmetric mode, 8.4 – 8.5 Hz.

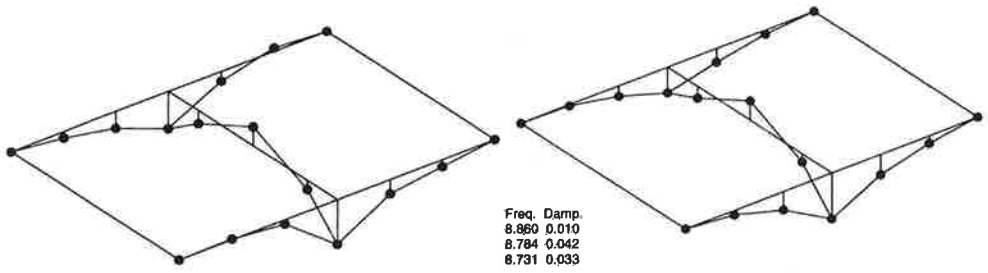


Figure 10. Mode 4 from spectral and ssi analyses. First lateral bending mode, 8.7 – 8.9 Hz.

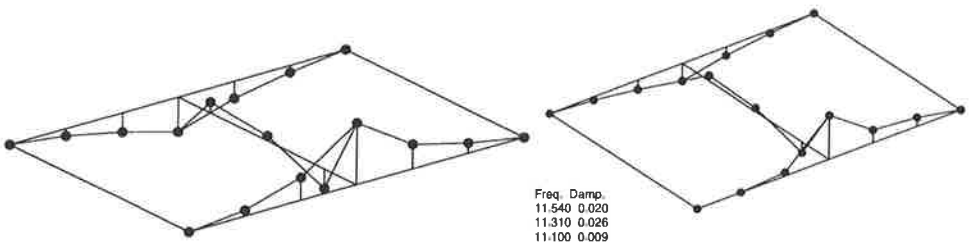


Figure 11. Mode 5 from spectral and ssi analyses. Second lateral bending mode, 11.1 – 11.5 Hz.

5. CONCLUSION

An experimental modal analysis of the Hämeensilta Bridge was performed using responses due to unknown traffic excitation. The stochastic subspace technique was used for the identification of the modal parameters. Since the signals from the accelerometers consisted of spurious frequencies, two analyses were performed – the first using two vibration transducers only, the signals of which were not corrupted, and the second using all five accelerometers. The latter analysis was made at a higher cost, because a higher model order had to be used, but it resulted in an additional structural mode. It also led to natural modes rather than operational modes. The estimated natural frequencies decreased from set-up 1 to set-up 3, which is probably due to temperature variation. The reason for the spurious frequencies was not clear. The corrupted signals should be detected and avoided already on site, but very often, spurious frequencies could not be avoided, e.g. rotating machines cause harmonic vibrations not corresponding to their natural frequencies. The ability to distinguish these vibrations in a rotating machine is an important application [12, 13] and is a subject for further study.

REFERENCES

1. Felber, A.J. and Cantieni, R. 1996. *Introduction of a new ambient vibration testing system — Description of the system and seven bridge tests.* EMPA report No. 156'521. Dübendorf, 154 p.

2. Brincker, R., Zhang, L. and Andersen, P. 2000. *Modal identification from ambient responses using frequency domain decomposition.* Proceedings of IMAC-XVIII: A Conference on Structural Dynamics, San Antonio, Texas, USA, February 7–10, 2000. Society for Experimental Mechanics, Bethel, Connecticut, 625 – 630.

3. Anderssen, P. 1997. *Identification of civil engineering structures using vector ARMA models.* Ph.D. Thesis, Aalborg University. 244 p.

Thermal hardening of glues [11] as well as cooking a potato are everydays examples. At its initial state the potato is composed entirely of virgin phase (a) which, when cooked sufficiently, transforms entirely to another phase (i). At its intermediate state the potato may be seen as a mixture of uncooked phase (a) with a mass fraction β^a and a cooked phase (i) with mass fraction β^i. It is clear that the cooking of the potato is irreversible and depends strongly on time history. It is impossible to get again a virgin non-cooked potato by only cooling a cooked one; *cf.* the practically reversible case of freezing and defreezing (liquid-solid phase transformation of the water content) a portion of salmon.

The formulation of the problem is made following the productive approach used in [3], [11], [12], [13], [14], [17] and [18] where the behaviour of the system is completely described, first, by the choice of relevant description scale, state variables, internal variables, an expression for the specific free energy, a pseudo-potential of dissipation and, finally, writing down the conservation laws and the constraints imposed by the Clausius-Duhem inequality [16] and by other physical and geometrical constraints.

There are other formulations for phase changes occurring within a temperature range [5]. A formulation of the reversible Stefan problem without computing the free boundary is presented in [3] where the water content is approximated by a homographic function[1] depending on temperature. In some numerical examples, presented in this paper, phase change occurs in reality all the time within a continuous range of temperatures.

Some other ways of treating phase change problems and related numerical aspects are presented in [5], [6], [15], [21], [22] and [25]. Another numerical method is shown in [23] for two-dimensions, where only regions close to the phase change surface are spatially refined. An enthalpy formulation was used in [7] conjuctively with a finite-difference method and a method tracking the interface.

2 Physical problem

Formulation of the problem: Find the temperature $T(\mathbf{x},t)$ and the converted mass fraction $\beta(\mathbf{x},t) \equiv \beta^i(\mathbf{x},t)$ of the solid mixture. The mixture of mass $m = m^i + m^a$ consists of active phase with mass fraction $\beta^a(\mathbf{x},t) = m^a(\mathbf{x},t)/m$ and of converted phase with mass fraction $\beta^i(\mathbf{x},t) = m^i(\mathbf{x},t)/m$. The constitutive equation $\beta(\mathbf{x},t) = f(T,\chi)$ gives the mass fraction of the converted inert phase as a function of the temperature and a history parameter χ. The second principle of thermodynamic constrains the admissible form of $f(T,\chi)$ for arbitrary non-monotonic heating including cooling and reheating cycles.

The rate of conversion $d\beta^i(\mathbf{x},t)/dt$ is then computable through the constitutive law $f(T,\chi)$. The time derivatives $d\beta^k(\mathbf{x},t)/dt$ are discontinuous because of the irreversibility and of the constraints on mass fractions. Because of causality, left time derivatives are used.

3 Variables and definitions

The state variable is the Kelvin temperature $T(\mathbf{x},t) > 0$. As internal variable, $\chi(\mathbf{x},t) \equiv \max(\{T(\mathbf{x},\tau)\}, 0 < \tau \leq t)$ is chosen. The mass fraction $\beta = 1 - \beta^a(\mathbf{x},t)$ is a dependent variable. The constraints $\beta = f(T,\chi)$, $\beta^i \geq 0$, $\beta^a \geq 0$, $\beta^i + \beta^a = 1$ and $d\beta(\mathbf{x},t)/dt \geq 0$ should

[1] Fonction, quotient de deux fonctions du premier degré, dont la courbe représentative est une hyperbole dont les assymptotes sont parallèles aux axes des coordonnées (Le Robert éléctronique).

hold. The thermodynamical properties of the mixture are: the apparent thermal conductivity $\lambda(\beta) = \beta\bar{\lambda}^i\rho/\bar{\rho}^i + (1-\beta)\,\bar{\lambda}^a\rho/\bar{\rho}^a$, the apparent specific thermal capacity $c(\beta) = \beta c^i + (1-\beta)c^a$ and the apparent density $\rho = \sum \rho^k$. For the phase $k \in \{a, i\}$, the volume is V^k, the apparent density $\rho^k = m^k/V$, the bulk density $\bar{\rho}^k = m^k/V^k$, the bulk thermal conductivity $\bar{\lambda}^k$ and the bulk specific thermal capacity $c^k \equiv \bar{c}^k$.

4 Energy balance and irreversibility

Free energy: The conservative behaviour of the mixture is described by the specific free energy $\psi(T, \beta)$, and its dissipative behaviour by a pseudo-potential of dissipation $\phi(\nabla T, \dot{\beta})$. The specific free energy of the mixture is

$$\psi = \beta\psi^i + (1-\beta)\psi^a + TI(\beta). \tag{1}$$

The chosen constitutive equation is such that the mass conservation of the mixture is always satisfied. The last term in (1) accounts for the constraint on the mass fraction. The indicator function $I(\beta) = 0$ when the constraints are satisfied, otherwise $I(\beta) = +\infty$. The contribution of the non-smooth part of specific free energy to the specific internal energy of the mixture vanishes since $\hat{e} = \hat{\psi} - T\frac{\partial\hat{\psi}}{\partial T} = TI(\beta) - T\frac{\partial TI(\beta)}{\partial T} = TI(\beta) - TI(\beta) = 0$. Following [11], [13] and [14], the specific free energy expressions are chosen as $\psi^i = -c^iT\ln(\frac{T}{T_0}) - \ell\frac{T-T_0}{T_0}$ and $\psi^a = -c^aT\ln(\frac{T}{T_0})$. The (smooth part of) the specific free energy of the mixture is then

$$\psi = -cT\ln(\tfrac{T}{T_0}) - \ell\tfrac{T-T_0}{T_0}\beta\,(T,\chi) = -cT\ln\left(\tfrac{T}{T_0}\right) - \ell\tfrac{T-T_0}{T_0}f\,(T,\chi)\,. \tag{2}$$

Specific internal energy: Substituting expression (2) in $e = \psi + sT = \psi - T\frac{\partial\psi}{\partial T}$, one gets

$$e = cT + \ell\beta, \tag{3}$$

where the specific thermal capacity of the mixture is $c\,(\beta) = \beta c^i + (1-\beta)\,c^a$. The quantity $e^i - e^a$ is the energy needed to convert (a) into (i). After some algebraic manipulations, one gets $\bar{\ell} \equiv e^i - e^a = \ell - \left(c^a - c^i\right)T$, where $e^i = \psi^i - T\frac{\partial\psi^i}{\partial T} = c^iT + \ell$ and $e^a = \psi^a - T\frac{\partial\psi^a}{\partial T} = c^aT$. $\bar{\ell}$ is the effective latent heat of phase change.

Energy conservation: The local energy balance [13] is

$$\rho\frac{de}{dt} = -\nabla\cdot\left(T\vec{Q}\right) + TR = \rho\sum_k \frac{d}{dt}\left(\beta^k e^k\right) = \rho\sum_k \beta^k\frac{de^k}{dt} + \rho\sum_k e^k\frac{d\beta^k}{dt}. \tag{4}$$

Inserting the internal energies in (4), one finally arrives to the energy equation

$$\rho c\frac{dT}{dt} + \rho\left(\ell - \left(c^a - c^i\right)T\right)\frac{d\beta}{dt} = -\nabla\cdot\vec{q} + r, \tag{5}$$

where the heat flux vector $\vec{q} = T\vec{Q}$ and the source term $r = TR$.

Clausius-Duhem inequality: The local form of the second principle of thermodynamic is

$$T\left(\rho\frac{ds}{dt}+\nabla\cdot\vec{Q}-R\right)\geq 0, \tag{6}$$

for any thermodynamically admissible process. The right side of (6) is usually chosen as a pseudo-potential of dissipation. By expressing (4), using the free energy $\psi(T,\beta)$, and remembering that $s=-\frac{\partial\psi}{\partial T}$, one obtains

$$\rho\frac{de}{dt}=\rho\left(\frac{\partial\psi}{\partial\beta}\frac{\partial\beta}{\partial t}+T\frac{ds}{dt}\right)=-T\vec{\nabla}\cdot\vec{Q}-\vec{Q}\cdot\nabla T+TR, \tag{7}$$

and finally — after reordering terms in order to identify the Clausius-Duhem inequality represented by (6) — one arrives to the useful inequality

$$T\left(\rho\frac{ds}{dt}+\nabla\cdot\vec{Q}-R\right)=-\rho\frac{\partial\psi}{\partial\beta}\frac{\partial\beta}{\partial t}-\vec{Q}\cdot\nabla T\geq 0, \tag{8}$$

which is used to constrain the constitutive equation. Requiring that both parts of the inequality (8) are separately positive, one obtains the classical constraint on the Fourier heat conduction constitutive law and the condition $-\rho\frac{\partial\psi}{\partial\beta}\frac{\partial\beta}{\partial t}\geq 0$. The remaining part of the constraint inequality will provide the necessary conditions to be satisfied by the constitutive equation. Assuming irreversibility $\dot{\beta}\geq 0$, one easily finds the constraint $-\frac{\partial\psi}{\partial\beta}\geq 0$ to be fulfilled by the constitutive equation $\beta=\beta\left(T,\chi\right)$.

5 Heating–cooling rule

The constitutive equation may be of the general form $\beta(\mathbf{x},t)=f(T,\dot{T},\chi)$. The formulation shown in the current paper applies also in this case. But, now we limit ourself to cases where β is given as $f(T,\chi)$. The irreversibility constraint $\dot{\beta}\left(x,t\right)\geq 0$ should hold for every thermodynamically admissible process. The constitutive law is assumed, by definition, admissible for monotonic heating ($\dot{T}\geq 0$). Consider non-monotonic heating with heating–cooling cycles. In cooling phase the rule $\frac{\partial\beta}{\partial T}=0$ for $\dot{T}<0$ is used to ensure that irreversibility condition $\frac{d\beta}{dt}=\frac{\partial\beta}{\partial T}\cdot\dot{T}+\frac{\partial\beta}{\partial\chi}\cdot\dot{\chi}\geq 0$ ($\dot{\chi}=0$) holds. It is supposed that cooling happens without thermal degradation. Thermodynamically, any number ϵ such that $\frac{\partial\beta}{\partial T}<\varepsilon\leq 0$, i.e. $\frac{\partial\beta}{\partial T}\in[0,\frac{\partial\beta}{\partial T}\left|_{\chi(\mathbf{x},t)}\right.]$ is admissible since then the irreversibility condition holds automatically. For the reheating case $\dot{T}>0$ of the material point — which have been already, in earlier stages, cooled down to the temperature $T<\chi$ — the value $\frac{\partial\beta}{\partial T}=0$ is retained, and this until the temperature reaches again the value $\chi\left(\mathbf{x},t\right)\equiv\max\left\{T(\mathbf{x},\tau),0\leq\tau\leq t\right\}$.

N.B. that, for instance, one may always, if relevant, refine the "description scale" (by opening the "black box" of the constitutive equation) and describe the remaining mass fraction of the converted phase by the basic rate equation $\dot{\beta}=-k(1-\beta)^n$, $k=Ae^{-E/RT}$, $A>0$ and $n>0$ which now intrinsically contains the irreversibility conditions. Doing this, explicativity but also complexity is "gained". N.B. that the description scale, which is part of the statement of the problem, is not unique; it is usually dictated by the questions posed by the problem and by the desired level of details.

6 Variational formulation of the energy equation

The unknowns are the temperature $T(\mathbf{x},t)$ and the mass fraction $\beta(\mathbf{x},t)$ of the inert phase at any point $(\mathbf{x},t) \in \Omega\times]0,t_{max}] \subset \mathbb{R}^3 \times \mathbb{R}^+$. The normal of the heat flux $\bar{q}_n$ is given at the boundary Γ_n (Neumann BC), and at the boundary Γ_d the Dirichlet boundary conditions $\bar{T}_d$ is imposed. The boundaries Γ_n and Γ_d are not overlapping. Initial conditions $T(\mathbf{x},0)$ and $\beta(\mathbf{x},0)$ are given.

The variational formulation is obtained by multiplying the local form of energy equation by the smooth test function $v \in V$ which vanishes at the boundary Γ_d. Integration by part leads to the weak formulation: *Find* $T(\mathbf{x},t) \in H^1(\Omega) \cap \{T \mid T(\mathbf{x} \in \Gamma_d, t) = \bar{T}_d\}$, *such that*

$$(\rho c\dot{T}, v) + (\lambda\nabla T, \nabla v) = -(\bar{q}_n, v)_{|\Gamma_n} - (\rho(\ell - (c^a - c^i)T)\dot{\beta}, v) + (r, v), \tag{9}$$

for $\forall t > 0$ and $\forall v(\mathbf{x}) \in V(\Omega) = \{v \mid v \in H^1(\Omega)$ *and* $v_{|\Gamma_d} = 0\}$, where $\beta(\mathbf{x},t) = f(T,\chi)$ is given by the constitutive equation satisfying the constraint $\dot{\beta}(\mathbf{x},t) \geq 0$. For shortness, the notation $(u,v) = \int_\Omega u(\mathbf{x})v(\mathbf{x})\,d\Omega$ and $(q_n,v)_{|\Gamma_n} = \int_{\Gamma_n} \vec{q}\cdot\vec{n}v\,d\Gamma$ is used.

7 On the stability of the continuous problem

Stability and error analysis of two-phase Stefan problem are treated in [23]. In the present paper, one stability aspect for some special cases of the solution of equation (9) is presented. Consider the case with adiabatic boundary conditions $((\bar{q}_n, v)_{|\partial\Omega} = 0)$ where external source terms are absent $((r,v) = 0)$. Only irreversible $(\dot{\beta} \geq 0)$ and endothermic $(\ell > 0)$ phase changes are allowed. Then, under which conditions the norm of the solution is decreasing with time? N.B. that the exothermic case may, under certain conditions, lead to physical instability [1]. The next constraints hold by definition: $T > 0$ and bounded. Thermal conductivity $\lambda(\beta)$ and thermal capacity $c(\beta)$ are strictly positive. Let assume that their norms are bounded. The norm $\|u\|(t) \equiv \|u(.,t)\|_{L_2(\Omega)}$ is used.

Stability analysis is performed by choosing $v = T$ in (9) which after reordering gives

$$(\rho c\dot{T}, T) = -\|\sqrt{\lambda}\nabla T\|^2 - (\rho\ell\dot{\beta}, T) + (\rho(c^a - c^i)\dot{\beta}T, T). \tag{10}$$

Considering the negativity of the signum of (10), one obtains the energy inequality. In (10), the terms $-\|\sqrt{\lambda}\nabla T\|^2$ and $-(\rho\ell\dot{\beta}, T)$ are both separately negative. Then inserting the identity

$$\frac{1}{2}\frac{d}{dt}\int_\Omega \rho c(\beta)T^2 d\Omega \equiv \frac{1}{2}\frac{d}{dt}\|\sqrt{\rho c}T\|^2(t) = (\rho c\dot{T}, T) - \frac{1}{2}(\rho(c^a - c^i)\dot{\beta}T, T) \tag{11}$$

into equation (10), one finally obtains the useful form of the energy equation

$$\frac{1}{2}\frac{d}{dt}\|\sqrt{\rho c}T\|^2(t) = -\|\sqrt{\lambda}\nabla T\|^2 - (\rho\ell\dot{\beta}, T) + \frac{1}{2}(\rho(c^a - c^i)\dot{\beta}T, T), \tag{12}$$

which negativity is discussed. The signum of the right side may change with the evolution inducing changing stability characteristics with time. Stability of the solution is garanted when the right side of (12) is negative. Since $\frac{d}{dt}\|\sqrt{\rho c}T\|^2(t) \leq 0$ means that $\|\sqrt{\rho c(\beta)}T\|(t)$ is a decreasing function of time, and therefore $\|\sqrt{\rho c}T\|(t_i) \leq \|\sqrt{\rho c}T\|(t_j) \leq \ldots \leq \|\sqrt{\rho c}T\|(t_0)$, for $t_0 > t_j > t_i$.

Particular case: $c^a = c^i$. Stability derives directly from (12) by dropping the last term, since then $\frac{1}{2}\frac{d}{dt}\|\sqrt{\rho c}T\|^2(t) = -\|\sqrt{\lambda}\nabla T\|^2 - (\rho\ell\dot{\beta}, T) \leq 0$. This case is similar to dissipative case of the standard diffusion equation with $\dot{\beta} = 0$.

General case: $c^a \neq c^i$. The stability holds if $-\|\sqrt{\lambda}\nabla T\|^2 - (\rho\ell\dot{\beta}, T) + \frac{1}{2}(\rho(c^a - c^i)\dot{\beta}T, T) \leq 0$. The stronger condition $-(\rho\ell\dot{\beta}, T) + \frac{1}{2}(\rho(c^a - c^i)\dot{\beta}T, T) \leq 0$ leads to $\ell \geq 1/2(c^a - c^i)T$.

8 Variational form of the discrete problem

The nonlinear variational problem (9) is discretised in time by the implicit Backward-Euler scheme, and in space by the finite element method. The trial functions, $v_h(\mathbf{x}) \in V_h(\Omega) \subset V(\Omega)$, are continuous and linear in each element. The semi-discretisation gives

$$(\rho c\dot{T}, v_h) + (\lambda\nabla T, \nabla v_h) = -(\bar{q}_n, v_h)_{|\Gamma_u} - (\rho(\ell - (c^a - c^i)T)\dot{\beta}, v_h) + (r, v_h). \tag{13}$$

For commodity, notation $v_{h|\Omega^e}^i = N_i^e$ is used. Semi-discretisation of (13) is achieved by seeking solutions from V_h as $T^e(\mathbf{x}, t) = \mathbf{N}^e(\mathbf{x})\mathbf{u}^e(t)$ using Galerkin method. The mass fraction $\beta^e(x, t) = \mathbf{N}^e(\mathbf{x})\mathbf{b}^e(t)$, the heat flux $\vec{q} = -\lambda(T, \beta)\nabla\mathbf{N}^e(\mathbf{x})\mathbf{u}^e(t) \equiv -\lambda(T, \beta)\mathbf{B}^e(\mathbf{x})\mathbf{u}^e(t)$, and the mass rate $\dot{\beta}^e(\mathbf{x}, t) = \mathbf{N}^e(\mathbf{x})\dot{\mathbf{b}}^e(t) \geq 0$.

9 Semi-discretisation

Rewriting (13), using matrix notation, one obtains the nonlinear set of ODEs

$$\mathbf{C}(\mathbf{u}, \mathbf{b})\dot{\mathbf{u}} + \mathbf{K}(\mathbf{u}, \mathbf{b})\mathbf{u} = \mathbf{f}_q(\mathbf{u}, \mathbf{b}) + \mathbf{f}_r(\mathbf{u}, \mathbf{b}) + \mathbf{f}_\beta(\mathbf{u}, \mathbf{b}, \dot{\mathbf{b}}), \tag{14.1}$$

$$[\mathbf{C}(\mathbf{u}, \mathbf{b}) + \mathbf{C}_\beta(\mathbf{u}, \mathbf{b}, \dot{\mathbf{b}})]\dot{\mathbf{u}} + \mathbf{K}(\mathbf{u}, \mathbf{b})\mathbf{u} = \mathbf{f}_q(\mathbf{u}, \mathbf{b}) + \mathbf{f}_r(\mathbf{u}, \mathbf{b}). \tag{14.2}$$

For brevity, TAR denotes formulation (14.1) which is used with Time Averaged mass Rate, while ERTAR is the Enthalpy-like Regularised formulation (14.2), also with Time Averaged mass Rate. The formulation ERTAR (Chap. 10 of this paper) is more stable than TAR.

Some stability aspects of non-autonomous IVPs can be found in [9, Chap. 10.4–10.6]. For a discussion on stability and error analysis *cf.* [8, Chap. "Nonlinear systems of ODEs"]. Some details on analysis of nonlinear ODEs can be found in [24]. For a discussion on error estimates of nonlinear parabolic problems *cf.* [10]. Mathematical aspects of Stefan problem are presented in [23]. In general, stability characteristics of the solution change with time.

The nonlinear system of equations (15), arising from time-integration of ODEs (14), is solved by the Picard fixed-point iteration scheme [20, Chap. 3.2] and [26, Chap. 18, pp. 450–452]. Within each time step solution of (15), $\mathbf{u}(t) = \mathbf{G}(\mathbf{u})(t)$, is seeked iteratively. This scheme converges if $\|\mathbf{G}'(\mathbf{u})\| < C \leq 1$ [20, Chap. 3.2]; *cf.* the incremental scheme of [25].

Using Newton-Cotes integration scheme diagonal capacity matrices are obtained; having this at hand, a typical Picard iteration step i, at time t_n, looks as

$$\mathbf{u}_n^{i+1} = \mathbf{u}_{n-1} + \Delta t_n \mathbf{C}^{-1}{}_n^i \left(-\mathbf{K}_n^i \mathbf{u}_n^i + \mathbf{f}_n^i\right), \tag{15}$$

$$\mathbf{C}^{-1}{}_n^i = \mathbf{C}^{-1}(\mathbf{u}_n^i, \mathbf{b}_n^i), \quad \mathbf{K}_n^i = \mathbf{K}(\mathbf{u}_n^i, \mathbf{b}_n^i),$$
$$\mathbf{f}_n^i = \mathbf{f}_q(\mathbf{u}_n^i, \mathbf{b}_n^i) + \mathbf{f}_r(\mathbf{u}_n^i, \mathbf{b}_n^i) + \mathbf{f}_\beta(\mathbf{u}_n^i, \mathbf{b}_n^i, \dot{\mathbf{b}}_n^i).$$

When using ERTAR, the capacity matrix $\mathbf{C}$ in (15) should be replaced by $\mathbf{C} + \mathbf{C}_\beta$, and the force vector $\mathbf{f}_\beta = \mathbf{0}$. In TAR formulation, $\mathbf{C}_\beta = \mathbf{0}$ and $\mathbf{f}_\beta \neq \mathbf{0}$. The i^{th} iteration is continued until the convergence criteria $\left\|\mathbf{r}^i\right\|_{L_2(\Omega)}(t_n) \equiv \left\|\mathbf{u}^i - \mathbf{G}\left(\mathbf{u}^{i-1}\right)\right\| = \left\|\mathbf{u}_n^{i-1} - \mathbf{u}_n^i\right\| \leq RTOL_u \cdot \|\mathbf{u}_n\| + ATOL_u$ and $\left\|\mathbf{b}_n^{i-1} - \mathbf{b}_n^i\right\| \leq TOL\beta$ are fulfilled. The absolute and the relative tolerances are, respectively, $ATOL_u$ and $RTOL_u$. $TOL\beta$ stands for both.

The L_2-norm of the residual of the global discrete equation, at time t_n, is $\|\mathbf{r}\|_{L_2(\Omega)}(t_n)$. At each time step, the Picard iteration is initialized with the converged solution of the previous time step, where the initial values of the matrices and the vectors are evaluated. Within the time step, all the quantities are updated during the iteration. Note that Picard iteration does not need evaluations of derivatives of matrices or vectors which is not the case in the Taylor expansions used in Newton's methods. Picard iteration scheme may have advantages in regions where these derivatives are discontinuous.

The global heat capacity matrix $\mathbf{C}$, conductivity matrix $\mathbf{K}$, as well as the global force vector $\mathbf{f}$ are assembled from the contribution of their respective elementary matrices (indexes n and i are dropped for clarity)

$$
\begin{aligned}
\mathbf{C}^e &= \int_{V^e} \rho c(T,\beta)\mathbf{N}^T\mathbf{N}dV, & {\mathbf{f}_q}^e &= -\int_{\partial V^e \cap \Gamma_n} \bar{q}_n \mathbf{N}^T dV, \\
\mathbf{C}_\beta^e &= \int_{V^e} \rho\bar{\ell}(T) sgn(\dot{\beta}(\mathbf{x},t)) \frac{|\nabla\beta(\mathbf{x},t)|}{|\nabla T(\mathbf{x},t)|}\mathbf{N}^T\mathbf{N}dV, & {\mathbf{f}_\beta}^e &= -\int_{V^e} \rho\bar{\ell}\,(T)\,\dot{\beta}(\mathbf{x},t)\mathbf{N}^T dV, \\
\mathbf{K}^e &= \int_{V^e} \lambda(T,\beta)\mathbf{B}^T\mathbf{B}dV, & {\mathbf{f}_r}^e &= \int_{V^e} \rho r(\mathbf{x},t)\mathbf{N}^T dV, \qquad (16) \\
& & \bar{\ell}\,(T) &= \ell - (c^a - c^i)T(\mathbf{x},t).
\end{aligned}
$$

The conductivity matrices are integrated with one Gauss-Legendre integration point, localised at the barycenter of the element.

10 On regularisation of the mass rate term

The time derivative $d\beta(\mathbf{x},t)/dt$ is evaluated using the backward Euler difference scheme. Regularisation of the rate term $\dot{\beta}$ is needed, see for instance [12, p. 235].

In the present paper a smoothed rate $\bar{\dot{\beta}}(t_n) = \theta\dot{\beta}(t_n) + (1-\theta)\dot{\beta}(t_{n-1})$ is tested with TAR, $\theta \in]0,1]$. N.B. that this stabilised rate $\bar{\dot{\beta}}$ fulfils automatically the irreversibility constraint when the non-regularised time derivative is formed such that it fulfils the positivity condition.

ERTAR-regularisation, one of the main results of the current work, is now presented. This scheme is formulated by analogy with the enthalpy formulation for reversible phase changes [6] and [22], but now with extension to irreversible case, and with use of mass fraction rate. Now, the regularised rate $\frac{d\beta}{dt} \approx sgn(\dot{\beta})\frac{|\nabla\beta|}{|\nabla T|}\dot{T}$ is developed for irreversible phase change. It is not difficult, for the irreversible case, to write a regularised expression for $\rho(c\dot{T} + \bar{\ell}\dot{\beta})$ as $\rho(c + \bar{\ell}\frac{\partial\beta}{\partial T})\dot{T} \approx \rho(c + sgn(\dot{\beta})\bar{\ell}\frac{|\nabla\beta|}{|\nabla T|})\dot{T}$. N.B. that this regularised expression still fulfils the irreversibility constraint. This regularisation scheme is powerful when treating irreversible (and reversible) phase changes as will be shown later. The regularised term gives in the weak formulation the "augmented" capacity matrix $\mathbf{C} := \mathbf{C} + \mathbf{C}_\beta$ with $\mathbf{C}_\beta$ defined in (16). The essence of the ERTAR-regularisation is the approximation of the time derivative of the converted mass fraction with the ratio of two smoother positive quantities respecting

the irreversibility constraint. Naturally, regularisation introduces regularisation errors; a compromise between consistency and stability should be found.

Quadrature of $\mathbf{C}_\beta$ and $\int_{V^e} \rho\bar{\ell}\,(T(\mathbf{x},t))\,\dot{\beta}^e(\mathbf{x},t)\mathbf{N}(\mathbf{x})^T dV \approx \sum w_i \rho\bar{\ell}(T(\mathbf{x}_i,t))\dot{\beta}^e(\mathbf{x}_i,t)\mathbf{N}(\mathbf{x}_i)^T$ is performed using one Gauss-Legendre integration point $\mathbf{x}_i$, localised at the barycenter $\bar{\mathbf{x}}$ of the element. Mass fraction rates are evaluated as $\dot{\beta}^e(\bar{\mathbf{x}},t) \approx \texttt{mean}(\sum_{i=1}\dot{\beta}^e(\mathbf{x}_i,t))$ where i=1:$NNOD$ is the index of the nodes of the element. The mass fraction β is calculated from the constitutive equation such that the irreversibility holds (Chap. 11 of this paper).

11 An algorithm for calculating the mass fraction and its rate

The notation $\beta^e(\mathbf{x}_i,t) \equiv \mathbf{b}(t)$ and $\dot{\beta}^e(\mathbf{x}_i,t) \equiv \dot{\mathbf{b}}(t)$ is used, where i=1:$NDOF$ is the index of the global nodal-points, $NDOF$ the number of global degrees of freedom. The current time step being t_n, then, within the Picard iteration step i:

DO `idof = 1 to NDOF` % `idof=1:NDOF` indixes of the global nodal points.
$\tilde{\chi}^i_n(\texttt{idof}) = T(\texttt{idof}, t_n)$
$\tilde{\beta}^i_n(\texttt{idof}) = f\left(T^i_n(\texttt{idof}, t_n), \tilde{\chi}^i_n\right)$
$\beta^i_n(\texttt{idof}, t_n) = \texttt{max}(\tilde{\beta}^i_n, \beta_{n-1}(\texttt{idof}, t_{n-1}))$
END DO `idof`
$d\beta^i_n(:,t_n)/dt = (\beta(:,t_n) - \beta(:,t_n))\,/\Delta t_{n-1}$ % not averaged rate
IF (time smoothing) **THEN**
$\dot{\beta}(:,t_n) := \theta d\beta^i_n(:,t_n)/dt + (1-\theta)d\beta^i_n(:,t_{n-1})/dt$ % averaged rate, $\theta \in]0,1]$.

12 Automatic time step control

A time step control scheme for nonlinear heat conduction problem with phase change can be found in [6]. A time step refinement strategy for linear case is presented in [19]. In [8], time step control was extended to nonlinear heat conduction (without phase change). In these strategies control of the size of the time step is based on a "maximum allowed" temperature increment at each time step.

In this paper, a simple time step control strategy is developed for nonlinear heat conduction problems with irreversible (and reversible) phase changes based on controlling simultaneously the increments of the temperature and the mass fraction. Strategies based only on temperature changes will fail in regions where phase changes happen almost with no temperature rise. Since then $\|\mathbf{T}_{n+1} - \mathbf{T}_n\|$ tends to zero, but on the other hand, $\dot{\beta}$ may be large and therefore time step refinement is needed in this region.

The algorithm proposed here uses only available information. This strategy works robustly with ERTAR even for phase changes occurring within very narrow temperature ranges. But with TAR, this refinement strategy does not work for narrow phase change temperature; as expected, this leads to relatively very small time steps.

The proposed algorithm: Choose Δt_{min}, Δt_{max}, tolerance $tol\Delta T$ for temperature change and $tol\Delta\beta$ for change in mass fraction. Let $n = 1$ be the initial step with Δt_n given. The solutions $\mathbf{T}_n$ and β_n are known at the previous time steps t_n. The current time is t_{n+1}.

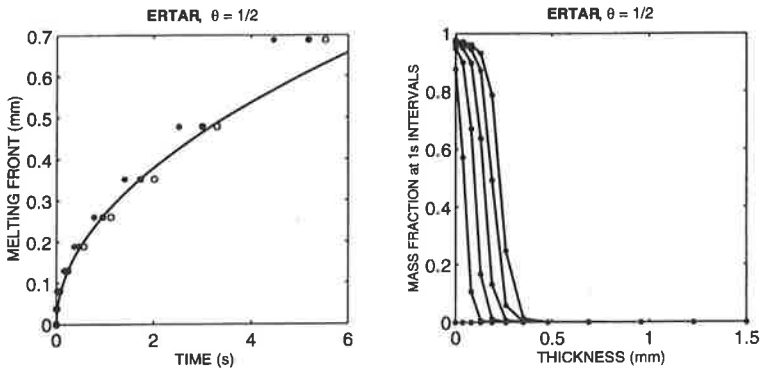


Figure 1: (a) Melting front, – - Anal., ∘ - Num. $T > 10^{-7}$ $(°C)$, $*$ - $T > 10^{-8}$ $(°C)$ and $\bullet$ - $\beta > 5 \times 10^{-6}$, (b) mass fraction profiles.

(**1**): using formulation ERTAR compute $\mathbf{T}_{n+1}$, β_{n+1} and $\dot{\beta}_n$
(**2**): `IF` $\|\mathbf{T}_{n+1} - \mathbf{T}_n\| \leq tol\Delta T/2$ `AND` $\|\beta_{n+1} - \beta_n\| \approx \|\Delta t_n \dot{\beta}_n\| \leq tol\Delta\beta/2$ `THEN`
Accept $\mathbf{T}_{n+1}$ and enlarge the time step $\Delta t_{n+1} := min(\Delta t_{n+1} \cdot INC, \Delta t_{max})$, $(INC > 1)$, increment $n := n + 1$ and $t_{n+1} = t_n + \Delta t_{n+1}$ `THEN GOTO` the next time step (1)
(**3**): `ELSE` refine $\Delta t_{n+1} = max(\Delta t_{min}, \Delta t_n/DEC)$, $(DEC > 1)$ and
`REPEAT` the computation of the solution at the current time (without incrementing n)
`UNTIL` condition (2) is satisfied
(`IF NO CONVERGENCE : STOP`).

Maximum, Euclidean or L_2- norms are used to control the size of the time steps.

13 EXAMPLES

Stefan problem: The melting of 1-D "ice" sheet is considered. Analytical solutions are available for some cases in [4, Chap. XI, Example III, p. 287]. The analytical solution for the melting front is $S(t) = 2\mu\sqrt{\kappa t}$, where κ is the thermal diffusivity of the water, and $\mu = 0.353$. In this example, Figure 1, the phase changes occur at $0(°C)$ within a very narrow temperature range. ERTAR is used with time step adaptivity; this scheme is efficient even though phase changes occur within a relatively small temperature range. But using TAR with time step adaptivity, it was practically impossible to compute this example since time steps were refined to practically "zero".

Next data is used: Adaptive $\Delta t \in [10^{-6}, 0.0071]$ (s), $tol\Delta T = 10^{-3}$ $(°C)$ and $tol\Delta\beta = 10^{-2}$. In the Picard iteration, $ATOL = 0$ and $RTOL = 10^{-3}$. The thickness was discretised using non-uniform mesh with eleven elements. Initial conditions: $T(\mathbf{x}, 0) \equiv T_0 = -10^{-6}$ $(°C)$, $\beta(\mathbf{x}, 0) = 0$, $\dot{\beta}(\mathbf{x}, 0) = 0$ (1/s). Boundary conditions at $x = 0$: Newton cooling law $q_n(0,t) = h(T_f(t) - T(0,t))$, with $T_f(t) = T_0 + 2(1 - e^{-10t})$ $(°C)$, and at the cold face

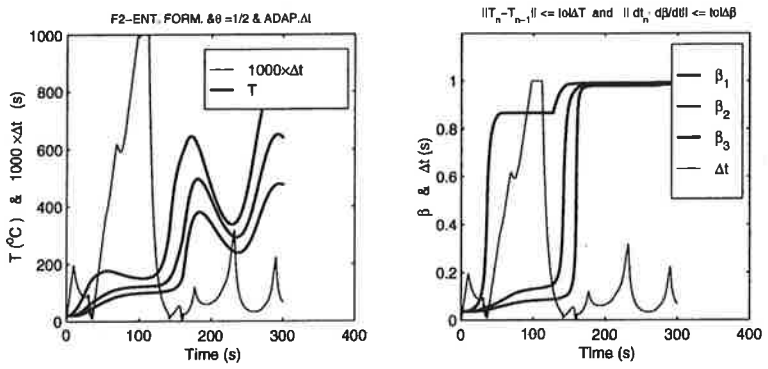


Figure 2: Adaptive time step (ERTAR). Time histories of T, β and Δt.

$q_n(d,t) = -h(T_0 - T(d,t))$. Convection coefficient h= 1000 $(W/m^2.K)$ for the both faces of the ice sheet ($\approx$ Dirichlet), no radiation, $\ell_{s\to l} = 308.15 \times 10^3$ (J/kg), $k_i = 2.219$ (W/m.K), $k_l = 0.603$ (W/m.K), $c_i = 1930.1$ (J/kg.K), $c_l = 4186.8$ (J/kg.K), and $\rho_i = \rho_l = 1000$ (kg/m^3). The constitutive equation $\beta(T) = T/(0.01 + |T|) - 1$ for $T > 0$ $(°C)$ and $\beta(T) = 0$ for $T \leq 0$ is used. N.B. that $\partial\beta(T = 0^-)/\partial T \neq \partial\beta(T = 0^+)/\partial T$. Newton's methods may have, in the neighbourhood of $T = 0$, difficulties to converge, since left and right derivatives are different. In this example, the Picard iteration converged "quickly". The results are shown in Figure 1. The position of melting front $S(t)$ is computed as: $sol._{x(t)}\left\{T(x,\tau) = 10^{-7} - 10^{-8}(°C),\ \ \beta(x,\tau) = 5 \times 10^{-6}\right\}$. In regions where $\dot{T} \approx 0$ or $\dot{\beta} \approx 0$ computing inversely the time, to cross a given temperature or mass fraction, is ill-posed.

13.1 Automatic time step control

The time-adaptivity makes sense only when ERTAR is used. The data for this example is the same as in those presented for the convergence analysis (Chap. 13.2 of this paper) except for time step adaptivity.

Adaptivity: Results are shown in Figure 2. In this example, maximum norm is used to control the time step. The time step is first refined, due to temperature rise at the beginning ($t \approx 10 - 30s$), and then at $t \approx 30 - 40s$ when the first phase changes happen. When there are practically no changes, neither in temperatures nor in mass fraction, the time step is kept constant at $t \approx 90 - 110s$. Within $t \in]120, 140[s$, where the temperature is approximatively constant, the time step is refined due to rapid changes in the mass fraction. In the region $170 < t < 300s$, the time step is refined again due to rapid changes in the temperatures.

Regularisation: Figure 3 shows an example on the effect of different regularisation schemes. Irreversible phase changes are considered. Adaptive time step is used. In this

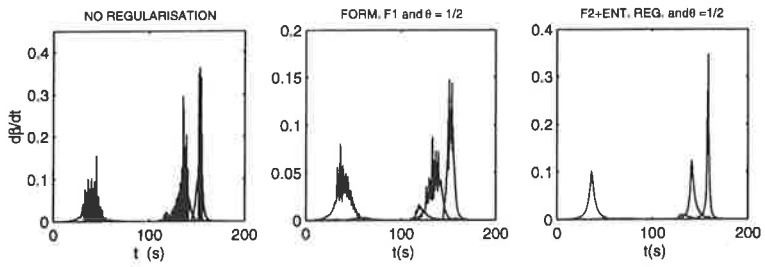


Figure 3: Regularisation of $\dot{\beta}$ for irreversible phase changes. (a) Formulation (14.1) and no regularisation, (b) formulation (14.1+TAR), (c) formulation (14.2+ERTAR).

example, the phase change region is not very narrow — about 23 % of the whole temperature range — and this introduces some stabilisation.

13.2 Convergence

Irreversible phase change: Two series of examples are studied. The first series has cyclic boundary condition (BC) and, the other one, a monotonic BC which asymptotically approaches a constant. Only convergence of TAR, with respect to the mesh, investigated by using successively refined meshes of size $h = h_0/k$, $k \in \{1, 2, 4, 8, 10, 16\}$, $h_0 = 5$ (mm), and fixed $\Delta t = 0.025$ (s). In all the examples (excluding the melting of ice and if not mentioned) the next thermodynamical constants are used: $\lambda^i = 0.1$ (W/m.K), $c^i = 900$ (J/kg.K), $\rho^i = 120$ (kg/m^3), $\lambda^a = 0.3$ (W/m.K), $c^a = 1800$ (J/kg.K), $\rho^a = 450$ (kg/m^3), $\rho = \rho^a$ (kg/m^3) and $\ell = 0.710^6$ (J/kg). $\beta(T) = 1/2 \cdot (T - T_{shift})/(\epsilon + |T - T_{shift}|)$, with $T_{shift} = 350$ $(°C)$, $\epsilon = 20$ $(°C)$ in the convergence example. In the examples on regularisation and time step adaptivity, $\epsilon = 10$ $(°C)$ and $T_{shift} = 150$ $(°C)$. The L_2-norm of the difference of two successives iterates — with respect to spatial coordinate — $\mathbf{u}^{i+1}(t) - \mathbf{u}^i(t)$ and $\beta^{i+1}(t) - \beta^i(t)$ is used with tolerances $ATOL_u = TOL_\beta = 10^{-3}$.

Cyclic BC: At $x = 0$, the generalized Neumann BC $q_{i,n} = \varepsilon_i \sigma(T_i^4 - T^4) + h_i(T_i - T)$, $i \in \{f, \infty\}$, where ∞ stays for the boundary with radiation-convection to zero Celsius, and f stays for the face at ambient temperature above zero. The constants $\epsilon_f = 0.9, \epsilon_0 = 0.8, h_f = 10$ (W/m^2) and $h_o = 25$ (W/m^2). The ambient temperature $T_f(t) = 20 + 700(1 - exp(-0.01t)) + 250sin(0.05t)$ $(°C)$. The calculation with sixteen elements does not converged. Typical temperature, mass fraction histories and profiles of an irreversible phase change example are shown in Figures 4 and 5. Some results on convergence in energy norm are shown in Figures 6 and 7.

Monotonic BC: In this series, relative tolerance 10^{-2} and $\theta = 1/2$ are used. $T_f(t) = 20 + 700(1 - exp(-0.08t))$ $(°C)$, $h_f = h_0 = 100$ $(W/m^2.K)$, and no radiation.

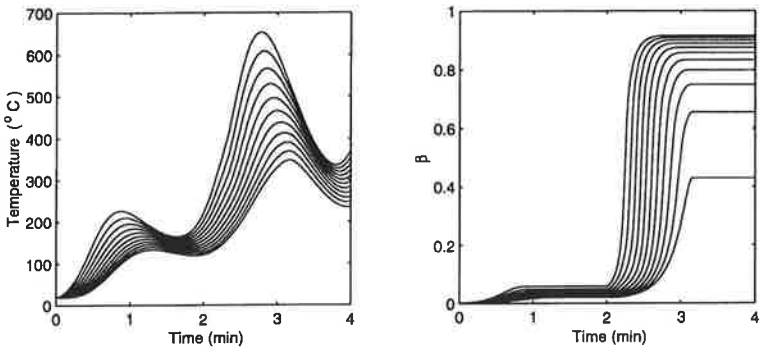


Figure 4: (a) Example of temperature and (b) mass fraction in an irreversible phase change.

The energy norm $\sqrt{(\lambda\nabla T, \nabla T)} \equiv \|\sqrt{\lambda}\nabla T(.,t)\|$ converges when refining the mesh (Figure 6). Convergence of the integrals $(\rho\ell\dot{\beta}, T)$ and $\frac{1}{2}(\rho(c^a - c^i)\dot{\beta}T, T)$ should be investigated. Analytical solutions are not accessible, this is why consistency errors are not studied. However, one can numerically check that the integrals in the weak form of the energy equations are Cauchy sequences with respect to the refinement of the finite element mesh (necessary condition). The same should be done for refinement of time step. It seems that there is some degradation of convergence for long times (Figure 7). Energy norm of two successive solutions

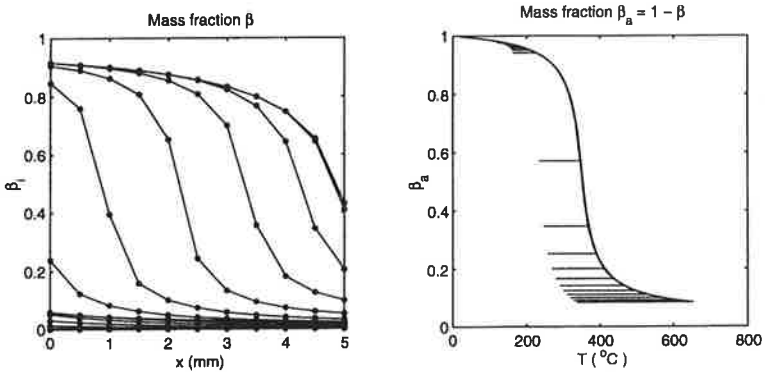


Figure 5: (a) Mass fraction profiles, (b) nodal evolution of β with T.

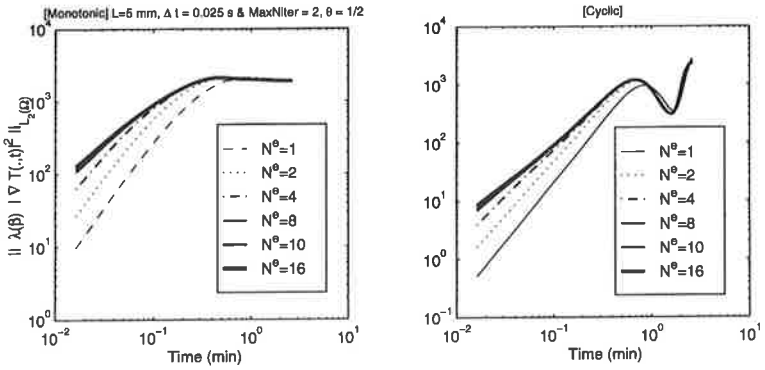


Figure 6: Energy norms, (TAR) + $\theta = 1/2$: (a) monotonic BC, (b) cyclic BC.

is calculated as $\|(u_{h/2} - u_h)(.,t)\|_E = \|\lambda(\beta_{h/2}(\mathbf{x},t))\nabla u_{h/2}(\mathbf{x},t) - \lambda(\beta_h(\mathbf{x},t))\nabla u_h(\mathbf{x},t)\|_{L_2(\Omega)}$. The norm of the error has contribution from spatial and temporal discretisation errors, *cf.* [9, eq. 16.33, p. 412] and [19, eq. 8.42, p. 159] for the case of a linear pure diffusion. To obtain numerically an estimate of the rate of convergence with respect to spatial discretisation parameter h, one should compute a stationary case. In this case there is no contribution to the error from the time-discretisation since, in linear cases, the error estimate at $t_n \leq t_N$ is $\max_{I_n} \|u(t) - U(t)\| \leq C_i S_c(\max_{n \leq N} \Delta t_n \, |\dot{u}|_{\infty,I_n} + \max_{I_n} h^2 \, \|u(t)\|_{H^2(\Omega)})$, where C_i is an

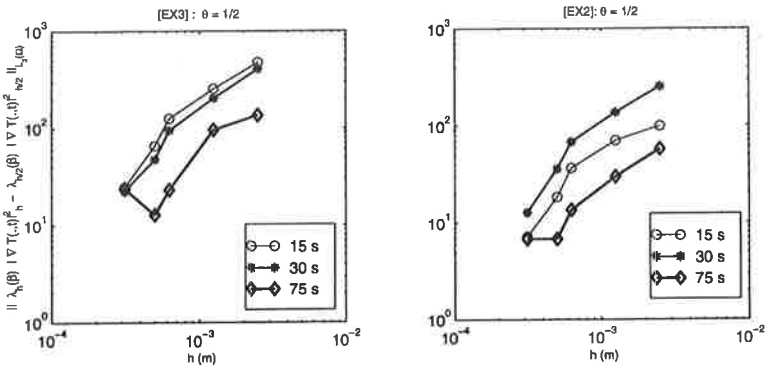


Figure 7: Energy norms of difference of two successive solutions: (a) monotonic BC, (b) cyclic.

interpolation constant not depending on h, and $S_c = (1 + \log \frac{t_n}{\Delta t_n})^{1/2}$ is a stability factor depending on the problem, *cf.* [8] and [19]. An automatic time step refinement strategy based on this error estimate is described in [19]. When f and λ are continuous (no phase changes and ρc not depending on temperature) this error estimate mentioned above is generalized to nonlinear cases [10] and [19].

14 Concluding remarks

The energy equation for a solid two-phase mixture undergoing irreversible phase changes is formulated. The irreversibility condition is used to constrain the constitutive law of the rate of the converted mass fraction. An "enthalpy-like" regularisation (ERTAR) is derived to be used together with a finite element formulation for the energy equation. An adaptive time-integration algorithm, based on controlling simultaneously the increments of temperature and of converted mass fraction, is proposed. At each time step, the nonlinear system of equations, rising from the full discretisation of energy equation, is solved by the Picard fixed point iteration scheme. This iteration scheme may have advantages in regions where the derivatives are discontinuous, since then methods needing derivatives have convergence problems.

15 Acknowledgments

The author wishes to thank Mr. Juha Hartikainen for many interesting and fruitful discussions on the subject. The present work was motivated by a course in non-smooth mechanics [12], organized at Helsinki University of Technology (1997).

References

[1] C. A. Anderson and O. C. Zienkiewicz, "Spontaneous ignition: Finite element solutions for steady and transient conditions", *Journal of Heat Transfer, ASME*, **96** (1974), 398-404.

[2] D. Baroudi, "Application of the finite element method to the Stefan problem in 1-D", in *Proceedings of the* 5^{th} *Finnish Mechanics Days*, Jyväskylä, Finland, 1994, pp. 159-166.

[3] D. Blanchard and M. Frémond, "The Stefan problem: Computing without the free boundary", *International Journal for Numerical Methods in Engineering*, **20** (1984), 757-771.

[4] H. S. Carslaw and J. C. Jaeger, *Conduction of heat in solids*, 2^{nd} ed., Oxford University Press, 1959.

[5] T.-F. Cheng, "A numerical simulation for two-dimensional moving boundary problems with a mushy zone", *Computational Mechanics*, **23** (1999), 440-447.

[6] G. Comini, S. Del Guidice, R. W. Lewis and O. C. Zienkiewicz, "Finite element solution of non-linear heat conduction problems with special reference to phase change", *International Journal for Numerical Methods in Engineering*, **8** (1974), 613-624.

[7] A. W. Date, "Novel strongly implicit enthalpy formulation for multidimensional Stefan problems", *Numerical Heat Transfer, Part B*, **21** (1992), 231-251.

[8] K. Eriksson, D. Estep, P. Hansbo and C. Johnson, "Introduction to adaptive methods for differential equations", *Acta Numerica* (1995), 105-158.

[9] K. Eriksson, D. Estep, P. Hansbo and C. Johnson, *Computational differential equations*, Studenlitteratur, Lund, 1996.

[10] K. Eriksson and C. Johnson, "Error estimates and automatic time step control for nonlinear parabolic problems", *SIAM Journal on Numerical Analysis*, **24** (1987), 12-22.

[11] M. Frémond, "Mixtures and phase changes in solid mechanics", in *Proceedings of the* 3^{rd} *Finnish Mechanics Days*, Otaniemi, Finland, 1988, pp. 1-33.

[12] M. Frémond, *Non smooth mechanics* (lecture notes), course held at Helsinki University of Technology and organized by the Laboratory of Structural Mechanics, 1997.

[13] M. Frémond and M. Mikkola, "Thermomechanical modelling of freezing soil", in *Proceedings of the Sixth International Symposium on Ground Freezing*, Balkema, Rotterdam, 1991, pp. 17-24.

[14] M. Frémond and P. Nicolas, "Macroscopic thermodynamics of porous media", *Continuum Mechanics and Thermodynamics*, **2** (1990), 119-139.

[15] E. Fried and A. Q. Shen, "Generalization of the Stefan model to allow for both velocity and temperature jumps", *Continuum Mechanics and Thermodyamics*, **11** (1999), 277-296.

[16] P. Germain, Q. S. Nguyen and P. Suquet, "Continuum Thermodynamics", *Journal of Applied Mechanics, ASME*, **50** (1983), 1010-1020.

[17] J. Hartikainen, M. Mikkola and R. Orama, *Maan routimisen termomekaaninen malli ja sen laskelmat*, Tielaitoksen selvityksiä 45/1994, Tielaitos, Helsinki.

[18] J. Hartikainen and M. Mikkola, "Thermomechanical model of freezing soil by use of the theory of mixtures", in *Proceedings of the* 6^{th} *Finnish Mechanics Days*, Oulu, Finland, 1997, pp. 1-23.

[19] C. Johnson, *Numerical solutions of partial differential equations by the finite element method*, Studenlitteratur, Lund, 1987.

[20] B. Lucquin and O. Pironneau, *Introduction au calcul scientifique*, Masson, Paris, 1996.

[21] Nelson O. Moraga and Carlos H. Salinas, "Numerical model for heat and fluid flow in food freezing", *Numerical Heat Transfer, Part A*, **35** (1999), 495-517.

[22] K. Morgan, R. W. Lewis and O. C. Zienkiewicz, "An improved algorithm for heat conduction problems with phase change" (short communications), *International Journal for Numerical Methods in Engineering*, **12** (1978), 1191-1195.

[23] R. H. Nochetto, M. Paolini and C. Verdi, "An adaptive finite element method for two-phase Stefan problems in two space dimensions. Part I: Stability and error estimates", *Mathematics of Computation*, **57** (1991), 73-108.

[24] J. Pitkäranta, *Differenssimenetelmät* (lecture notes), Helsinki University of Technology, 1994.

[25] W. Donald Rolf III and Klaus-Jürgen Bathe, "An efficient algorithm for analysis of non-linear heat transfer with phase changes", *International Journal for Numerical Methods in Engineering*, **18** (1982), 119-134.

[26] O. C. Zienkiewicz, *The Finite Element Method*, 3^{rd} ed., McGraw-Hill, London, 1986.

DESIGN RULES AGAINST FRETTING FATIGUE IN CONNECTING RODS

ROGER RABB
Wärtsilä NSD Corporation, P.O. Box 244, FIN-65101 Vaasa, Finland

ABSTRACT

This investigation was initiated after the failure of a connecting rod in a large medium speed diesel engine. An examination of the causes to this accident revealed that a fatigue crack had initiated in the inner surface of the small end of the connecting rod. The crack was located in a region with relatively small cyclic stresses and with a high nominal safety factor against fatigue. This surprising circumstance made it necessary to conduct a thorough nonlinear FE analysis including contact with friction between the bearing bush and the small end. This analysis revealed that cyclic slip occurred between the bearing bush and the small end during every working cycle of the engine. Maximum slip was predicted to occur at the location where the fatigue crack had initiated thus indicating fretting as the most likely crack initiation mechanism. To avoid similar failures in the future a combination of experimental and analytical work was begun to create detailed design rules against fretting.

Experiments were carried out using ring/bushing pairs cut from the small ends of forged production connecting rods. The interference fit between the ring and the bearing was closely controlled and precise measurements of the slippage during the cyclic loading were used to correlate the nominal friction coefficient obtained from nonlinear FE analyses. Various load ranges, ring/bushing material combinations and degrees of interference fit were investigated. It was found that cyclic loading caused a significant increase in the nominal friction coefficient between the mating surfaces. In some cases the surfaces would actually stick together eliminating all slip. Increasing the cyclic load, as might occur during service, resulted in the surfaces being torn apart again. This initiated again significant slip along the ring/bushing contact surfaces. Microscopic analysis assisted in determining the extent of fretting damage. This simple test scheme will hopefully make it possible to understand the factors influencing slip and fatigue crack initiation and to determine threshold values of the friction coefficient and contact pressure and finally to create the desired design rules.

1 CONNECTING ROD FAILURE

During the year 1995 a type of the large medium speed diesel engines designed and manufactured by the Wärtsilä company was faced with two serious connecting rod failures. The engines where these accidents occurred were running electric generators for power produc-

tion and were therefore running at a high load. The appearance of the failed connecting rod is shown in Fig. 1. Figures 2 and 3 of one the failed connecting rods give some idea of the severe consequences of a connecting rod failure.

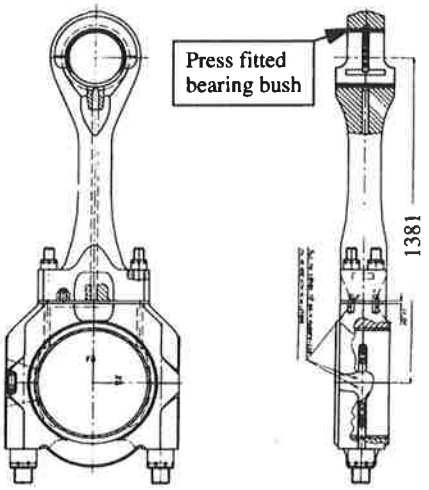


Figure 1. Geometry and design of the failed connecting rod.

Figure 2. Broken connecting rod.

Figure 3. Fracture surface in the small end of the broken connecting rod.

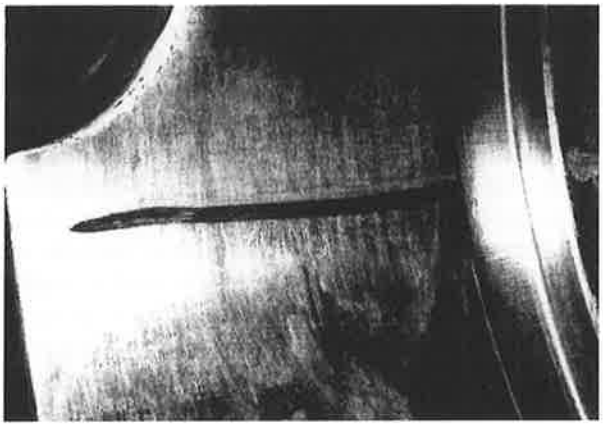


Figure 4. Appearance of a machined groove to test the fatigue resistance of the small end.

1.1 Preliminary examination of the causes to the failure

The failures of these connecting rods came as a total surprise. It was found already in the preliminary check that a fatigue crack had initiated in the inner side of the small end of the con-

necting rod, at a location about 60° from the centre line of the shaft towards the big end of the connecting rod. These connecting rods had been designed and calculated in the midst of the eighties. The design of the small end with a press fitted bearing bush was based on earlier design and there had never been any problems with this design before. The stress distribution in the small end had been calculated with a linear finite element analysis and the calculated safety factor against fatigue was above 2 even in the most critical point. At the position where the cracks had initiated a safety factor of above 3 had been calculated. At the time in the midst of the eighties the available computer memories put severe limitations upon the size of a nonlinear analysis and an analysis with contact and friction between the bearing bush and the small end was not possible at that time.

This failure seemed at first so mysterious that it was decided to carry out an engine test with a sharp ground groove at the observed location of the crack, see figure 4. The theoretical stress concentration factor, of this 3...3.5 mm wide and 1.5 mm deep scratch, was estimated to 2...2.5. The test engine in the laboratory was then run with this connecting rod mounted for 165 hours during which time approximately $2.5 \cdot 10^6$ load cycles were accumulated. In the inspection of the groove after this endurance test no signs of crack initiation was found in the bottom of it. During the same laboratory test the calculated stress field in the small end were measured with several strain gauges, see figures 5 and 6. Nothing unexpected was measured. This test proved that the original stress analyses were reliable and that it was necessary to look for the causes to this crack initiation from something exceptional circumstances.

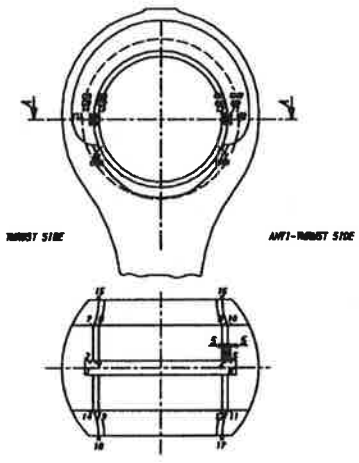


Figure 5. Strain gauge locations.

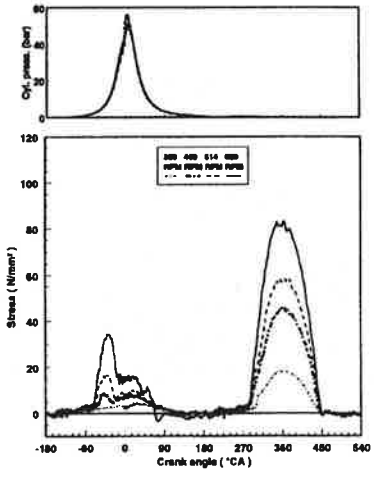


Figure 6. Measured stress at location No. 2.

1.2 Detection of signs of fretting damage

The design of the small end with a press fitted bearing bush was old and well tried and it had never caused any problems before. However, as the examination of the failed connecting rods went on, some evidences began to accumulate which indicated that these failures could be connected to fretting initiated fatigue. In figure 7 the upper part and the lower part of the

fractured small end are matched. Under the microscope it is difficult to detect any clear signs of fretting in the inner side of the small end in the neighbourhood the crack. However, the crack has initiated in the inner surface close to the oil groove in the middle as can be seen from figures 8 and 9. An examination of the backside of the bearing bush showed clear signs of fretting damage as can be seen in figure 10.

Figure 7. Cracked connecting rod. The total length of the crack is about 110 mm.

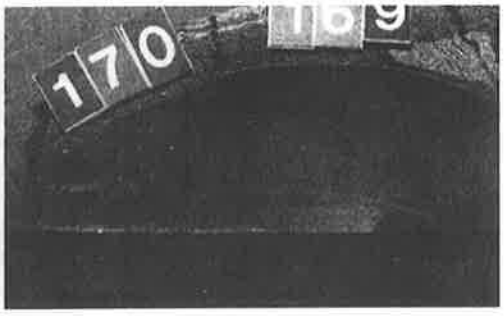


Figure 8. Crack initiation point near to the oil groove.

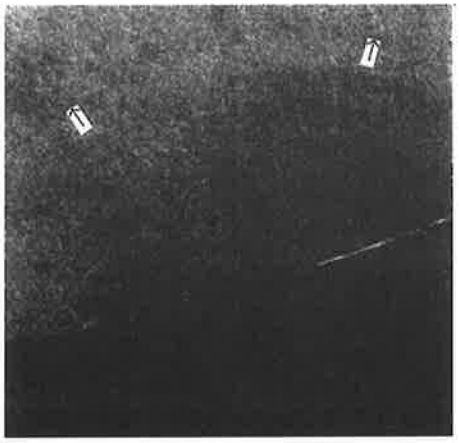


Figure 9. Crack propagation direction.

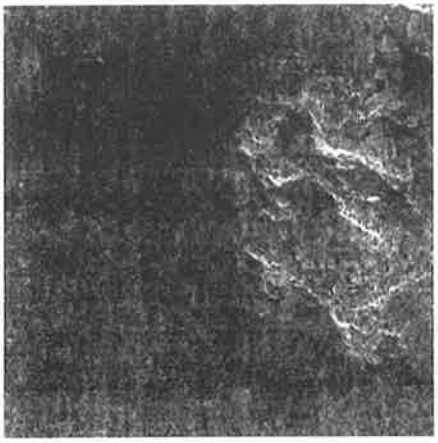


Figure 10. Fretting marks on the outer surface of the bearing close to the oil groove.

2 EXAMINATION WITH A NONLINEAR FE ANALYSIS

Serious suspicions had awaked at this stage that the fatigue crack had been initiated by fretting and not by for instance a material defect or some kind of overload. It was therefore decided to carry out a nonlinear FE analysis with contact and friction included between the small end of the connecting rod and the bearing bush of it. The purpose of this nonlinear FE analysis was above all to calculate the slip, the contact pressure and the shear stress in the contact surface at different positions of the partition surface between the small end and the

bearing bush. At the same time an evaluation of what would be the needed coefficient of friction and the needed interference fit to prevent slippage was done. The nonlinear FE model together with some results are shown in figures 11 and 12.

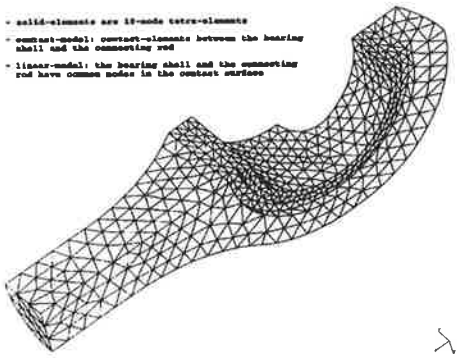


Figure 11. Mesh of the nonlinear FE model with contact and friction.

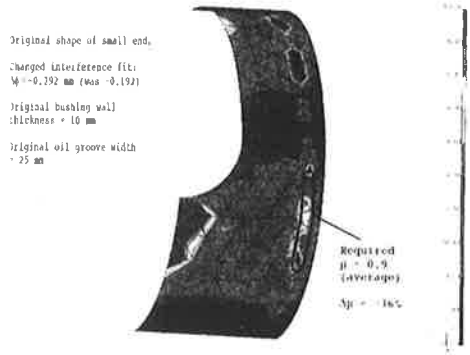


Figure 12. Required coefficients of friction to prevent slippage.

It began now to seem clear that fretting had been the cause to the crack initiation. To get a good model to investigate the distribution of the slip around the periphery of the contact surface the model in figure 13 of a slice of the small end was made. The calculated displacements caused by the inertia loading along the periphery of the two contact surfaces show clearly that the maximum slip will occur at approximately the position where the crack had been initiated, see figure 14.

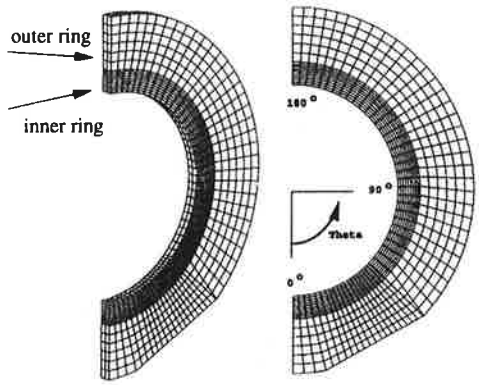


Figure 13. Nonlinear FE model

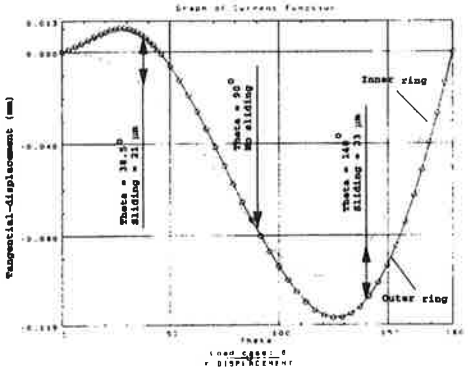


Figure 14. Maximum slip occurs at about $\Theta = 40^{o}$.

3 RING FRICTION TEST

It was obvious at this stage that the fatigue crack had been initiated by fretting. It is reported in the literature that fretting can considerably influence the fatigue limit. For instance R. B. Waterhouse [1] mentions a strength reduction factor of 4.14 for the fatigue limit of a speci-

men of 0.6Cr2.5Ni0.5Mo clamped against the same material. However, the available literature gives very little really useful information about fretting fatigue. Much research work is done but for the time being there is no really reliably theory that could predict the reduction of the fatigue limit due to fretting.

It was decided that, to prevent any further fretting induced cracking in the future, design rules had to be created of how to avoid this phenomena. The main emphasize should be on determining the following issues:

1) the determination of the true coefficients of friction in the contact plane between the bearing bush and the small end of the connecting rod
2) the determination of the needed interference fit to prevent slip
3) the determination of the threshold value for the contact pressure, below which a surface crack will not initiate

3.1 Test arrangement

To make the test as realistic as possible it was decided to cut the small end of a connecting rod with mounted bearing shell into slices with a width of 10 mm. These slices were loaded with a point load which gave approximately the same stress distribution and slip conditions in the partition plane as during operation of the engine. The test arrangement is illustrated in figures 15 and 16.

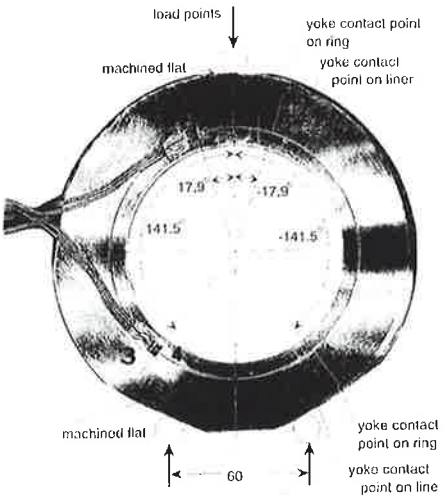


Figure 15. Test specimen with strain gauges in the ring friction test. Load and supports are sketched.

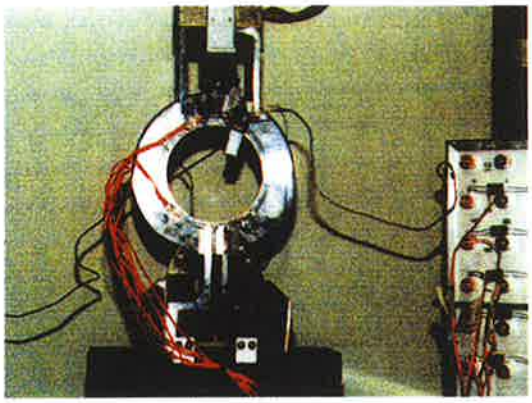


Figure 16. Slippage was also measured directly with an extensometer.

The shrinkage fit between the inner and the outer ring was ϕ105 s6/H6. The friction surface of the outer ring was machined and it was made of quenched and tempered steel 34CrNiMo6/

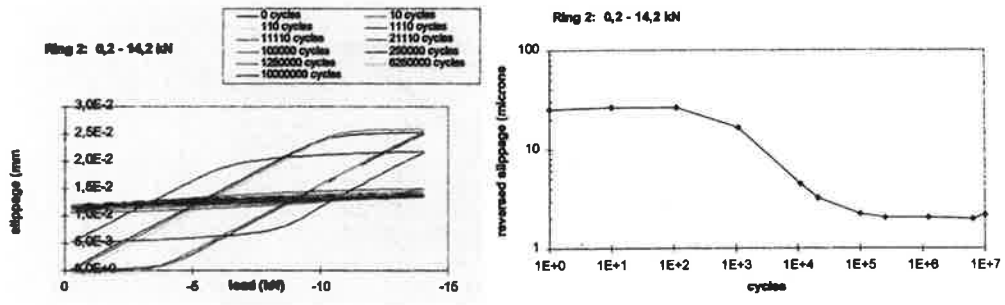
EN 10083-1. The inner ring was made of soft steel St3 with a tin flash on its outer surface. The surface finishes in the partition plane before testing were R_a/R_z = 0.8/7.2 µm for the inner ring and 0.8/3.9 µm for the outer ring.

In addition to the realistic test situation there were some further advantages of this test arrangement. It was possible to measure the influence of the slip both indirectly by measuring the stress in the bearing shell at the locations with expected maximum slip and also by measuring the slip directly with an extensometer as shown in figure 16. By comparing the measured values to the corresponding calculated values of the nonlinear FE analysis it was possible to make an accurate estimation of the coefficient of friction at the slip location. Before the tests it was not known how well the proposed ways of measuring would work. However, both ways of measuring worked well and this gave further confidence in the test arrangement.

The two MTS model 632.26 extensometers were fixed between the yokes so that one extensometer knife edge was supported by the yoke fixed to the ring and the other knife edge was supported by the yoke fixed to the liner. These extensometers had a nominal gage length of 8 mm and were calibrated so that 1 µm displacement of the knife edges corresponded to a signal of 62.5 mV and a theoretical resolution of better than 0.01 µm. These tests were made at VTT, the Finnish Research Centre in Espoo.

3.2 Results of the test

In terms of slippage between ring and liner, both ways of measuring show a very similar trend, figures 17 and 18. The degree of slippage is constant or even slightly increasing during the first 100 cycles. This value decreases significantly, by nearly a factor of 10, during the period between 1000 and 100000 cycles. Between 10^5 and 10^7 cycles, the degree of slippage remains virtually constant. Closely associated with the change in slippage there is a dramatic reduction in the strain hysteresis for each of the strain gages. This is consistent since an increase in surface roughness and friction between the ring and liner is expected to result in a more and more linear relation between stress and load.

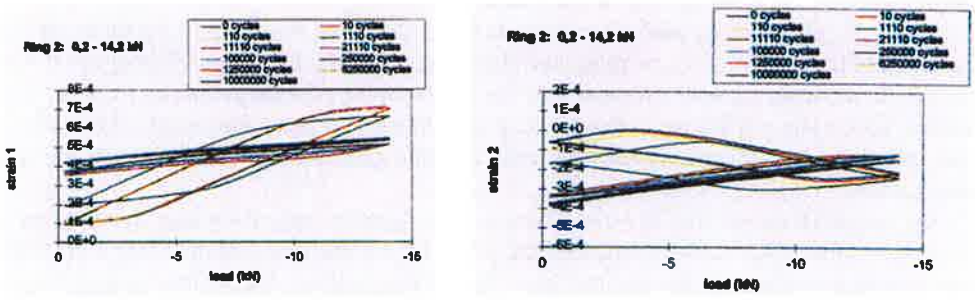


a) slippage loops

b) change in slip

Figure 17. Slippage as a function of the number of cycles measured with the extensometers. Machined ring against a liner with tin flash. Load range from 200 to 14200 N.

One important matter that can be concluded at once from the result is that the coefficient of friction is growing with the number of cycles. This is probable a following of the increasing adhesion between the mating surfaces when the oxide layer is destroyed by the slippage process. As a following of this increasing coefficient of friction it is possible to imagine that there must be situations where the slip must be halted after a sufficient number of slip cycles.



a) ring b) liner

Figure 18. Measured strain hysteresis curve corresponding to the case in figure 17.

3.3 Evaluation of the coefficient of friction

The actual coefficient of friction between the mating surfaces was evaluated by comparing the measurements to the slippage loops calculated in a nonlinear friction and contact analysis. The FE model of the friction ring is shown above in figure 13. The measured slippage and the calculated one at different coefficients of friction are shown in figure 19.

The agreement between the calculated and measured slippage loops is excellent with the adapted coefficients of friction. It can be concluded that the coefficient of friction is in the beginning close to a value of $\mu = 0.16$ for this surface combination. When the tin flash layer was removed from the liner the starting value of the coefficient of friction increased to $\mu = 0.21$, see figure 20.

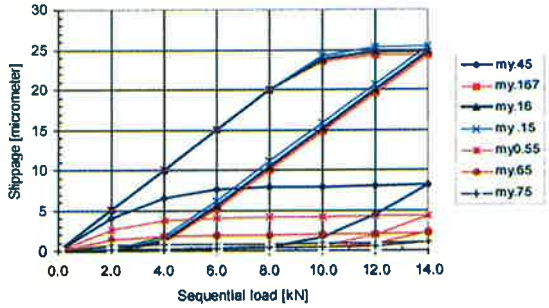
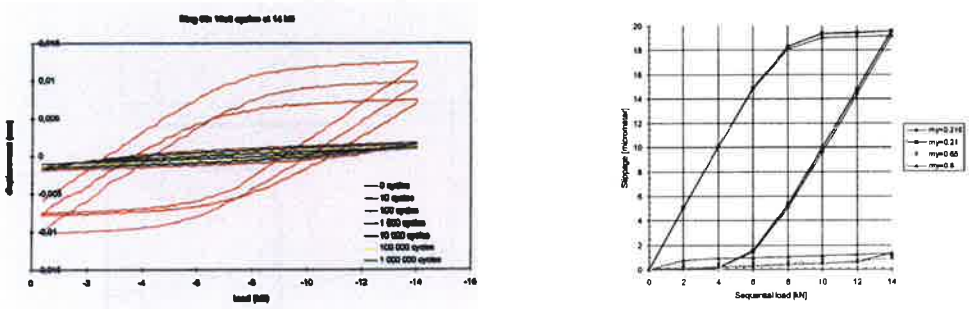


Figure 19. Comparison of measured and at different coefficients of friction calculated slippages for the measurement with a machined ring against a liner with tin flash shown in fig. 17.

A surface combination of shot peened ring against liner with tin flash was also tested. It verified the assumption that shot peening, in addition to provide a beneficial influence on the

fatigue strength, also increase the coefficient of friction. The evaluated coefficient of friction was in this case also about $\mu = 0.21$ as for the case with machined ring against liner with the tin flash removed. The evaluated coefficients of friction for different surface combinations are summarized in table 1.



a) measured b) calculated

Figure 20. Measured and calculated slippage when the tin flash is removed from the liner. Evaluated coefficient of friction $\mu = 0.21$.

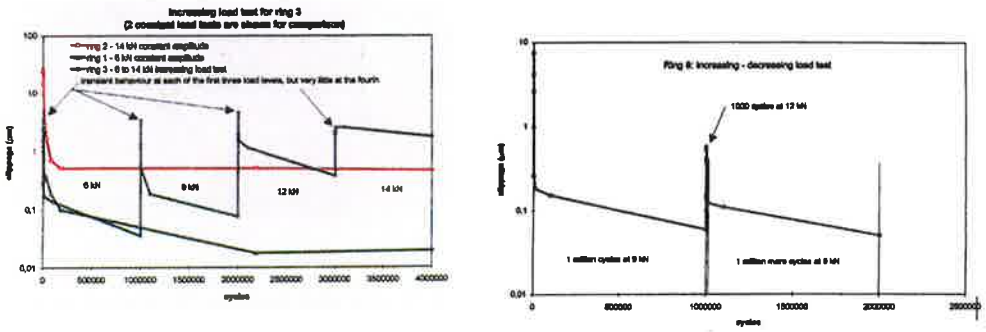
A test with stepwise increasing load and with a total of 10^6 cycles at each load level was also conducted for the surface combination machined steel against tin plated soft steel. A tendency that a increasing load tore apart the surfaces and increased the slip from the low level it had reached at the end of the preceding load level could be observed. However, the slippage levels remained quite low at every increase of the load and showed the decreasing tendency as observed in the tests with constant load ranges. However, also a tendency for a higher and higher remaining slippage at every increase of the load was observed, figure 21a. A similar behaviour can be observed in the increasing-decreasing test shown in figure 21b for a surface combination machined ring against a liner with the tin flash removed. It can therefore be concluded that a halted slippage can be reborn by an occasional overload.

Table 1. Evaluated coefficients of friction .

Surface combination	Interference fit	Evaluated coefficient of friction
Machined steel 34CrNiMo6 TQ+T against tin plated soft steel St3	Normal s6/H6	0.16
Machined steel 34CrNiMo6 TQ+T against soft steel St3 with removed flash layer	Normal s6/H6	0.21
Shot peened steel 34CrNiMo6 TQ+T against tin plated soft steel St3	Normal s6/H6	0.21

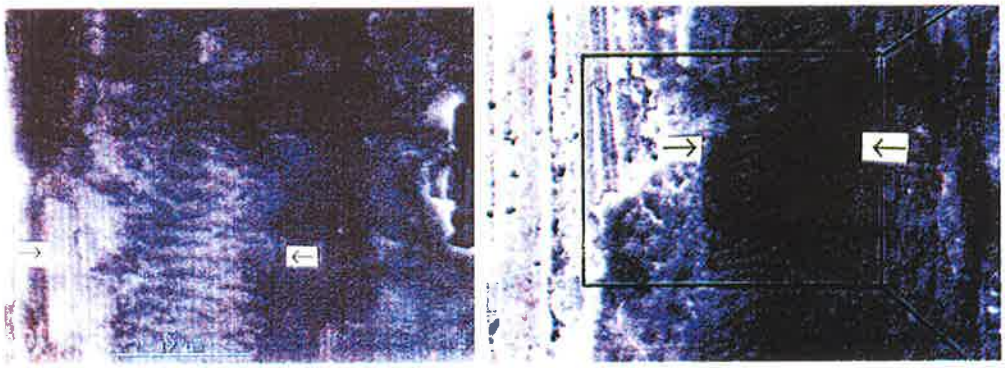
An important observation was that with the chosen cutoff limits for the load cycling, none of the specimens tested failed by fretting or showed any signs of cracking. According to the

SEM investigation the real contact area between the mating surfaces is mostly concentrated to the asperities of the machining marks. In these areas transfer of the liner flash material (tin Sn) to the ring surface is found. Indications of slippage rest lines in the surfaces were found in the direction $\theta \approx 40^{o}$, figure 22. The distance between adjacent rest lines is approximately about 1.8 ... 2.2 µm. The areas containing slippage rest lines were not found to contain any indications of micro cracks.



a) increasing load test b) increasing-decreasing load test

Figure 21. Surface combination machined ring against a tin plated liner in both an increasing and an increasing-decreasing load test.



a) ring surface b) liner surface

Figure 22. Micro slippage area after 10^7 cycles for the specimen for which the corresponding slippage loops are shown in figure 17.

4 CONCLUSIONS

Fretting initiated fatigue failures are a major problem for the medium speed diesel engine. Connecting rod failures, as the one that initiated this investigation, are particularly devastating. It is therefore extremely important to learn to understand and prevent these types of fatigue failures. Regrettably, in spite of much effort nowadays to examine this phenomena and a lot of published books [2,3,4,5,...] and articles [6,7,8,9,10,11,...], there are little completely useful design rules available. Because of this lack of a detailed understanding of fretting in-

itiated fatigue, the best way available is to try to determine the critical parameters for the application in question. Among these critical parameters is of course the coefficient of friction. If the true value of it is known, it would in principle be possible to make the design in such a way that no slippage can occur. In reality, this is not always possible. Under these circumstances it is very important to investigate if there for instance exists a threshold contact pressure below which no fretting damage will occur despite of the slippage.

The true meaning of the growing coefficient of friction dependent on the number of cycles is also unclear. Could it be that the slippage is halted before a critical surface crack has initiated? This question should also be investigated with the goal to be able to take it into account already in the design stage. The number of cycles needed to initiate a fretting crack would also be important to know. The experiences from fretting initiated fatigue failures from the field indicate that usually a very high number of cycles is needed to initiate a crack. The number of needed cycles is in some cases reported to be hundred of millions, whereas in normal fatigue only some millions of cycles are needed to initiate a propagating crack.

The ring friction test program corresponded to the expectations put on it. It is possible with this method to measure the slip in the partition surface between the ring and the liner very accurately. It is also easy to make a nonlinear FE analysis of this simple test specimen and by this determine the actual coefficient of friction. The purpose is to go on with this test program in a trial to determine the above mentioned open questions.

REFERENCES

[1] R. B. Waterhouse. Fretting Corrosion. Pergamon press. Biddles Ltd., Guildford, Surrey 1975.

[2] D.A. Hills and D. Nowell. Mechanics of Fretting Fatigue. Kluwer Academic Publishers 1994.

[3] K.L. Johnson. Contact Mechanics. Cambridge University Press 1985.

[4] F.F. Ling and C.H.T. Pan. Approaches to Modeling of Friction and Wear. Springer-Verlag New York Inc. 1988.

[5] ESIS Publication 18, Fretting Fatigue. Edited by R.B. Waterhouse and T.C. Lindley. European Structural Integrity Society 1994.

[6] D.A. Hills. Mechanics of fretting fatigue. Wear, 175 (1994) 107-113.

[7] D. Nowell and D.A. Hills. Crack Initiation Criteria in Fretting Fatigue. Wear, 136 (1990) 329-343.

[8] Saeed Adibnazari and David W. Hoeppner. A fretting fatigue normal pressure threshold concept. Wear, 160 (1993) 33-35.

[9] R.D. Mindlin and H. Deresiewicz. Elastic Spheres in Contact under Varying Oblique Forces. Journal of Applied Mechanics. 1953, 20, 327-344.

[10] C.A. Bergman, R.C. Cobb and R.B. Waterhouse. The Effect of Shot-peening on the Friction and Wear of a Carbon Steel and an Austenitic Stainless Steel in Fretting Conditions. Wear of Materials 1987. ASME Conference Houston Texas.

[11] Ernest Rabinowicz. The Nature of the Static and Kinetic Coefficients of Friction. Journal of Applied Physics. Volume 22, Number 11 November 1951.

RASKASKULJETUSTEN ANALYSOINTI

Ilari Silvola, Jyri-Pekka Arjava, PI-Rauma
Timo P.J. Mikkola, Timo Kukkanen, VTT Valmistustekniikka

TIIVISTELMÄ
Syvänmeren öljyntuotantoon tarkoitetut SPAR rakenteet kuljetetaan Porista Meksikonlahdelle raskaskuljetusaluksen kannella. Rakenteiden mitoittamiseksi kuljetusten aikaisia kuormituksia vastaan on kehitetyssä laskentaproseduurissa merenkäynnin aiheuttamat kuormitukset lasketaan nk. 3D paneelimenetelmällä ja rakenteen jännitystasot 3D FE-mallilla. Laskentamallien tuottama valtava tulosmäärä saadaan tehokaasti tiivistettyä spektraalimenetelmään perustuvan tilastollisen tarkastelun avulla äärilujuus, lommahdus ja väsymistarkasteluihin soveltuvaksi. Analyysin useita työläitä vaiheita on voitu automatisoida.

Käytännössä rakennesysteemi ei toimi lineaarisesti ja lineaarisuusoletus johtaa helposti ylikonservatiiviseen mitoitukseen erityisesti ääritilanteen osalta. Usein systeemin epälineaarinen käyttäytyminen nimenomaisesti estää ylikuormittumisen ja tekee koko kuljetuksen mahdolliseksi. Kuljetusanalyysiryhmä on onnistunut kehittämään toimivan laskentaproseduurin, jossa käytettävillä linearisoinneilla ja mallitustavoilla päädytään järkevään lopputulokseen. Työn aikana on useita valittuja ratkaisuja jouduttu perustelemaan ja osoittamaan niiden luotettavuus loppuasiakkaalle. Käytännössä toteutetut kuljetukset ovat osoitus myös onnistuneista kuljetusanalyyseista.

1 JOHDANTO

Tehokas ja kannattava öljyntuotanto on mahdollista yhä syvemmistä vesistä ja valtaosa öljyntuotannon lisäkapasiteetista tulee jo nyt syvän veden alueelta. Eräs uusi öljyntuotannon konsepti on SPAR rakenne, jossa tuotantotasanne on runkorakenteen muodostavan yhden suunnattoman pystyssä kelluvan lieriön päällä. SPAR konseptilla useita hyviä ominaisuuksia, jotka liittyvät rakenteen stabiiliuteen merenkäynnissä sekä nopeuteen, jolla päästään rakennusvaiheesta tuotannon aloittamiseen.

Öljyntuotannossa tai valmisteilla on useita SPAR rakenteita, joiden kaikkien runkorakenteen valmistus ja installointi on ollut Aker Rauma Offshore OY:n vastuulla. Runkorakenteet on valmistettu Aker Mäntyluodon telakalla Mäntyluodossa ja PI-Rauma OY:n vastuulla on ollut merkittävä osa rakenteiden suunnittelusta. SPAR rakenteen lieriöosa on suurimmillaan ollut halkaisijaltaan n. 40 m, pituudeltaan n. 200 m ja painoltaan yli 30 000 tn. Tällaisen rakenteen valmistus sinänsä asettaa jo suuria vaatimuksia tuotannolle ja suunnittelulle. Lisäksi öljyntuotantoon liittyvien rakenteiden suunnittelussa käytetään tiukkoja turvallisuuskriteereitä, jotka entisestään lisäävät projektityön vaativuutta.

Eräs osa projektin suunnittelua on osoittaa, että SPAR rakenne kestää vaurioitumatta kuljetuksenaikaiset rasitukset. Kuljetus Mäntyluodosta sijoituspaikalleen Meksikon lahdelle tehdään raskaskuljetusaluksilla. Suurimmat valmistetut SPAR rakenteen runko-osat on jouduttu kuljettamaan kahdessa osassa. Kuljetusanalyysit on tehty yhteistyössä PI-Rauma OY:n ja VTT:n kesken Aker Rauma Offshore OY:n osallistuessa työn ohjaukseen ja valvontaan. Lisäksi loppuasiakkaat, öljy-yhtiöt, valvovat työn etenemistä.

Kuljetusanalyysissa tarkastellaan SPAR rakenteen ja aluksen selviämistä äärimmäisissä olosuhteissa sekä matkan aikaisten aaltokuormien aiheuttamaa väsyttävää rasitusta. Aaltokuormien aiheuttama aluksen ja sen lastin liiketila lasketaan hydrodynamiikan ohjelmilla nk. 3D paneelimenetelmällä. Aaltokuormien aiheuttamat rasitukset siirretään 3D FE menetelmän rakennemalliin kiihtyvyys- ja painekuormina. Laskentatulokset käsitellään taajuustason tilastollisen tarkastelun avulla nk. spektraalimenetelmällä, (Kukkanen 1996), lujuus- ja väsymistarkastelua varten.

Laskentaproseduuri on kehitetty PI-Rauma OY:n, VTT:n ja Aker Rauma Offshore OY:n välisessä yhteistyössä, (Mikkola & al. 1998). Pääosa laskennasta tehdään kaupallisilla ohjelmistoilla mutta tärkeä osa laskentaproseduuria ovat lukuisat omat tiedonsiirto- ja tulosten prosessointiohjelmat, jotka tekevät massiivisen laskennan hyödyntämisen toisaalta käytännössä mahdolliseksi ja toisaalta muuntavat lasketut 'raakatulokset' käytännön analyyseihin sopiviksi.

2 RAKENTEIDEN KUORMAT KULJETUKSEN AIKANA

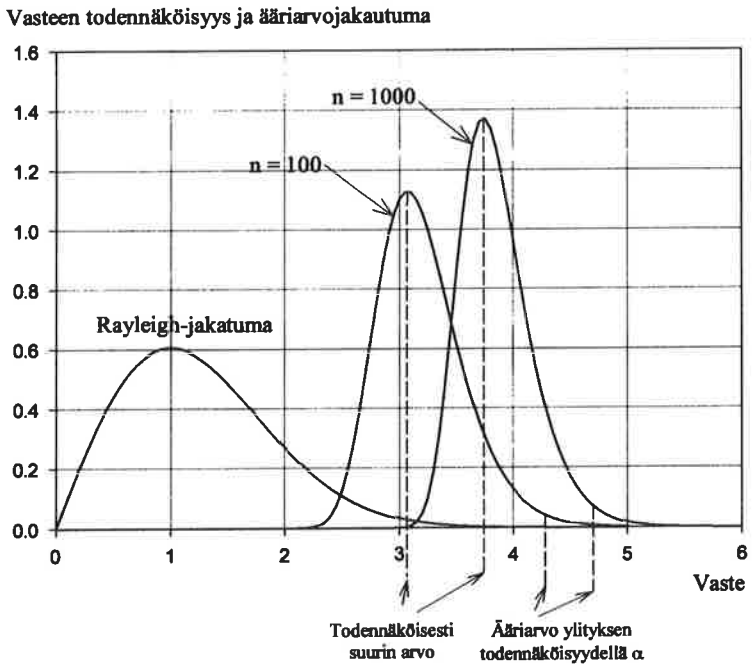
Kuljetuksen aikana SPAR rakenne on tuettu aluksen kannelle kahdella pituussuuntaisella tukirakenteella lähes koko SPAR rakenteen matkalta. Tukirakenteiden suunnittelulla pyritään jakamaan syntyvät viivakuormat mahdollisimman leveälle sektorille SPAR rakenteen kehäkaarille ja toisaalta aluksen kaarille. Toteutetuissa kuljetuksissa SPAR rakenteen asema ja suunta aluksen kanteen nähden on vaihdellut ja lisäksi aluksen ja SPAR rakenteen kaarijaot ovat aina erilaisia. Tukirakenteen on siis kyettävä jakamaan välittyviä kuormia myös tukirakenteen pituussuunnassa.

Kuljetuksen aikainen staattinen kuormitus aiheutuu pääasiassa lastin painosta ja dynaaminen rasitus aaltokuormista. Aaltokuormien periodi on 5 - 15 s luokkaa, joten kuormitus voidaan hyvin käsitellä kvasistaattisena. Aluksen ja SPAR rakenteen sekä niitä yhdistävän tukirakenteen joustavuus ja massajakauma määräävät kuormitukset kuljetuksen aikana.

Lujuusmitoituksessa äärikuormitustilanteena käytetään merenkäyntitilannetta, joka määräytyy kuljetukselle asetetusta riskitasosta P_f ja kuljetusreitin merenkäyntitilastoista. Riskitaso ilmoittaa todennäköisyyden, jolla aluksen sallitaan joutuvan kerran kuljetuksen aikana mitoitustilanteena käytettyä pahempaan merenkäyntitilaan. Yleinen käytäntö on tarkastella merenkäyntiä 3 tunnin vakiotilanteiksi oletetuissa jaksoissa ja kuvata sitä merkitsevän aallonkorkeuden, -taajuuden ja aaltospektrin avulla. Olettamalla aluksen kohtaamien merenkäyntien riippumattomuus ja se, ettei kapteeni lue sääennusteita tai välitä keliolosuhteista reittiä valitessaan saadaan mitoitusmerenkäynnin aallonkorkeus laskettua

matkan aallonkorkeustilastojen ja riskitason avulla, Kuva 1. Mitoitusmerenkäynnin taajuus valitaan siten, että se aiheuttaa pahimmat rasitukset rakenteille ja toisaalta tilastoista löytyväksi. Valittu mitoitustilanteen määrittely on yksinkertainen ja suhteellisen selkeä, mutta se johtaa nopeasti suurilta tuntuviin mitoitusaallonkorkeuksiin. Käytännössä aluksilla on käytettävissään luotettavat lyhyen ajan sääennusteet, joiden avulla raskaiden merenkäyntiolosuhteiden riskiä voidaan olennaisesti pienentää. Ongelmana on, ettei tämän reunaehdon huomiointi mitoitustilanteen määrittelyssä ole kovin yksinkertaista. Käytännössä alkuperäisestä mitoitusmerenkäynnistä on joissakin kuljetuksissa rajoitetusti tingitty juuri tähän vedoten.

Väsyttävä kuormitus syntyy kaikista kuljetuksen matkan aikana kohtaamista aallokoista. Nämä saadaan matkareitin aaltotilastojen perusteella.



Kuva 1 Mitoitusmerenkäynnin kehittyminen aluksen kohtaamien merenkäyntitilanteiden lukumäärän n kasvaessa. Kuvassa vaste on mitoitustilanteen merkitsevä aallonkorkeus.

3 LASKENTAPROSEDUURI

Rakenteen kuormat mitoitusmerenkäynnissä laskentaan hydrodynamiikan mallilla. Teoriassa mitoituksessa käytetään kullekkin rakenneosalle mitoitussuureen todennäköisintä maksimiarvoa 3 tunnin mitoitusmerenkäynnissä. Mitoitussuure on joko lommahduksen käyttökerroin tai maksimijännitys. Muuttujina ovat merenkäynnin suunta alukseen nähden ja merenkäynnin periodi. Käytännössä laskennassa valitaan jokin referenssivaste, jolle 3 tunnin ääriarvo lasketaan. Rakenne analysoidaan kuormitustilanteessa, joka aiheuttaa referenssivasteelle 3 tunnin ääriarvon.

Ensimmäisissä kuljetusanalyyseissa referenssivaste valittiin SPAR rakenteen jäykän kappaleen liiketilaa kuvaavista suureista, kuten vaakakiihtyvyys tai kallistuskulma. Näiden suureiden avulla määritettiin useita säännöllisiä mitoitusaaltoja, joiden avulla eri rakenneosien lujuustarkastelu tehtiin. Viimeisissä kuljetusanalyyseissa lujuustarkastelut on tehty suoraan jokaisen tukirakennealueen jännityskomponenttien 3 tunnin ääriarvojen perusteella. Tukireaktiolle laskettujen ääriarvojen perusteella laskenta on voitu rajata yhteen aallokon suuntaan ja periodiin kullakin tukialueella. Laskenta perustuu nk. spektraalimenetelmään, jossa lasketaan aluksen jännitysvaste säännöllisissä yksikön korkuisissa aalloissa, joiden pituus (periodi) ja suunta alukseen nähden vaihtelee. Viimeisimmissä kuljetusanalyyseissa käytetty menetelmä tiivistyy lujuusanalyysin osalta yhden kuormitustilanteen tarkasteluksi, joka vastaa hyvin suurella tarkkuudella haluttua teoreettista 3 tunnin äärivastetta. Aiemmissa, mitoitusaaltoon perustuvissa menetelmissä, huonona puolena oli tarve laskea useita kuormitustilanteita ja toisaalta niiden ja halutun teoreettisen kuormitustilanteen huono vastaavuus.

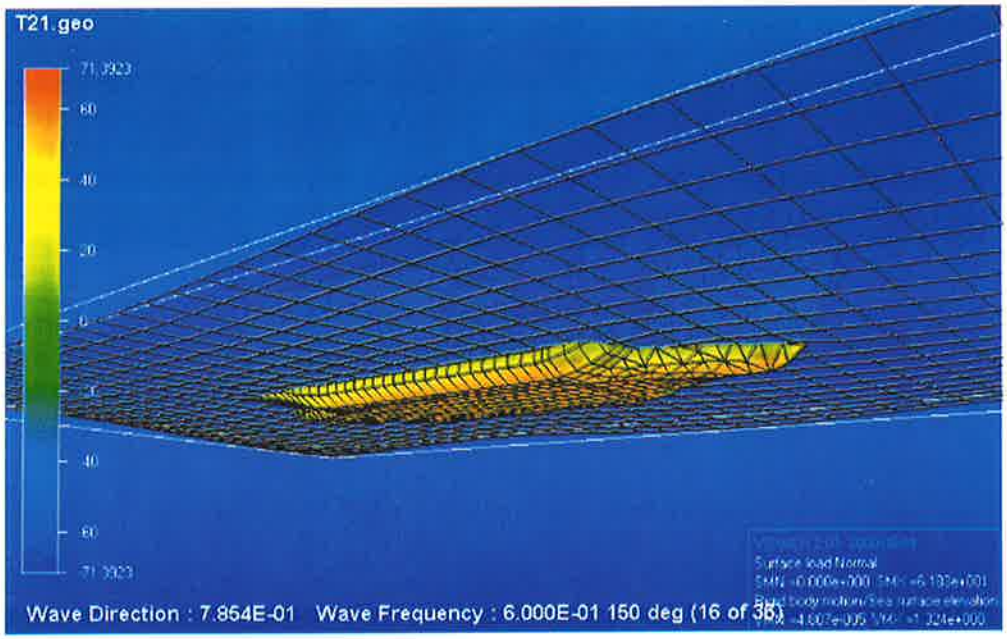
Väsymisen suhteen mitoituksessa käytetään aluksen kuljetusreitin aaltotilastojen mukaista kuormitusta. Laskenta perustuu spektraalimenetelmään, jossa yksikköaalloille laskettujen jännitysvasteiden ja kuljetusreitin yhdistetyn aaltotilaston perusteella lasketaan jokaiselle tarkastelupisteelle erikseen väsymisvaurio. Väsymisanalyysin ongelmana on, että laskentamallin jännitykset vastaavat lähinnä nimellisjännitystä, jonka yhteys eri rakenneosien liitosten väsymisvaurioon on epämääräinen. Laskennassa otettiin alusta lähtien käytännöksi tehdä koko rakennetta kuvaavan nk. globaalimallin avulla screening väsymistarkastelu. Tämän tarkastelun avulla saadaan eri rakenneosille arvioitua paikallisten jännityskonsentraatiokertoimien sallitut raja-arvot ja toisaalta eri rakennealueiden suhteellinen kriittisyys väsyttävän kuormituksen suhteen. Tulosten perusteella tehdään varsinainen väsymistarkastelu rajatulle määrälle rakenteen detaljeja. Tässä tarkastelussa käytetään huomattavan paljon tarkempaa laskentamallia.

4 AALTOKUORMIEN ANALYSOINTI

Raskaskuljetusalukset ovat tyypillisesti leveitä ja suhteellisen hitaasti kulkevia aluksia. SPAR rakenteen paino rajaa mahdollisten kuljetusalusten määrää, mutta toisaalta se ei sinänsä ole lähellä suurten raskaskuljetusalusten kapasiteettia. Rakenteen pyöreä ja pitkä muoto on kuljetuksen kannalta kuitenkin hankala. Kuljetettava massa nousee suhteellisen korkealle ja systeemin keinuntahitaus on melko pieni. Tuloksena on helposti kuljetuksen

suuri keinuntakulma sivuaallokoissa ja sen vuoksi lastin suuri vaakakiihtyvyys mitoitusmerenkäynnissä.

Aaltokuormien analysointi on tehty paneelimenetelmään perustuvalla AQWA ohjelmistolla, Kuva 2. Rakenteen mallitus ja varsinainen laskenta ei nykyisellä tietokonekalustolla ole suuri ongelma ja laskenta-ajat pysyvät kohtuullisina. Varsinainen ongelma analyysissa liittyy keinuntaliikkeeseen ja sen vaimennukseen. Lujuus- ja väsymisanalyysit halutaan tehdä taajuustasolla, jolloin joudutaan käyttämään linearisoitua laskentamallia. Keinuntaliikkeen vaimennus on voimakkaasti liikkeen amplitudin funktio lähellä keinunnan resonanssialuetta, joka käytännössä aina osuu analysoitavalle taajuusalueelle. Keinuntaliike myös vahvasti määrää rakenteen kuormituksia mitoitustilanteessa.



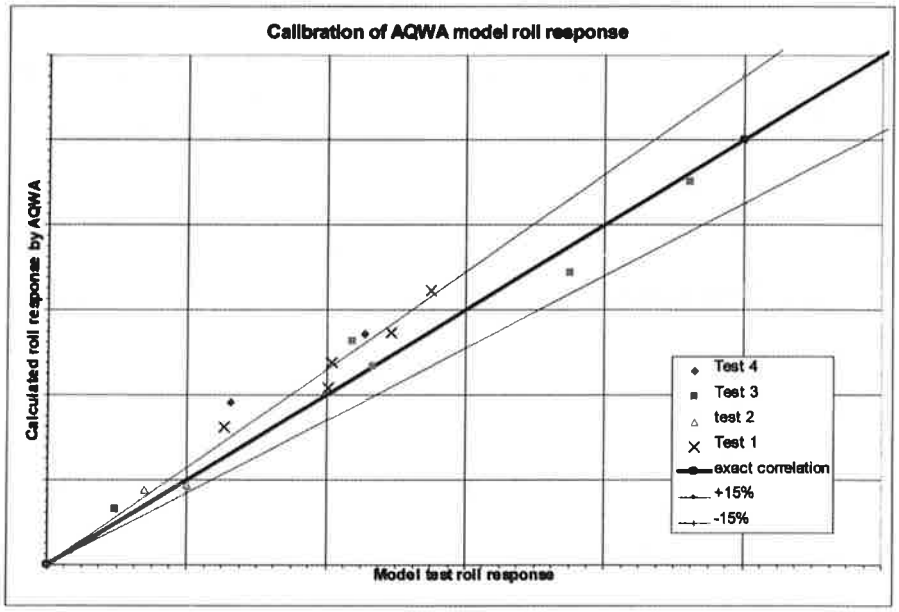
Kuva 2 Kuljetuksen laskentamalli.

Keinuntaliikkeen vaimennus kuljetusalukselle määritetään erilaisissa lastitilanteissa tehtyjen mallikokeiden avulla. Käytännössä kokeet tehdään lastitilanteita vaihdellen epäsäännöllisessä aallokossa sivuaallokossa muutamilla aallonkorkeuksilla. Näiden tulosten avulla lasketaan vaimennus muissa aaltotilanteissa. Kuljetusanalyysien yhteydessä keinunnan vaimennuksen mallittamiseksi tehdään aikatason analyysi mallikokeita vastaavissa olosuhteissa ja kalibroidaan mallin vaimennus tulosten vastaavuuden saavuttamiseksi. Mallikokeita ei yleensä ole tehty suoraan kuljetusta vastaavalle lastitilanteelle. Tulosten perusteella aluksen vaimennus keinuntaliikkeelle saadaan kuitenkin kokemusten perusteella hyvin hallintaan, Kuva 3, ja ekstrapolointi kuljetustilanteeseen voidaan tehdä luotettavasti.

Laskentamallin kalibrointi on kullekin alukselle kertaluontoinen tehtävä. Mikäli samaa alusta käytetään uudessa kuljetuksessa, voidaan aiemmassa kalibroinnissa saatuja tuloksia suoraan soveltaa uudena lastitilanteen analysointiin.

Lujuus- ja väsymisanalyysia varten joudutaan erikseen valitsemaan aaltotilanne, johon vaimennuksen linearisointi tehdään. Ääritilanteen mitoitusta varten riittää kalibrointi mitoitusaallonkorkeuteen. Väsymisanalyysia varten tehdään keinunnan kalibrointi aallonkorkeudelle, jonka kontribuutio väsymisvaurioon on suurin. Tämä voidaan arvioida aaltotilastojen ja jännitysten ja aallonkorkeuden lineaarisen korrelaation perusteella.

Useat kuljetetut SPAR rakenteet on jouduttu sijoittamaan aluksen kannelle siten, että ne ulottuvat pitkälle aluksen partaan yli. Seurauksena on, että SPAR rakenteen ulkovaippa voi joutua alttiiksi aaltoiskuille. Tämän kuormitustilanteen analysointi on edelleen suuri haaste jo teorian tasolla eikä siihen ole tarjolla luotettavia ohjelmia. Tähänastisissa kuljetusanalyyseissa on tavoitteena ollut laskea merkitsevän aallonkorkeuden raja aallokon suunnan funktiona siten, että rajoitetaan SPAR rakenteen altistuminen aaltoiskuille. Analysointi tehdään laskien aluksen liiketilaa epäsäännöllisessä aallokossa ja tarkastellen SPAR rakenteen vaipan eri alueiden ja aallokon suhteellista liikettä. Analysoitava matriisi erilaisia aaltotilanteita ja tarkasteltavia pisteitä kasvaa nopeasti kohtuullisen suureksi. Käytännössä laskennan avulla on pystytty tuottamaan ohjeelliset aallokon suunnasta riippuvat kelirajat, joita noudattamalla vältetään aaltoiskukuormitus SPAR rakenteeseen.



Kuva 3 Erään kuljetusaluksen mallikoe- ja laskentatulokset keinuntaliikkeelle.

5 RAKENTEEN MALLITUS

Kuljetuksen aikana raskaskuljetusaluksen ja SPAR rakenteen muodostamaa kokonaisuutta kuormittavat aallokon aiheuttamat rasitukset. SPAR rakenteeseen kuormat välittyvät aluksen kannella olevien tukirakenteiden kautta. Eri aaltotilanteissa syntyvä tukivoimajakauma määräytyy toisaalta aluksen ja kannen jäykkyydestä ja toisaalta tukirakenteiden ja itse SPAR rakenteen jäykkyydestä. Globaali taivutusjäykkyys on SPAR rakenteella noin viisinkertainen laivapalkin taivutusjäykkyyteen nähden. Toisaalta molemmat rakenteet muodostuvat jäykistetyistä levykentistä joiden kehyskaariväli on erilainen. Tukirakenteiden suunnittelulla pyritään välttämään voimakkaita kuormituskeskittymiä, joiden täydellinen välttäminen on kuitenkin mahdotonta.

Kuljetuksen aikana esiintyvien kuormien suuruus ja jakauma riippuu SPAR rakenteen, tukirakenteiden seka aluksen jäykkyyksistä ja massajakaumista. Lisäksi aluksen uppoumajakauma vaikuttaa erityisesti staattiseen eli tyynen veden voimajakaumaan. Sekä SPAR rakenteen että aluksen jäykkyys muodostuu rakenteiden käyttäytymisestä globaalisti että lokaalisti. Globaalia käyttäytymistä voidaan mainiosti kuvata palkkimallilla. Toisaalta lokaali käyttäytyminen, joka muodostuu aluksen kannen sekä SPAR rakenteen kehäkaarien paikallisesta joustosta, vaikuttaa voimien jakautumiseen olennaisesti. Tämän kuvaaminen voidaan periaatteessa tehdä jousien avulla mutta jousivakioiden määritys tekee tästä lähestymistavasta käytännössä helposti yhtä hankalan kuin suora 3D mallitus.

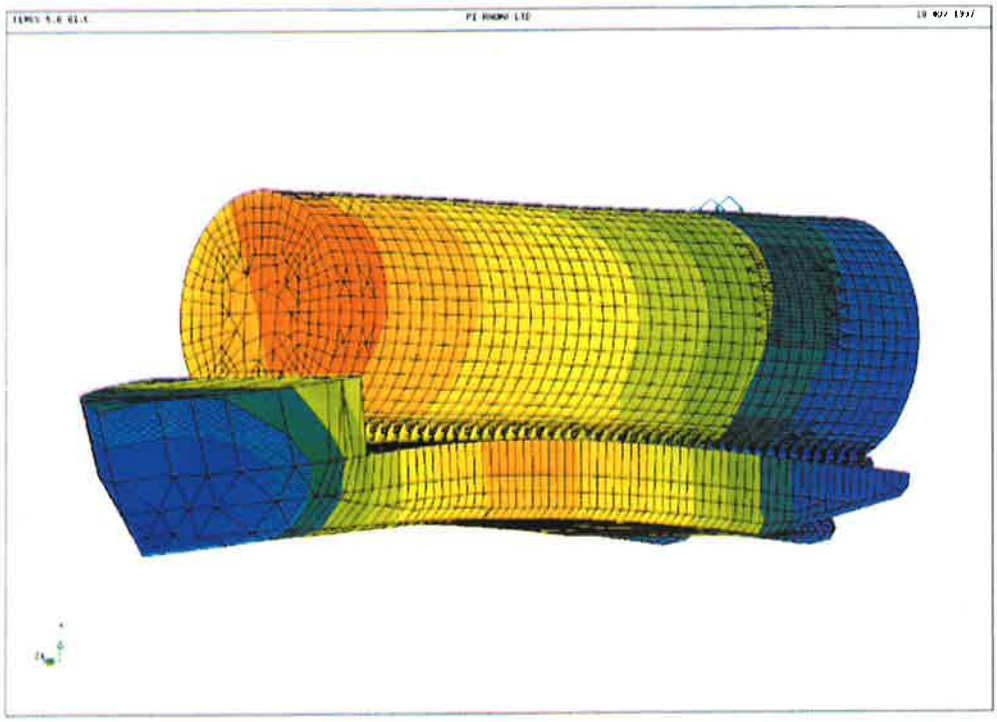
Rakenteesta muodostetaan aluksi nk. globaali FE malli, joka käsittää SPAR rakenteen, tukirakenteet ja aluksen, Kuva 4. Mallitus tehdään pääosin kuorielementeillä käyttäen elementtijakona SPAR rakenteen ja aluksen kaarijakoa. Mallitustarkkuutta kasvatetaan jonkin verran tukirakenteiden lähialueella SPAR rakenteen osalta. Käytännössä globaalimallin mallitustarkkuus on riittävä rakenteen äärilujuus- ja lommahdustarkasteluun. Väsymistarkastelun osalta se antaa eri rakennealueiden jännitystason mutta ei eri hitsiliitosten paikallisia jännityksiä.

Tukirakenteiden FE mallin liittäminen aluksen FE malliin muodostaa käytännössä hankalan mallitusyksityiskohdan, kun yhdistettävien FE mallien topologia ei ole yhteensopiva. Käytännössä tukirakenteiden ja aluksen kannen välillä on puupetaus ja tukirakenteiden liike on sidottu vaakatasossa pisteittäin. Rakenneanalyysissa ei ole lupa ottaa kitkavoimia huomioon. Tosin tältä osin määräykset vaihtelevat jonkin verran. Kuljetusanalyyseissa eri rakennemallit on kytketty toisiinsa siirtymien välisillä sidosyhtälöillä, jotka on muodostettu elementtimenetelmän muotofunktioilla yhdistettävien solmujen asemien perusteella. Kitkaa ei ole huomioitu, joten poikittaiset ja pitkittäisliikkeet on sidottu kuljetuksenakaista tuentaa vastaavasti.

Rakenteiden käyttäytyminen ääritilanteessa on epälineaarista, kun poikittaistuenta ottaa todellisuudessa vastaan vain puristuskuormia ja kuljetettava rakenne saattaa osittain irrota aluksen kannelta. Todellisen ääritilanteen tarkastelemiseksi on joissain tapauksissa

rakenteen reunaehtoja ja aluksen ja tukirakenteen välisiä sidosyhtälöitä muokattu todellista tilannetta paremmin vastaavaksi.

Väsymisanalyysia varten rakenteen mallituksen tulee olla huomattavasti tarkempi mutta koko rakenteen mallittaminen väsymisanalyysin edellyttämällä tarkkuudella on mahdotonta. Käytännössä on rakenteesta valituille detaljeille tehty paikallinen kuorielementtimalli, jonka reunoille on siirtymäreunaehdot otettu globaalimallista. Mallin tarkkuus vastaa laivaalan käytäntöä, jossa elementtikoko on levynpaksuuden suuruusluokkaa väsymisen kannalta kriittisissä kohdissa.



Kuva 4 Rakennemalli ääritilanteen kuormituksessa.

6 LUJUUS JA VÄSYMINEN

6.1 Lujuustarkastelu

Rakenteen lujuustarkastelu käsittää sekä lommhdustarkastelun että äärilujuustarkastelun. Käytännössä kuljetuksen aikaisissa kuormituksissa tulee vastaan lähinnä levykenttien lommahdus. Teoriassa tarkastelussa tulisi laskea lommahdusriskiä kuvaavalle käyttökertoimelle todennäköisin maksimiarvo kolmen tunnin mitoitusmerenkäynnin aikana, joka on

määritelty ainoastaan merkitsevän aallonkorkeuden avulla. Merenkäynnin periodi ja aallokon suunta valitaan siten, että ne maksimoivat tarkasteltavan rakenteen rasituksen.

Viimeisimmissä kuljetusanalyyseissa aallokon periodin ja suunnan valinta tehtiin jokaiselle tukialueelle tämän tuen tukireaktion perusteella. Valituissa aaltotilanteissa laskettiin rakenteiden jännityksille komponenteittain 3 tunnin ääriarvo ja tehtiin näiden perusteella lommahdustarkastelu. Lopputuloksena oli, että kunkin alueen tarkastelu voitiin tehdä yhden jälkikäsittelyllä muodostetun jännitystilanteen perusteella. Eri rakennealueiden osalta tarkastelutilanne vaihteli hieman. Etuna aiempiin menetelmiin on varmuus siitä, että yhden tilanteen tarkastelu riittää. Jännitysten laskenta elementtimenetelmällä joudutaan tekemään tässä noin 400 yksikkökuormitustapaukselle. Tulosten jälkiprosessointi voidaan kuitenkin hoitaa kohtuullisessa ajassa eikä se vaadi suurta tietokonekapasiteettia.

Jännitysten käsittely komponenteittain lisää tuloksen konservatiivisuutta, kun eri jännityskomponenttien vaihetieto menetetään. Tarkastelua voidaan kuitenkin tarvittaessa tarkentaa muodostamalla eri jännityskomponenttien ääriarvoja referenssisuureina käyttäen mitoitusaaltoja. Nämä kuormitustilanteet saadaan jo aiemmin laskettujen yksikköaalto-kuormitustapauseten perusteella, joten uutta FE analyysia ei tarvita. Muodostettujen mitoitusaaltojen perusteella voidaan tarkastella realistisemmin rakenneosan käyttäytymistä ääritilanteessa. Yleensä rajatun rakennealueen eri jännityskomponenttien ääriarvot saadaan tuotettua yhdellä mitoitusaallolla.

6.2 Väsymistarkastelu

Offshore rakenteille vaadittu väsymistarkastelu on konservatiivinen ja rakenteilta vaaditaan melko suuria varmuuksia väsymisen suhteen. Tämä on perusteltua rakenteen hankalan ja suuria kustannuksia aiheuttavan korjaamisen vuoksi. Rakenteet ovat myös monimutkaisia käsittäen valtavan detaljimäärän. Käytännössä kaikkien rakennedetaljien väsymislujuuden tarkka analysointi on mahdotonta ja mitoituksessa on valittava kompromissi menetelmän konservatiivisuuden ja tarkastelun mutkikkuuden välillä. Valittua taajuustason nk. spektraalimenetelmää sovelletaan tällä hetkellä enenevässä määrin merirakenteiden väsymistarkasteluun ja tehdyt analyysit ovat osoittaneet sen käyttökelpoisuuden.

Rakenteen väsymistarkastelu edellyttää aaltokuormien aiheuttamien jännitysvaihteluiden tuottaman väsymisvaurion laskentaa. Kyseessä on hitsattu teräsrakenne, joten tarkastelu on riittävää tehdä rakenteen eri hitsiliitoksille. Tarkastelu pohjautuu liitosten S-N käyrään ja Palmgren-Minerin lineaariseen väsymisvaurion kumuloitumismalliin. Kuormitusspektri muodostetaan kuljetusmatkan yhdistetyn aaltotilaston perusteella. Aaltokuormien tarkastelu tehdään taajuustasossa Rayleigh'n jakauman perusteella, jolloin väsymisvaurio D voidaan laskea kaavalla, ks. (Kukkanen 1996)

$$D = \sum_i \frac{n_i}{C} E\left[S_i^m\right] p_i$$

missä C ja m ovat S-N käyrän vakiot ja n_i ja p_i merenkäyntitilanteen i aaltojen lukumäärä ja todennäköisyys tarkastelujaksossa. Merenkäyntitilanteen i jännitysvaihteluista S_i tarvitaan odotusarvo suureelle S_i^m, joka Rayleig'n jakaumaa käyttäen on

$$E[S_i^m] = (8m_0)^{m/2}\Gamma(m/2+1)$$

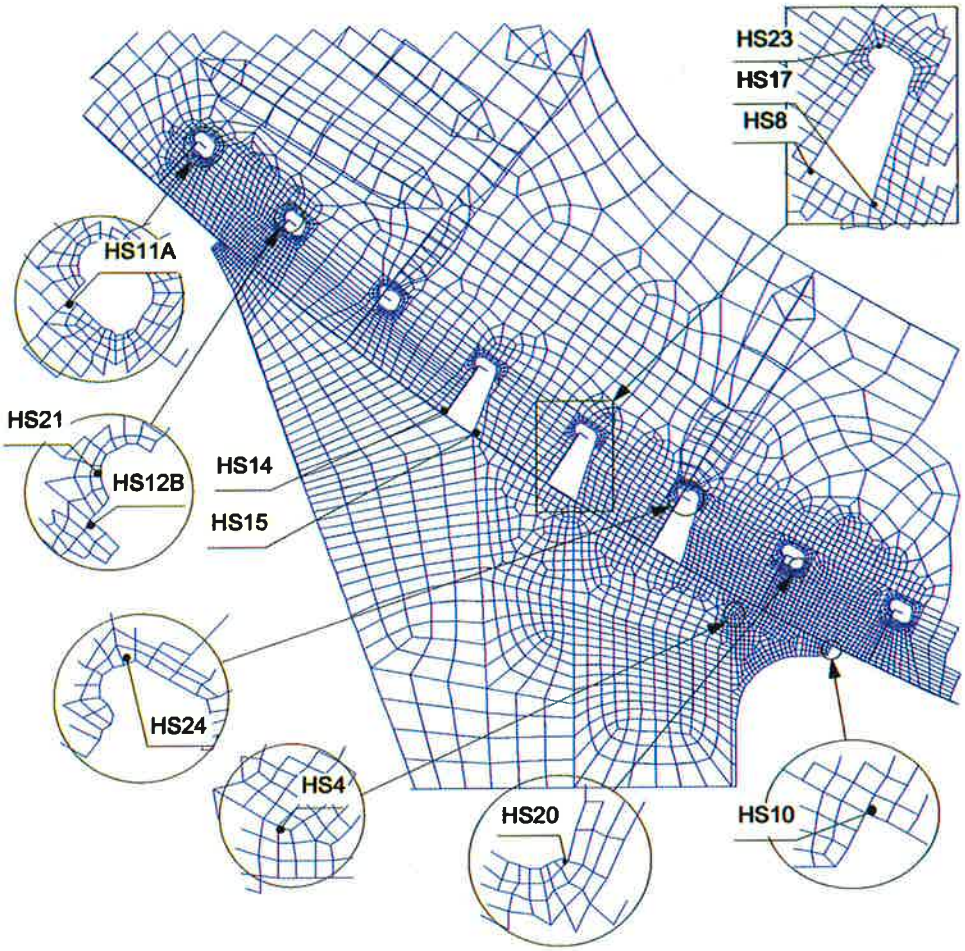
missä m_0 on merenkäyntitilan jännitysspektrin varianssi tai 0:s momentti.

Käytännössä rakenteelle lasketaan jännitysvasteet yksikköaalloille, joiden ja aaltospektrin avulla voidaan muodosta kullekkin tarkastelupisteelle jännitysspektrit ja laskea varianssi. Ongelmaksi muodostuu rakennemallin antamien jännitysten ja hitsiliitosten jännitysten vastaavuus.

Kuljetusanalyyseissa väsymisanalyysi on tehty tarkastelun tarkkuutta vaiheittaen kasvattaen mutta samalla tarkastelukohtien lukumäärää rajaten. Karkeimmassa nk. screening-tason tarkastelussa käytetään globaalista rakennemallista saatavia elementtitason jännitysten avulla laskettuja maksimipääjännityksiä. Nämä lasketaan kullekin kohdalle erikseen pääjännityksen suunnasta välittämättä. Tämän jännityksen ajatellaan edustavan lähialueen hitsiliitosten nimellistä jännitystä. Screening-tason väsymisvaurio lasketaan käyttäen jännityskonsentraatiokertoimena esimerkiksi arvoa 3 ja F2 luokan S-N käyrää. Tulokset saadaan kohtuullisella laskenta-ajalla koko rakenteelle. Tulosten käsittelyssä rajataan mallin singulaarisuudesta johtuvat näennäiset kriittiset alueet ja selkeästi mitättömän väsymisvaurion alueet jatkotarkastelun ulkopuolelle.

Väsymisanalyysin tarkentaminen voidaan tehdä laskemalla valituille rakennealueille muutamille kuormitustilanteille jännityskonsentraatiokertoimet, Kuva 5. Eri liitoksille saadaan laskettua jännityskonsentraatiokertoimien, niitä vastaavien S-N käyrien sekä screening-tason tarkastelun antaman väsymisvaurion avulla tarkennettu arvio väsymisvauriosta. Tarkastelu yksinkertaistuu huomattavasti, jos kunkin detaljin jännityskonsentraatiokerroin on aaltokuorman taajuudesta ja aallon suunnasta riippumaton vakio. Käytännössä useimmat detaljit voidaan tarkastella tällä tavoin karkeasti.

Väsymistarkastelun tarkentamista voidaan jatkaa tuottamalla detaljien jännityskonsentraatiokertoimet aaltokuorman taajuuden ja suunnan funktiona. Periaatteessa tämä tarkastelu on yksinkertainen, mutta se edellyttää huomattavan suuren lähtötietomäärän syöttämistä väsymisvaurion laskentaan. Tämä koostuu eri aaltokuormien jännityskonsentraatiokertoimista, jotka lasketaan kukin erillisellä FE mallilla. Laskentaa voidaan viedä tasolle, jolla kaikille yksikköaalloille lasketaan oma jännityskonsentraatiokerroin. Tällöin on helpointa prosessoida paikallisen tarkennetun FE-mallin jännitykset kuten globaalimallin jännitykset screening tarkastelussa.



Kuva 5 Jännityskonsentraatiokertoimien analysoinnissa käytetty rakenteen detaljimalli.

7 YHTEENVETO

Rakenteiden mitoittamiseksi kuljetusten aikaisia kuormituksia vastaan on kehitetty taajuustason nk. spektraalimenetelmään perustuva laskentaproseduuri. Proseduurissa aaltokuormien laskenta tehdään paneelimenetelmällä, josta saadaan rakennemallin kuormitukset kiihtyvyyksinä ja aluksen kylkipaineina. Rakenteen jännitystasot lasketaan FE menetelmällä rakenteesta muodostetulla 3D mallilla. Laskentatulokset prosessoidaan tilastollisen tarkastelun avulla äärilujuus, lommahdus ja väsymistarkastelua varten. Laskentamallien tuottama valtava tulosmäärä saadaan tehokaasti tiivistettyä eri tarkasteluihin soveltuvaksi. Analyysin useita työläitä vaiheita on voitu automatisoida.

Käytännössä rakennesysteemi ei toimi lineaarisesti ja lineaarisuusoletus johtaa helposti ylikonservatiiviseen mitoitukseen erityisesti ääritilanteen osalta. Usein systeemin epälineaarinen käyttäytyminen nimenomaisesti estää ylikuormittumisen ja tekee koko kuljetuksen mahdolliseksi. Kuljetusanalyysiryhmä on onnistunut kehittämään toimivan analyysiproseduurin, jossa käytettävillä linearisioinneilla ja mallitustavoilla päädytään järkevään lopputulokseen. Työn aikana on useita laskentaproseduurissa valittuja ratkaisuja jouduttu perustelemaan ja osoittamaan niiden luotettavuus loppuasiakkaalle. Käytännössä toteutetut kuljetukset ovat osoitus myös onnistuneista kuljetusanalyyseista.

8 LÄHDEVIITTEET

Kukkanen, T. (1996). Spectral fatigue analysis for ship structures. Uncertainties in fatigue actions. Licentiate's Thesis, Helsinki University of Technology, Ship Laboratory, 94 p. + app. 7 p.

Mikkola, T.P.J., Karppinen, T., Kukkanen, T., Silvola, I., Arjava, J.-P., Kahala, T. & Peltomaa, J. (1998). Transport analysis of spar hull structures. Structural Design '98 Espoo 1998, Teknillinen korkeakoulu, Laivalaboratorio, pp. 110 - 128.

MICROMECHANICAL MODELS IN DUCTILE FRACTURE ASSESSMENT

H. TALJA
VTT Manufacturing Technology
P.O.Box 1704
FIN - 02044 VTT

ABSTRACT

Fracture mechanics is "classically" used to describe material failure. To overcome the theoretical limitations of the one-parameter fracture mechanics, several triaxility or constraint parameters have been introduced. More elegant methodologies are based on the local approach. Ductile fracture is commonly described using the modified Gurson model, based on micromechanical studies of void initiation, growth and coalescence. The parameters of the material model can be determined by matching tensile test results by finite element analysis, and then be applied e.g. to J-R curve prediction. In the paper, this approach is applied to the German nuclear reactor pressure vessel steel 20 MnMoNi 5 5 with slightly anisotropic material property values.

1. INTRODUCTION

Classical fracture mechanics describes the effect of material defects or cracks on structural integrity using a single global parameter, usually either the stress intensity factor K_I introduced by Irwin in 1957 or the J-integral introduced by Rice (1968). The basic idea is that the fracture mechanics parameter controls the (singular) stress and strain fields at the vicinity of the crack tip.

Unfortunately, this is true only under very severe restrictions like for stationary (non-growing) cracks. If certain size requirements assuring plane strain conditions are not met, a so-called loss of constraint situation occurs and the measured fracture toughness values K_{IC} and J_{IC} become dependent on specimen geometry and size.

The required specimen sizes and crack lengths for determining valid K_{IC} and J_{IC} values are often unpractically and uneconomically large. The valid toughness values are lower bound values, and as such they may be overly conservative when applied in practical cases with smaller structural dimensions and realistic crack sizes. The transferability of the parameters from specimens to real structures needs close attention. Especially in the case of irradiated nuclear reactor pressure vessel (RPV) materials, only a very limited amount of material is available for testing. In such cases the required size criteria may even be impossible to fulfil.

In order to expand the limits of the classical approach, various methodologies have been proposed. So-called two-parameter approaches typically introduce an additional parameter, which characterises the constraint or size effect, e.g. by considering the second term of the series expansion of the stress field. Besides, empirical correlations have been found between the stress state triaxiality and the slope of the ductile fracture resistance curves. A review of such methods can e.g. be found in (Talja 1998).

A more general way to treat ductile failure is to use the local approach, which can be divided into continuum damage mechanics (CDM) and micromechanical models. The CDM models take the material deterioration into account in the constitutive equations, whereas the micromechanical models describe the actual void growth and coalescence phenomenoma in the material. One of the main advantages of local approach methods is that they can be applied just as well to initially intact structures as to cases where there is an initial crack.

2. MICROMECHANICAL MODELLING OF DUCTILE FAILURE

2.1. Micromechanics of ductile failure

Ductile failure mechanism in metals occurs by one of two competing mechanisms: either shear fracture with a purely tangential relative displacement of fracture surfaces, or a mechanism caused by microvoids, yielding a "dimpled rupture" appearance of the fracture surface. Sometimes extensive fractographic studies are needed to distinguish between the two failure mechanisms.

As shown schematically in Figure 1, the ductile material deterioration due to microvoids develops in three phases: nucleation of cavities, followed by void growth and coalescence. The cavities may nucleate either as brittle fracture or decohesion of inclusions of second-phase particles. The diameter of large particles, like MnS and Al_2O_3, is typically a few microns and their spacing is of the order of 100 microns. Plastic growth of microvoids continues with increasing plastic strain until a critical condition is reached due to localised plastic instability or internal necking of the intervoid matrix.

2.2. The modified Gurson model

Material models like the von Mises plasticity model assume that the plastic deformation does not depend on the hydrostatic stress. Such models cannot describe the dilatational behaviour of porous materials. Currently the most popular material model for porous materials has become the constitutive model first presented by Gurson (1977). Gurson derived his constitutive model for an aggregate of ductile matrix material and voids by an

upper bound analysis for a spherical void in a spherical body ("unit cube") of a rigid-plastic material. Tvergaard (1981 and 1982) presented modifications to the original model to incorporate the interaction of neighbouring voids and to consider hardening materials, resulting in the currently used form of the model:

$$\Phi = \frac{\sigma_e^2}{\sigma_f^2} + 2\, f\, q_1 \cosh\left(\frac{3 q_2 \sigma_m}{2\sigma_f}\right) - (1 + q_3 f^2) = 0 \qquad (1)$$

where σ_e is von Mises equivalent stress, σ_f current yield stress of the hardening matrix material, σ_m the mean stress (hydrostatic stress), f the current void volume fraction and q_1, q_2, and q_3 parameters introduced by Tvergaard. At a porosity value of 0, Eq. (1) reduces to the von Mises yield condition. When the porosity is equal to 1, all stress components have to vanish to fulfil the equation and the material loses its load carrying capacity. In reality this occurs already at a much lower value of porosity. Figure 2 shows how the porosity f and the parameter q_1 affect the yield surface.

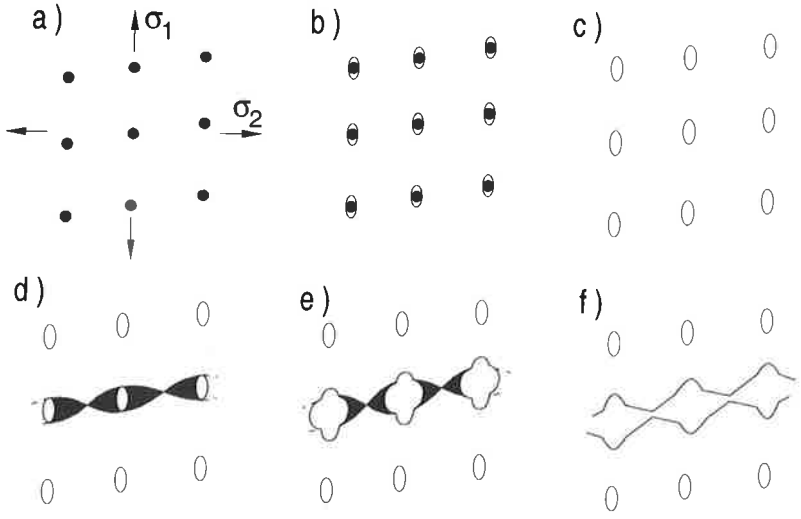


Figure 1 Development of ductile failure by void initiation, growth and coalescence in metals (Thomason 1990).

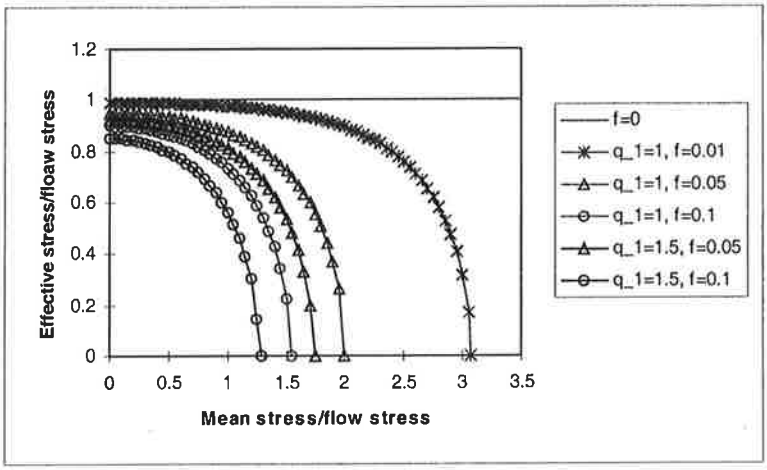


Figure 2 The yield criterion of Eq. (1) at different values of f and q_1 when $q_2 = 1$ and $q_3 = q_1^2$. The case f = 0 corresponds to the von Mises yield criterion.

To take the effect of void coalescence into account, Tvergaard and Needleman (1984) introduced a modified void volume fraction f* substituting f in Eq. (1):

$$f^* = f; f \le f_c$$
$$f^* = f_c + K_F(f - f_c); f > f_c \quad (2)$$

where f_c is the critical void volume fraction corresponding to the beginning of void coalescence. The void volume fraction f_f at final material failure determines parameter K_F

$$K_F = \frac{f_u^* - f_c}{f_f - f_c} \quad (3)$$

where the "ultimate void volume fraction" f_u^* at zero stress is $f_u^* = 1/q_1$. As $f \rightarrow f_f$, $f^* \rightarrow f_u^*$ and the material loses its stress carrying capacity.

The evolution of the void volume fraction f consists of nucleation and growth (Chu & Needleman 1980):

$$\dot{f} = \dot{f}_{growth} + \dot{f}_{nucl} \quad (4)$$

where, due to the incompressibility of matrix material

$$\dot{f}_{growth} = (1 - f)\dot{\varepsilon}_{kk}^p. \quad (5)$$

The model has been presented in more detail e.g. by Talja (1998), with a summary of some most common void nucleation models.

2.3. Determination of the modified Gurson model parameters

Perhaps the most important question in applying the modified Gurson model in practical analyses is how to determine the parameter values. The number of parameters is large: in addition to the parameters q_1, q_2 and q_3 they include also the *initial parameters* f_o and f_n as well as the *critical parameters* f_c and f_f (or K_f). f_o is the initial value of porosity and f_n the volume fraction of void nucleating particles. If the void nucleation is modelled, additional parameters for the nucleation model have to be determined.

The q_i parameters may be determined by cell model calculations (e.g. Tvergaard 1981, Koplik & Needleman 1988). Often the values $q_1 = 1.5$, $q_2 = 1$ and $q_3 = q_1^2$, suggested by Tvergaard (1982) for a matrix material with a strain hardening exponent of 10, are chosen.

To determine the initial and critical parameters a method consisting of both experimental and numerical work is necessary, as a direct measurement of the parameters would be extremely tedious if not altogether impossible. In practice the parameter values are obtained by matching the experimental results through numerical simulation. Very simple specimens like smooth tensile specimens can be used and the specimen size can be very small. Before direct application to a real structure, the transferability of the parameters has to be confirmed, e.g. by simulation of a fracture mechanical test. This approach has been successfully applied even to a multimaterial sub-cladding crack case (Schmitt et al. 1995).

If there is a strong gradient in the stress distribution, the finite element analysis results become dependent on mesh density and the element size has to be considered as an additional parameter. Often the spacing of adjacent inclusions is understood as a critical length. So, typically element sizes of 50 ... 200 μm are used to model the area where the crack growth is expected to occur.

3. APPLICATION OF THE MODEL TO REACTOR PRESSURE VESSEL STEEL

3.1. Material properties

Tensile and fracture mechanical tests were performed at Fraunhofer-Institut für Werkstoffmechanik (IWM) in the "Size effect" project (Größeneinfluß) using specimens made of several RPV steels, including the German steel 20 MnMoNi 5 5. The microstructure of the material is bainitic. At a small magnification (40x) the micrographs from the etched surface show darker and lighter bands resulting from macrosegregation in the material (Figure 3). However, microstructural features like the grain size showed no remarkable differences between the different orientations. The amount of inclusions was found to be small. Some of the MnS inclusions were strongly elongated in cross-sections T and S, with length

length values up to several hundreds of microns. In cross-section L they had an oval shape with a length of a few microns.

The effect of anisotropy can also be seen in the mechanical properties. Their values, measured from smooth tensile specimens with a diameter of 10 mm, are shown in Table 1. Orientation L is the axis of forging, T the long transverse direction and S the short transverse direction. R_{eL} is the lower yield strength, R_m ultimate strength, A_5 fracture elongation and Z reduction of area.

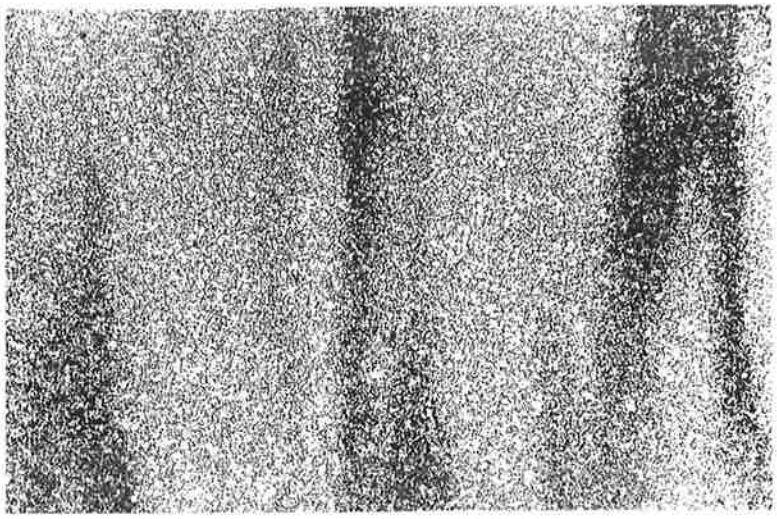


Figure 3 Steel 20 MnMoNi 5 5: Microstructure of the material in the cross-section perpendicular to direction T (40x).

Table 1 Steel 20 MnMoNi 5 5: tensile test data in different orientations (average values from three experiments).

Orientation	R_{eL} (MPa)	R_m (MPa)	A_5 (%)	Z (%)
L	552	683	24.4	74.5
S	523	666	20.3	60.0
T	520	665	21.4	67.7

3.2. Simulation of tensile tests

The tensile tests were simulated by finite element analyses to fit the values of Gurson model parameters. The specimens were modelled using 8-noded axisymmetric iso-parametric elements with reduced (2^{nd} order) Gaussian integration. An initial imperfection of 0.5 % was modelled and there were 12 equally spaced elements over the radius. The idea was to match as well as possible the point where the rapid load drop due to necking starts. Diameter reduction at the necking location, which is quite sensitive to the local instability, is a more suitable parameter for this purpose than the specimen elongation. The scatter of the material parameter values was also studied by fitting the scatter band of tensile test results.

The fitting was done twice, in two different "directions", varying either initial or critical f values, and the correspondence between the scatter bands in tensile and fracture mechanical test results was studied. In all cases the values of f_0 and K_F were taken as 0 and 4, respectively, and the values of the initial and critical parameters f_n and f_c were determined. In orientation T, the obtained value of f_c varied from 0.018 to 0.10 when the f_n the parameter was fixed as f_n = 0.004. Figure 4 shows a comparison of measured and computed results corresponding to these values. Corresponding to a fixed f_c = 0.03, f_n varied in the range of 0.002 to 0.006.

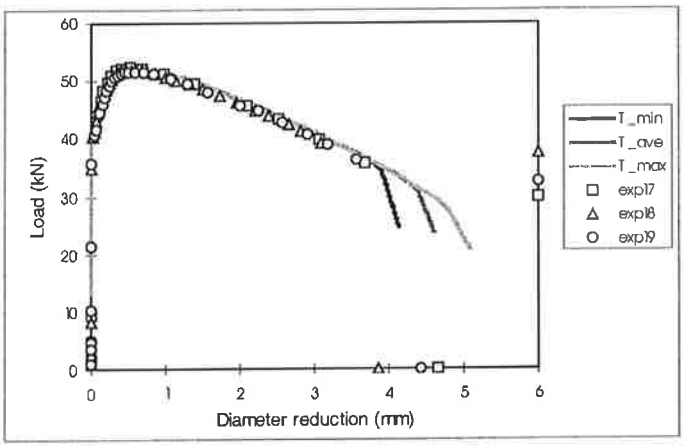


Figure 4 Steel 20 MnMoNi 5 5: Comparison of calculated and measured load values in tensile tests as a function of diameter reduction in orientation T. f_n = 0.004 and f_c is varied in the computations.

3.3. Simulation of fracture mechanical tests

Although the J-integral should strictly speaking not be applied to growing crack cases, the J-R curve approach has proved in practice very successful for the estimation of a limited amount of crack growth. Essential parameters of the J-R curve are the J-integral value corresponding to crack growth initiation J_i, which often is replaced by $J_{0.2}$ corresponding to a crack growth of 0.2 mm, and the tearing modulus T_J, which is the normalised slope of the J-R curve.

The parameter values obtained by matching the tensile tests were applied to simulate the behaviour of 50 mm thick side-grooved CT specimens, which also had been tested in three different orientations. Each analysis was performed using the stress-strain curves and damage parameters corresponding to the direction of tensile loading. As usual for side-grooved specimens, these computations were made using a plane strain finite element model. At the crack tip area 8-noded isoparametric elements with a length of 0.2 mm and a height of 0.12 mm were used. J-R curves representing the correspondence between J-integral values and crack growth were estimated. The crack growth was defined in the computation by porosity values exceeding f_f.

Table 2 shows measured and estimated results for orientation T-L. The computation with the "average" parameter values agrees quite well with the experiment. The scatter of the fracture mechanical test results is strongly overestimated, when the value of the critical parameter f_c is varied. The estimated fracture toughness seems to be much less sensitive to the initial parameter f_n than to the critical parameter f_c.

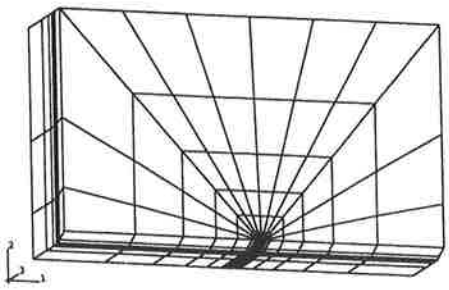
Table 2 Material 20 MnMoNi 5 5: Measured and estimated values of $J_{0.2}$ and tearing modulus (T_J) in orientation T-L.

	Measured	**f_c varied**			**f_n varied**	
		min	**ave**	**max**	**min**	**max**
$J_{0.2}$ (N/mm)	330	220	316	466	228	248
T (-)	210	210	232	457	320	325

An additional computation was performed using the 3D model shown in Figure 5a. Because elements with very small in-plane dimensions have to be used at the crack tip area, heavily non-linear 3D-analyses with micromechanical models are still unpractically time consuming. Though the modelling here was rather coarse, especially in the through-the-thickness direction, useful results were obtained. The analysis confirmed the applicability of a 2D model when global values like J-R curve results are considered (Figure 5b).

Besides, the 3D analysis was able to reproduce quite well also the local variation in crack growth along the crack front, which was clearly seen as enhanced growth near the side-groove (Talja 1998).

a)



b)

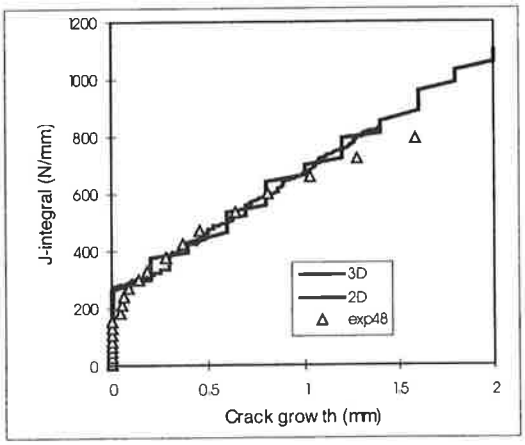


Figure 5 a) The 3D finite element model for a quarter of a side-grooved CT-specimen and b) a comparison between estimated and measured J-R curves.

4. CONCLUSIONS

In the paper an application of the modified Gurson model to tensile and fracture mechanical test specimens was presented. The importance of model parameter definition was shown: the amount of model parameters is large and their values cannot be measured directly. So the common procedure is to apply numerical fitting of experimental results. Very careful fitting is needed in order to achieve results which are transferable from one specimen type to another.

Because very detailed modelling is necessary and the simulation of the non-linear phenomenoma during crack growth is rather time consuming, very few applications for real component assessment with the Gurson model have been presented to date. It would be especially interesting to investigate the transferability of the parameters to a case of pressurised thermal shock where biaxial loading is of significant importance.

In future, one very interesting application area would also be to study the constraint effects in fracture mechanics specimens, i.e. to predict the dependence of the J-R curves on

specimen geometry and size. Some examples of use of the model in mixed mode fracture cases have been presented in the literature (e.g.. Laukkanen 1999); however, the feasibility of this approach should be evaluated.

ACKNOWLEDGEMENTS

The main part of this research was performed within the Finnish Research Programme on the Structural Integrity of Nuclear Power Plants (RATU2) and funded by the Ministry of Trade and Industry and VTT Manufacturing Technology. The finite element analyses were partly funded by Fraunhofer-Institut für Werkstoffmechanik (IWM) in Freiburg, Germany. Their funding and the additional financial support provided by VTT Manufacturing Technology and the foundations of IVO, Jenny and Antti Wihuri and Guerillot is gratefully acknowledged.

REFERENCES

Chu, C. C. & Needleman, A. 1980. Void nucleation effects in biaxially stretched sheets. Journal of Engineering Materials and Technology, vol. 102, pp. 249 - 256.

Gurson, A. L. 1977. Continuum Theory of Ductile Rupture by Void Nucleation and Growth: Part 1 - Yield Criteria and Flow Rules for Porous Ductile Media, Transactions of the ASME, Jan 1977, pp. 2 - 15.

Koplik, J. & Needleman, A. 1988. Void Growth and Coalescence in Porous Plastic Solids. Int. J. Solids & Structures, vol. 24, pp. 835 - 853.

Laukkanen, A. 1999. Applicability of Gurson-Tveergard constitutive model to characterise mixed-mode and mode II crack propagation. VTT Manufacturing Technology, VALB350, 71 p.

Rice, J. R. 1968. A path independent integral and the approximate analysis of a strain concentration by notches and cracks. Journal of Applied Mechanics, Transactions of the ASME, June, pp. 379 - 386.

Schmitt, W., Sun, D. - Z. & Blauel, J. G. 1995. Schädigungsmechanische Analysen (GURSON Modell) und experimentelle Verifizierung des Fehlerverhaltens in einem schweissplattierten Bauteil, 21. MPA-Seminar, 5.-6.10. 1995.

Talja, H. 1998. Ductile fracture assessment using parameters from small specimens. Espoo: Technical Research Centre of Finland, VTT Publications 353. 140 p.

Thomason, P.F. 1990. Ductile Fracture of Metals. Oxford: Pergamon Press. 219 p.

Tvergaard, V. 1981. Influence of Voids on Shear Bend Instabilities under Plane Strain Conditions. Int. J. of Fracture, vol. 17, pp. 389 - 407.

Tvergaard, V. 1982. On localization in ductile materials containing spherical voids. Int. J. of Fracture, vol. 18, pp. 237 - 252.

Tvergaard, V. & Needleman, A. 1984. Analysis of the cup-cone fracture in a round tensile bar. Acta Metallurgica, vol. 32, pp. 157 -169.

NUMERICAL SIMULATION OF THERMAL RESIDUAL STRESSES IN A FUNCTIONALLY GRADED MATERIAL COMPONENT

K. CALONIUS
VTT Manufacturing Technology
Kemistintie 3
P.O.Box 1704
FIN-02044 VTT, Finland

ABSTRACT

The Hot Isostatic Pressing manufacturing process of a functionally graded material component has been numerically simulated in this work. The main concern is the residual stress state in the component. During the manufacturing process the temperature range varies widely from room temperature to 1180°C. Non-linear thermo-visco-plastic analyses have been carried out using ABAQUS/Standard [1] finite element method computer code. The material parameters needed for these analyses are partly compiled from the literature and partly acquired through experimental work. The values of the material parameters needed to model the creep are obtained by curve fitting from the experimental data generated in this project.

1. INTRODUCTION

The joining of two steel materials is usually done with additional threaded fittings or by welding. Sometimes those methods cannot be used and an alternative way is to use a transition part where two powder metallurgically processed materials are joined together using a high pressure and an elevated temperature enabling diffusion. The above-mentioned process is called hot isostatic pressing (HIP).

Consolidation of two different materials induces thermal residual stresses, which develop during the cooling down stage especially at their interface. First of all, the mechanical and physical properties of the two different materials change in a complex way as a function of temperature according to their own characteristics and secondly, the cooling is never uniform over the whole material.

There are different ways to reduce thermal residual stresses in multimaterial components. One possible way is to use a seamless bonding, in other words, a section of Functionally Graded Material (FGM). The purpose of this study is to numerically simulate thermal residual stresses in a certain FG-material component. This component is a 250 mm long tube consisting of two base metals: a stainless steel and a ferritic steel. The material gradient between these two different metals has been made as smooth as possible using powder metallurgy based on FGM technology. The second goal is to find other ways to minimize the residual stresses.

The component and its manufacturing are explained. The most important material parameters are defined. The data is collected partly from the literature and partly obtained from the experimental work. A suitable finite element mesh is created and FEM calculations are carried out using four different heights of the FG-region. Finally, the results of the FEM-analyses are discussed and conclusions drawn.

2. COMPONENT

An FG-material can be defined as a material whose composition changes in a desired way as a function of position, usually microstructurally from a base material to another one. It is a composite material without any sharp interfaces. The aim is mainly to reach a gradual variation of mechanical properties. It has long been assumed that interface mismatch stresses can be reduced by replacing a sharp interface with an intermediate composite layer and in a similar case to this study the analogy has been demonstrated by FEM-calculations [2]. Nevertheless, understanding the stress reduction mechanisms, especially with the materials studied here, is limited.

Altogether six different materials are used in the numerical simulations. Five of them are powder materials manufactured using a HIP process: a ferritic steel, an austenitic steel and their three alloys. The sixth material, a conventionally manufactured austenitic steel, is used in the process itself for the container and for the central core. It is not a part of the finished component. The three gradient alloys are 75/25 steel, 50/50 steel and 25/75 steel, which are used to model the FG-region. The numbers before and after the slash are volume percentages of ferritic steel and austenitic steel, respectively. The lower part of the HIPped tube is ferritic steel and the upper part is austenitic steel, as can be seen in Figure 1.

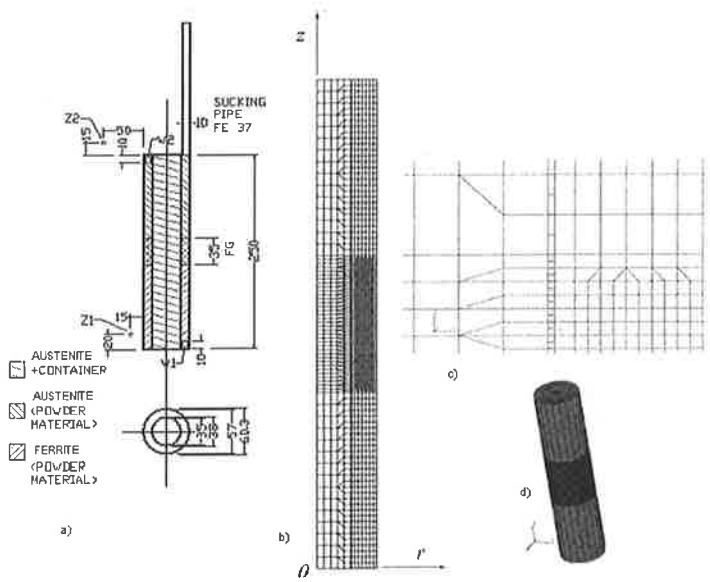


Figure 1. Real component and its dimensions (a). The whole mesh with the coordinate system (b). A mesh detail (c) and a circumferentially swept 3-dimensional model (d).

The initial dimensions of the component and its container are shown on the left in Figure 1. The central core has been machined slightly cone-shaped, so that it would be easier to separate from the inner surface of the tube after the process and there would be room for the lubricant between them. The temperature was measured during the process in the furnace at points referred as Z1 and Z2.

3. MANUFACTURING

In the beginning of the HIP process the material is consolidated under high isostatic pressure and at high temperature using gas as a medium substance for pressing. The initial state of the material can be solid, porous, or as in this case powder. Because the surface of the powder must be dense in HIPping, thus allowing the component to consolidate properly, the powder is tightly packed and sealed in a container, in this case in a metal can. In order to transfer the pressure to the powder, the container must be thin enough. It is then loaded into the HIP chamber, vacuum is created in the container and temperature and pressure are applied typically simultaneously in the chamber according to a program.

First both the temperature and the pressure were elevated nearly linearly to 1180°C and 100 MPa correspondingly in about an hour. Next they were kept at the set temperature for about 3 hours, in which time also the inner parts reached the same temperature and the powder fully consolidated. Finally, the furnace was cooled down and pressure discharged as shown in Figure 2.

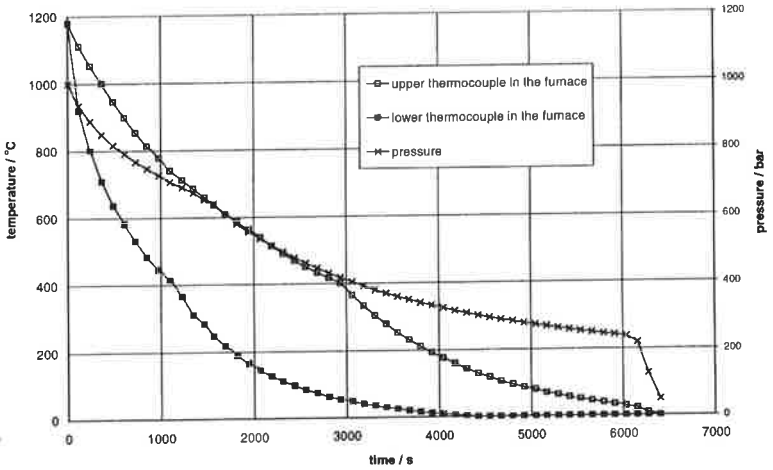


Figure 2. Temperature and pressure in the HIP process on 14th of March 1996. The pressure transient and the temperature transient measured by the upper thermocouple in the furnace were used in the FEM analyses.

4. MATERIAL PROPERTIES

The heat transfer analysis is based on the basic energy balance equation and the material parameters needed are thermal conductivity (λ), specific heat (c) and density (ρ).

All the parameter values for the heat transfer analysis are compiled from the literature. The values for interlayer materials have been estimated using the Law of Mixtures [3]:

$$X = \sum V_i \cdot X_i = V_{AUSTENITE} X_{AUSTENITE} + V_{FERRITE} X_{FERRITE}, \quad (1)$$

where V_i is the volume fraction of the particular base material and X_i its value. This assumption is justifiable since the thermal properties are to a large extent based on material specific factors, i.e. proportion of ingredients.

The basic feature of the whole constitutive system is that the deformation is divided into four different parts: thermal, elastic, plastic and creep deformations. Every material is isotropic.

The thermal expansion coefficient (Figure 3) is the most fundamental property, when studying the development of thermal residual stresses in a multimaterial system. The difference of the values of this parameter between two materials produces substantial stress peaks, whenever the part is exposed to a thermal cycle, such as processing.

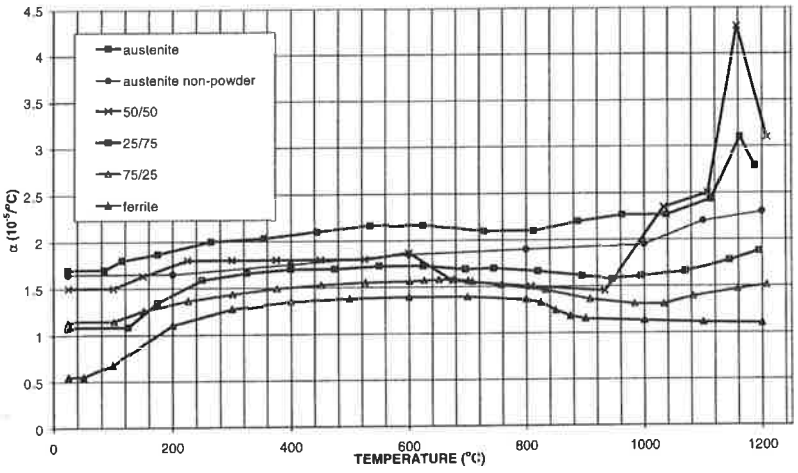


Figure 3. Coefficient of Thermal Expansion of every material as a function of temperature. [4]

The metals studied have approximately linear elastic behavior at low stress levels and the Young's modulus (E), is constant at constant temperature. Figure 4 shows the Young's modulus for every material used in simulations as a function of temperature. Each material

has relatively close values to each other between temperatures RT and 400°C, after which the values diverge.

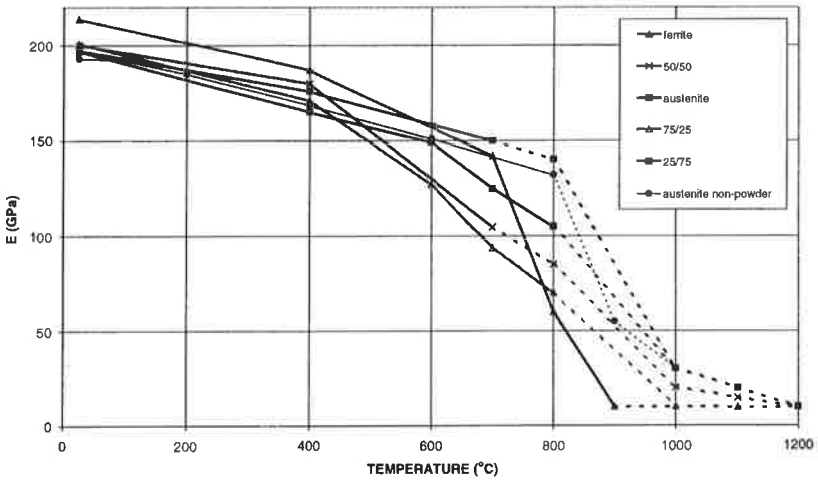


Figure 4. Young's Modulus of every material as a function of temperature. The approximated or extrapolated points are connected by dashed lines. [4]

A classical metal plasticity model based on the standard von Mises yield surface with associated plastic flow is used. The hardening in the model is isotropic, which means that the yield surface changes size uniformly in all directions such that the yield stress increases as plastic straining occurs.

Since the structure is exposed to continuous load at high temperatures, the importance of creep becomes quite obvious especially for rate-sensitive materials such as metals studied in this project. In this model the creep and plasticity occur simultaneously. Especially as both are isotropic, their interaction can properly be handled and a coupled system of constitutive equations integrated. The strain-hardening version of the uniaxial power-law creep model has been chosen. It is introduced in the following form:

$$\dot{\bar{\varepsilon}}^{cr} = \left(A\bar{q}^{n}\left[(m+1)\bar{\varepsilon}^{cr}\right]^{m}\right)^{\frac{1}{m+1}}, \qquad (2)$$

where $\dot{\bar{\varepsilon}}^{cr} = \sqrt{\frac{2}{3}\dot{\boldsymbol{\varepsilon}}^{cr} : \dot{\boldsymbol{\varepsilon}}^{cr}}$ is the uniaxial equivalent creep strain rate, $\bar{\varepsilon}^{cr}$ is the equivalent creep strain and $\bar{q}$ is the uniaxial equivalent deviatoric stress (von Mises). A, n and m are defined by the curve-fitting program and are all dependent on the temperature [1]. The strain curves from the creep tests at 800°C are shown in Figure 5.

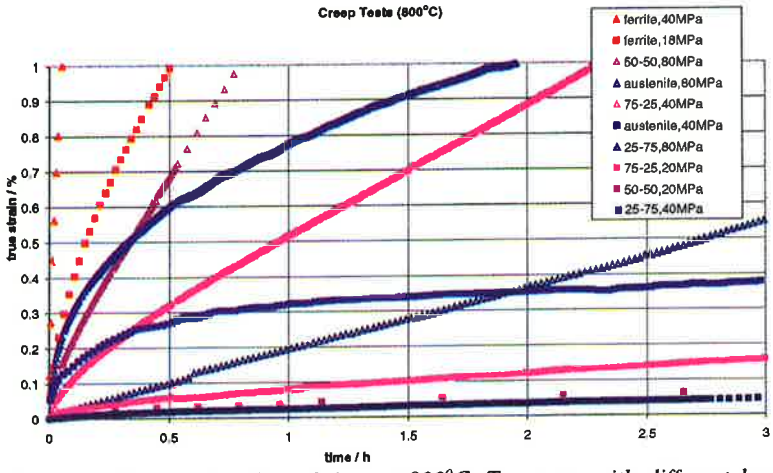


Figure 5. Creep strain as a function of time at 800°C. Two tests with different load levels are carried out for each material tested. [4]

5. NUMERICAL SIMULATION

Due to axially symmetric loading and boundary conditions, the model is simplified to two dimensions using axisymmetric solid elements. The coordinate system and an example of the finite element mesh are shown in Figure 1. The main model dimensions are equal to the dimensions of the component. The left edge of the mesh corresponds to the axis of symmetry and is thus fixed in the radial direction.

Exactly the same temperature and pressure transient is used in every case. The initial and stress free temperature for every node is 1180°C and the temperature transient is set to every node in the surface of the model. The isostatic pressure is modeled with a uniformly distributed pressure load on every element face on the edge: in the top, side and bottom of the model. The load is also set both on the inner and outer surface of the narrow gap between the core and the tube. The pressure transient is also obtained from the same manufacturing process than the temperature transient (Figure 2). Because the analysis was decided to be ended to the room temperature, the total time of the analysis was 6158 seconds, the final surface temperature was 25°C and the final pressure was 22.02 MPa.

The mesh configuration is made up with eight-node biquadratic elements with four integration points and reduced integration. Some triangular elements have to be used in regions where the mesh is refined. They are quadratic with 6 nodes and 3 integration points. The number of degrees of freedom is 11563. The clearance between the central core and the tube is modeled with gap elements provided in ABAQUS program [6]. They are shown in Figure 1. The gap elements allow for contact between the opposite nodes of the interfaces. The initial separation distance between those nodes is 0.5 mm. Quadratic diffusive heat transfer (correspondingly 6- or 8-node) elements are used in the heat transfer analysis and their nodes have only one active degree of freedom: temperature.

The main variable property in the model is the height of the FG-region. There are four different analysis cases, in which the only difference is that in Case 1 the height is 10 mm, in Case 2 it is 20 mm, in Case 3 it is 30 mm and finally in Case 4 it is 40 mm. The aim is to find out the lowest possible height with which the overlapping interfaces do not have an effect on each other's stress distribution, at least a distinct one that elevates the stresses.

6. RESULTS

The most important stress components are the axial tensile (σ_{22}) and shear stresses (σ_{12}) near the material interfaces on the radial free edge. In the interior region the main stress components of interest are the in-plane stresses, radial (σ_{11}) or circumferential (σ_{33}). The interior region is believed not to be important in this case. The contour plots of von Mises stress distribution in Figure 7, for example, show that the stress peaks are at the interfaces, near both the inner and outer surface.

The residual stresses and strains in integration points of certain corresponding elements in every case are calculated in Table 1. Only the highest value of the values of four integration points of those elements is included in the tables. The concerned elements are in both sides of every four material interface in the second vertical element layer from the outer container wall. They are from the final state of the analysis. The locations of these elements for Case 3 are painted red in Figure 6. Also yield stress ($R_{0.2}$) and shear strength (τ_m) are shown in Table 1.

Table 1. von Mises stress, shear stress (σ_{12}) and both the equivalent creep (CEEQ) and plastic strain (PEEQ) values of Cases 1-4 in the both sides of material interfaces, near the free edge.

	von Mises [MPa]					σ_{12} [MPa]					CEEQ [%]				PEEQ [%]			
Cases	1	2	3	4	$R_{0.2}$	1	2	3	4	τ_m	1	2	3	4	1	2	3	4
AUSTENITE					341					197								
	360	365	360	360		-145	-155	-160	-160		1.3	1.4	1.4	1.5	0.5	0.6	0.5	0.5
	300	335	350	360		-145	-155	-160	-160		0.6	0.8	0.9	1	0.6	0.8	0.8	0.9
25/75					357					206								
	110	30	20	30		5	10	-5	-10		0.5	0.5	0.5	0.6	0.8	1	1.5	2.1
	55	60	65	65		10	10	-5	-15		3.3	3.5	3.5	3.5	5.2	6.3	6.8	6.9
50/50					478					276								
	95	120	140	160		-45	-55	-60	-60		3.4	3.5	3.4	3.3	1.8	1.9	1.7	1.5
	220	185	150	130		-45	-55	-60	-60		1.1	1.1	1.2	1.3	6	6.3	6.8	7.3
75/25					577					333								
	135	125	150	160		-60	-65	-65	-65		0.4	0.2	0.2	0.2	0.4	0	0	0
	280	210	175	160		-55	-60	-65	-65		1	0.8	0.8	0.8	0.1	0.1	0.1	0.1
FERRITE					359					207								

The final displaced shape of the FG-region of the tube is shown in Figure 6. That kind of shape distortion is mainly due to the inequal values of the coefficient of thermal expansion within the whole temperature range of the process. Austenite steel and 50/50 steel have higher coefficient of thermal expansion values than their neighboring material layers and thereby contract more.

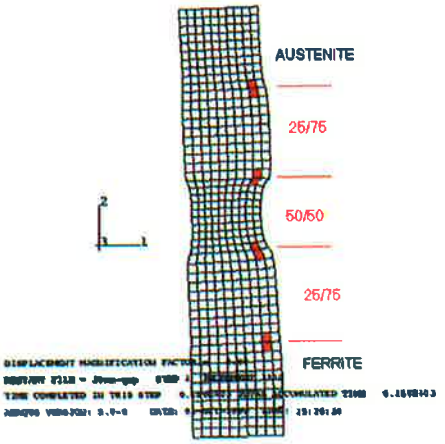


Figure 6. Displaced shape of the FG-region of the component. The displacement magnification factor is 5. The interfaces between each material have been marked. The elements used in Table 1 are painted red. Case 3.

Von Mises stress distribution contour plots from the final state in every case are shown in Figure 7. The stress distributions are very similar, except in Case 1 there are also rather high stresses near the lowest interface.

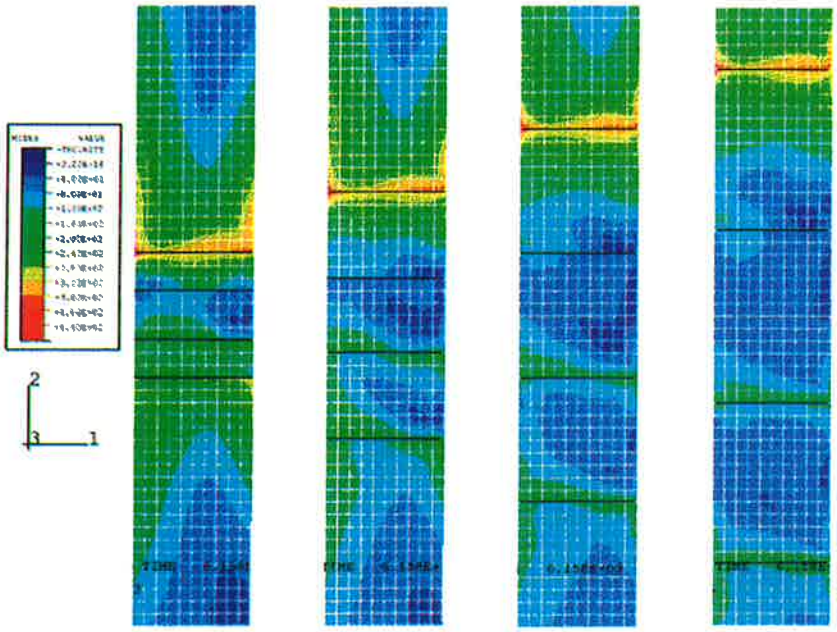


Figure 7. Contour plots of von Mises stress distribution in Cases 1-4, from left to right. The material interfaces have been marked. The scale of the legend is from zero to 440 MPa and the maximum values in Cases 1-4 are 482 MPa, 452 MPa, 449 MPa, 455 MPa, respectively.

It is naturally very essential to know the mechanisms how the stresses and strains have developed. The causal connections are very complex in the system. Different stress and strain variables as a function of time give a general insight into the mechanisms. The development of the variables shown in Table 1 in an integration point of a certain element is shown in Figure 8. The values are from the element in the austenite layer, near the 25/75 steel (the highest shaded element in Figure 6), in Case 3.

First of all, the clear connection between $\overline{q}$ and σ_{12} is seen. The curves look otherwise similar except $\overline{q}$ is over twice as high by magnitude. The stresses behave very unstable at least during the first half an hour in all the 8 elements studied in Table 1. This seems to be a situation in the whole model. Both $\overline{q}$ and σ_{12} stabilize, as shown in Figure 8, after approximately 1000 seconds (when the temperature is approx. 800°C) in a sense that they increase continuously. Since they both develop in a very similar way, no other stress component most probably has a very different behavior. Just when $\overline{q}$ and σ_{12} stabilize, the creep straining stops and $\overline{\varepsilon}^{cr}$ stays constant thereafter. Plastic deformation mainly takes place during the first 300 seconds (until approx. 1000°C). After 2500 seconds the material starts to yield again. That occurs only near the examined element and seems to slow the development of stresses. It is noteworthy that some of the material properties at high temperatures were approximated or extrapolated.

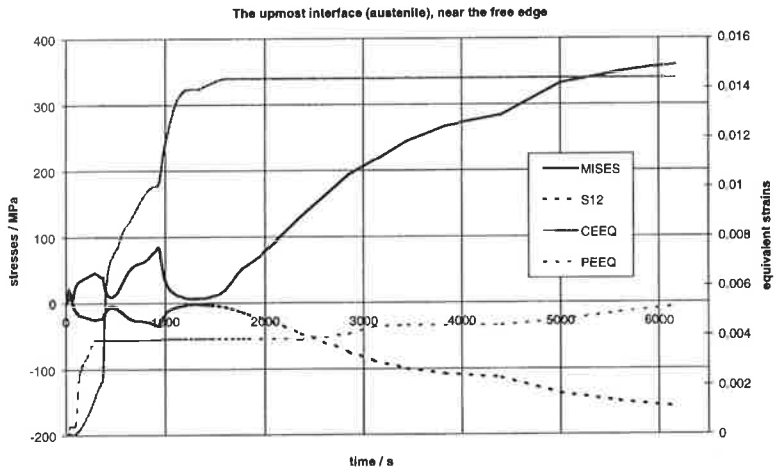


Figure 8. Von Mises stress (MISES on the legend), shear stress (S12), the equivalent creep (CEEQ) and plastic strain (PEEQ) as functions of analysis time. Case 3. Values are from the highest element painted red in Figure 6.

7. CONCLUSION

The height of the functionally graded region between 20 mm and 40 mm was found to have no considerable significance in this study. Anyway, 30 mm is at present the minimum

height, because shorter regions are not practical to manufacture for technical reasons. Longer FG-regions than 40 mm probably have only slightly lower residual stresses.

The upmost interface between the austenite steel and 25/75 steel is the most critical one concerning the joint strength. This would be the situation most probably also with more powder alloys; with smaller steps in proportions for the powder materials. The highest stress peaks develop thus in reality very near to the austenite steel.

The differences in the thermal expansion coefficients of the materials is the most important factor causing residual stresses. An ultimate aim would be an uniform expansion throughout the body. The 50/50 test material has a very problematic gradient of the coefficient of thermal expansion during the manufacturing process.

The plastic and creep deformations are relatively low in the simulated component and take place very early in the process. Both the plastic and creep deformations are concentrated around the 50/50 material and that is why no clear stress peaks are found in the middle part of the FG region. [4]

REFERENCES

[1] ABAQUS / Theory Manual, Version 5.8. Hibbitt, Karlsson & Sorensen, Inc., 1998.

[2] Saarenheimo, A., Heikinheimo, L., Holmström, S., Santaoja, K., 1999. Numerical Simulations of Thermal Residual Stresses in a Multimaterial Component - A Hot Isostatically Pressed Steel to MMC Assembly. Report VALB393.

[3] Heikinheimo, L., 1999. Materials Properties for Assessment of Thermal Residual Stresses. Report VALB360.

[4] Calonius, K., 2000. Numerical Simulation of Thermal Residual Stresses in a Functionally Graded Material Component. Report VALB440.

CALCULATION OF VOLUME CHANGE OF A POLYESTER RESIN DURING CURING

H. LAHTINEN
Department of Mechanical Engineering
University of Oulu
P.O.BOX 4200
90014 University of Oulu, FINLAND

ABSTRACT

A procedure for calculation of volume change of a polyester resin is presented. A kinetic model for simulation of chemical reactions of the polyester resin is developed and coupled with thermal conduction analysis. The subsequent structural analysis takes into account the chemical shrinkage, the thermal expansion phenomenon and linear viscoelastic material behaviour for the polyester resin.

INTRODUCTION

Residual stresses evolve in composite materials during their manufacturing process. There the properties of matrix resin change from liquid state to solid state while the properties of reinforcing fibres remain practically unchanged. In the manufacturing process thermosetting resin material is usually hardened by thermal treatment accentuated by exothermic chemical reactions. The residual stresses develop in the material mainly due to volume change of the matrix polymer, non-uniform degree of cure and temperature fields, and spatially varying material properties, which are caused by thermal and pressure cycles used in the process, chemical reactions containing volumetric shrinkage, thermal expansion phenomenon and lamination parameters like ply thicknesses and fibre directions.

Traditionally the analyses of residual stresses were based on thermal expansion mismatch and uniform temperature difference between the cure and the room temperature. Later Bogetti *et al.* [1] improved the analysis by solving a thermal conduction problem with a phenomenological model for reaction kinetics followed by a mechanical model based on the lamination theory for strain development. White *et al.* [2] carried on the analysis by assuming the material to exhibit linear viscoelastic behaviour. Hilton *et al.* [3] used nonlinear viscoelastic constitutive relations and applied the finite element method for laminates under a generalized plane strain state.

In this study a procedure for the calculation of volume change of a polyester resin is presented. The analysis is carried out by using the finite element method. A model for the chemical reaction simulation is developed and coupled with the thermal conduction analysis and the results obtained are used in subsequent structural analyses as body forces. The structural analysis takes into account the chemical shrinkage, the thermal expansion phenomenon and linear viscoelastic material behaviour for the polyester resin. The coal for the development of the calculation procedure is to predict the residual stresses of composite laminates. The residual stresses can be analysed already during the cooling phase of the manufacturing process [4]. However, the actual problem in the analysis that must be solved is the description of the material properties during curing.

CHEMICAL REACTION MODEL FOR POLYESTER RESIN

A kinetic model for the simulation of chemical reactions of a polyester resin is developed. For the model isothermally measured reaction heat H_T was predicted at 50, 60, 70, 75, 80, 90, 100 °C by differential scanning calorimetry (DSC) and after each run a dynamic temperature scan from the room temperature to 200 °C at a 10 °C/min heating rate was done to measure residual reaction heat H_R [5]. At different temperatures the cure rate is

$$\frac{d\alpha}{dt} = \frac{dH/dt}{H_T+H_R} \quad , \tag{1}$$

where H_t is the heat evolved at time t. According to the results the reaction rate increased strongly as the temperature was raised. The maximum degree of cure was reached at 75 °C below which perfect curing was not reached even after dynamic scanning.

Lam [6] studied the isothermal cure kinetics of an unsaturated polyester resin and parameters for a kinetic model by DSC. In his model the cure rate was written in the form

$$\frac{d\alpha}{dt} = Ae^{(-E/RT)}\ \alpha^m(1-\alpha)^n \quad , \tag{2}$$

where m and n are reaction orders, A and E parameters of the Arrhenius equation and α is the degree of cure. Here, the reaction orders m and n are constant and they are coupled together which causes the shape of the cure rate curve to be constant at all temperatures against the experimental evidence. Also, with this model the maximum degree of cure is always reached.

To improve the accuracy of the simulation, the equation of the cure rate was modified to a more complicated form

$$\frac{d\alpha}{dt} = Ae^{(-E/RT)}\ \alpha^m(\alpha_{max}-\alpha)(1-0.2\alpha)^2 \quad . \tag{3}$$

Below the glass transition temperature of fully cured polyester the maximum degree of cure

α_{max} is assumed to depend on the maximum curing temperature $T_{max}(t)$ at time t and the glass transition temperature $T_g(\alpha)$ as follows

$$\alpha_{max} = \frac{T_{max}(t) - T_g(0)}{T_g(1) - T_g(0)} \quad . \tag{4}$$

When $T_{max} > T_g(1)$, $\alpha_{max} = 1$. Differentiating Eq. 3 with respect to α and setting the derivative equal to zero, we obtain the value for the reaction order m at different isothermal temperatures when α in the equation is the degree of cure at the moment of maximum cure rate. Thus, the relation between the reaction order m and the temperature is obtained and an appropriate curve can be fitted to get a formula for m(T). Finally, the model is fitted to the experimental data at different isothermal temperatures and the relation between the Arrhenius equation and the temperature is found which enables us to solve values for the parameters A and E.

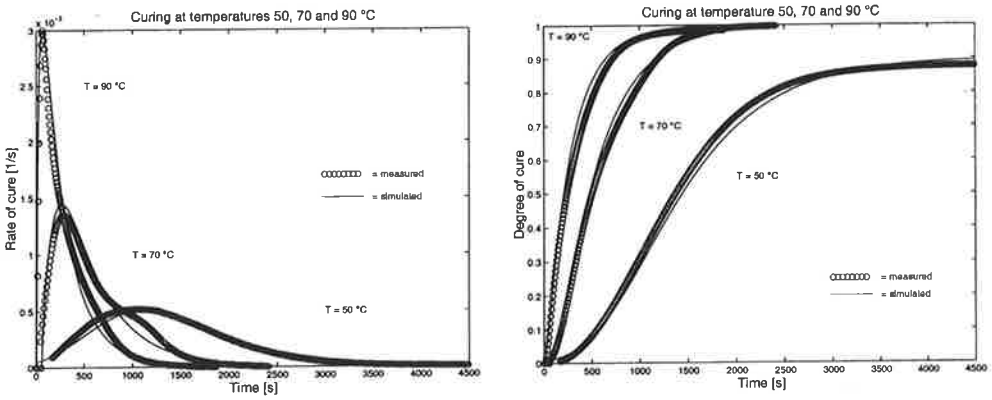


Fig. 1: The simulated rate of cure and degree of cure at different temperatures compared with the measured values.

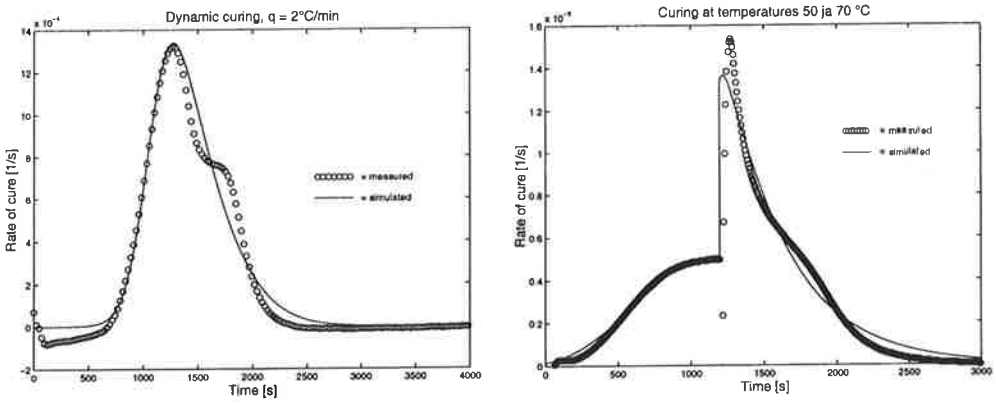


Fig. 2: The simulated rate of cure with dynamic scanning and with isothermal scanning at two different temperatures compared with the measured values.

In Fig. 1 the isothermally measured and calculated results are compared at different temperatures. The results fit together well except during the small step which appears in the measured curve of the cure rate and which is probably caused by an extra reaction type of polyester resin showing itself best at temperatures 70 - 80 °C. The "extra" step can be well seen in Fig. 2 where the resin is dynamically cured at heating rate 2 °C/min. The last comparison of the results was carried out by curing the resin isothermally first at 50 °C and then changing the temperature suddenly to 70 °C (see Fig. 2). As can be seen the simulation follows the measured data well.

THERMAL ANALYSIS

The thermal analysis contains a transient thermal conduction problem where the heat production of the chemical reactions acts as loading. At every time step and at every node of the model the cure rate, the degree of cure and the production of heat are calculated. The instantaneous heat production per unit volume of a polymer is

$$\dot{q} = \rho H_{TOT} \frac{d\alpha}{dt} \quad , \tag{5}$$

where $H_{TOT} = H_T + H_R$ is the total heat of reaction released during a complete cure.

Material properties of the polyester resin change during the cure while properties of reinforcing fibres can be considered constant. Here, the linear material properties of the resin are assumed to change linearly as a function of degree of cure and temperature. For example, the heat capacity is expressed as

$$c_p(\alpha,T) = c_p(0,T) + \alpha[c_p(1,T) - c_p(0,T)] \quad . \tag{6}$$

The temperature dependence of the heat capacity and the thermal conduction coefficient of the uncured and fully cured polyester were experimentally tested. The density was measured as a function of degree of cure and the result showed linear relationship.

A cylindrical structure of the polyester resin (height 75 mm, diameter 40 mm) was cured at the room temperature in a silicone mould and an optical fibre was embedded in the resin to measure its change of length during the cure (see Fig. 3). Furthermore, the temperature field history was measured by thermoelements from nine locations.

The specimen was modelled by 8-noded axisymmetric elements with 4*15 element mesh of the rotational plane (see Fig. 3). Also, the silicone mould was modelled and in its boundaries the convective heat transfer coefficients with the ambient temperatures were used as boundary conditions. The reaction heat liberated from the resin was the only loading of the problem. The material properties were updated in every time step according to the average degree of cure and temperature of elements. The equations were solved by full Newton-Raphson algorithm and the size of time steps was varied according to the change of temperature.

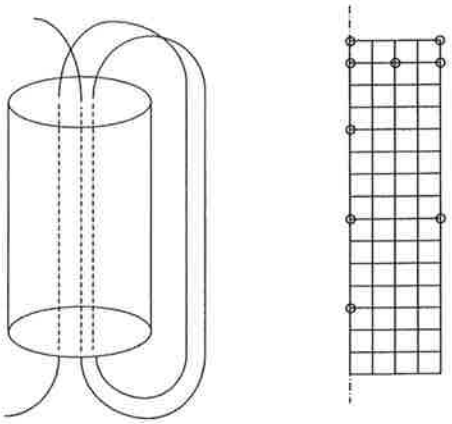


Fig. 3: The geometry of the specimen, the location of the optical fibre and the element mesh with the nine locations from which the temperature histories were measured.

In Fig. 4 the difference between the measured and calculated temperature histories can be qualitatively observed and in Fig. 5 the calculated cure rate and degree of cure are presented with respect to the temperature change of the specimen. The curing of the resin starts slowly accelerating strongly towards the complete cure. The specimen warms from the room temperature to the maximum temperature (160 °C) approximately in 10 minutes.

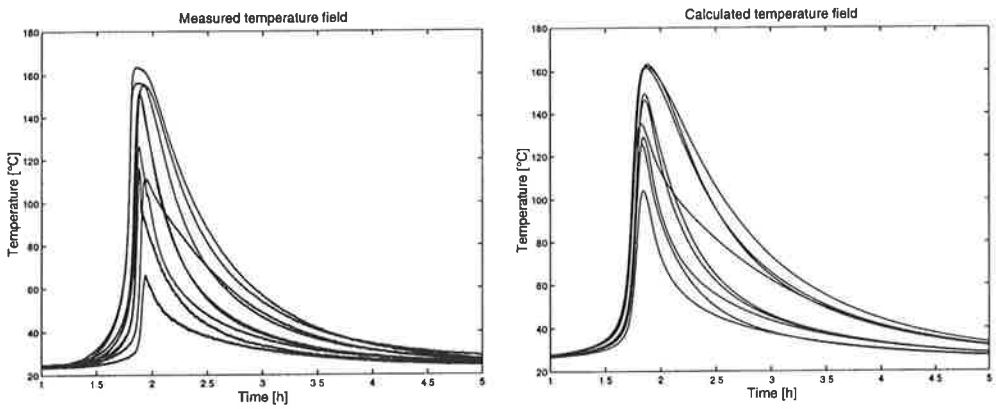


Fig. 4: The measured and calculated temperature histories of the specimen.

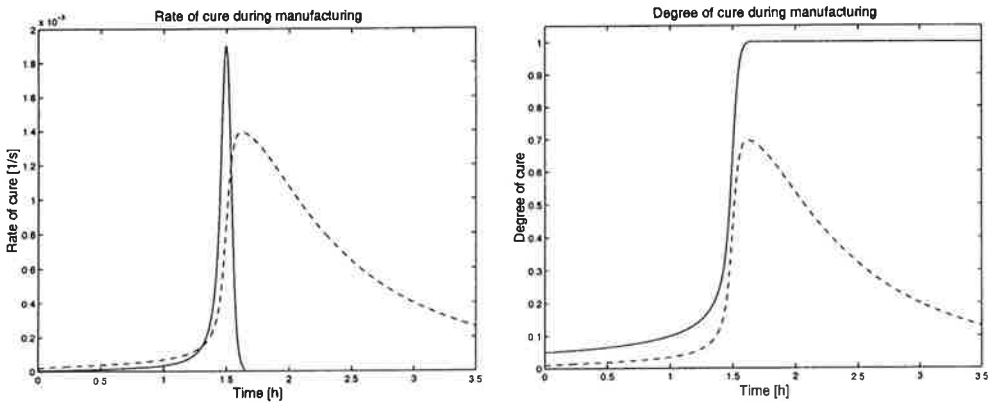


Fig. 5: The cure rate and the degree of cure compared with the temperature history (dashed line).

STRUCTURAL ANALYSIS

The thermal analysis is followed by a structural analysis that uses the previously calculated temperature histories as body loads. The structural analysis is a time dependent static problem where large displacements and strains can occur due to the chemical shrinkage of the cure. The instantaneous material properties of the resin change according to temperature and time when they are loaded by an external force or a change of temperature. The material nonlinearity is now described by a linear viscoelastic material model. For the model the shear relaxation modulus of the polyester resin was measured and the same relaxation times were assumed to the shear modulus, the bulk modulus and the structural relaxation. Also, the thermal expansion coefficient of the resin was measured as a function of temperature. The properties of the reinforcements were considered constant.

Unfortunately, the change of viscoelastic material properties during calculation is not possible in the finite element program used in the study. The restriction prevents a proper description of the polymer stiffness during the curing process. Also, when the change of thermal expansion coefficient during the analysis is not allowed, the chemical shrinkage and the thermal expansion cannot be simultaneously taken into account. However, the problem concerning the change of volume is bypassed by defining a new thermal history for the curing phase of the manufacturing process such that the new thermal history produces a volume change that is equal to the cooperative action of the chemical shrinkage and the thermal expansion.

The chemical shrinkage of the polyester resin is 10 % and it is measured to depend linearly on the degree of cure. In the thermal analysis the temperature and the degree of cure histories are solved which enables us to predict the volume change of the resin at each time step taking into account the chemical shrinkage and the thermal expansion. Thereafter, using the thermal expansion coefficient measured for the resin in the structural analysis, the new temperature history can be predicted backwards from the moment where the resin becomes fully cured. In

Fig. 6 the original and the new thermal history for the cylindrical specimen made of the polyester resin are presented.

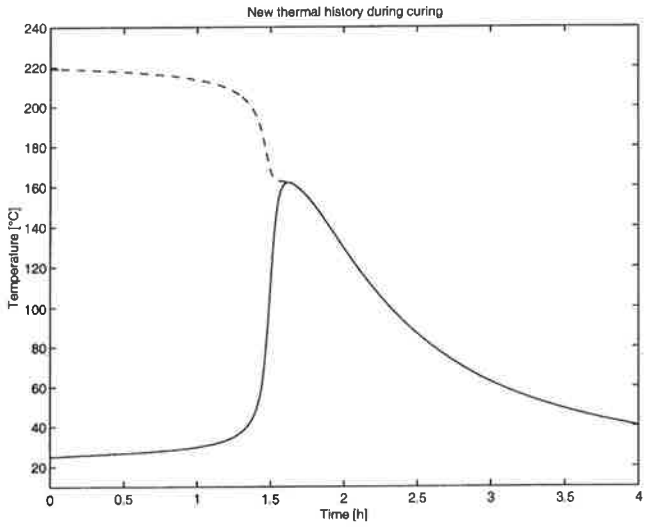


Fig. 6: The original and the new thermal history for the cylindrical specimen of the polyester resin.

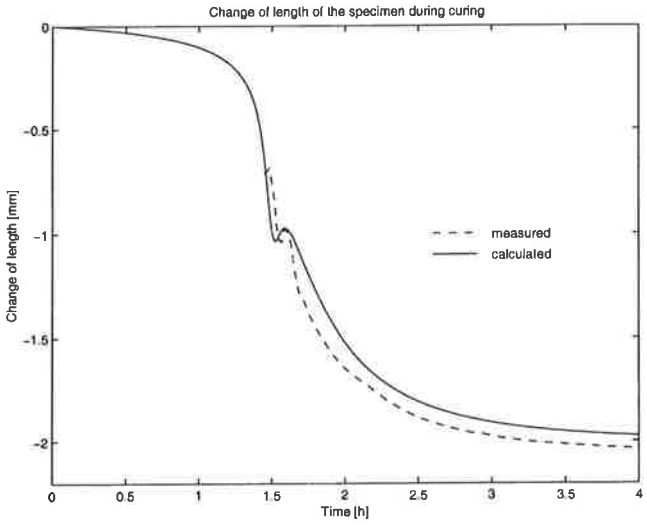


Fig. 7: The calculated and the measured change of length of the specimen.

In Fig. 7 the calculated and the measured change of length of the specimen are presented. In the end of the cure there is a small turn in the curves that might result from a small phase difference between the chemical shrinkage and the thermal expansion. Then, the chemical

shrinkage ceases while the temperature of the specimen continues to grow for a moment which causes the sudden growth of the specimen. Once the specimen starts to cool, the length of the specimen diminishes again.

An optical fibre was embedded to the specimen and the change of its length was measured. Also, the fibre was added to the model presented in the Fig. 3. In Fig. 8 the measured and the calculated results are compared. In the response of the optical fibre there is again a small turn in the end of the cure, i.e. 1.5 h's from the beginning of the cure. The peak of the response is caused by the chemical shrinkage of the resin which forces the fibre to shorten. The relaxation time of the resin is still small in the end of the cure and, therefore, most of the compression of the fibre is released. During the cooling phase the length of the specimen decreases and the fibre is again forced to shorten. This time the shortening is permanent. When the structural analysis is carried out with the measured material properties of the fully cured resin, the optical fibre shortens permanently also during the cure which markedly differs from the experimental result. After the cure the calculated and the measured results are quite close to each other. When the stiffness of the resin, especially the long-term shear modulus, was diminished remarkably, the calculated result is similar to the experimental result.

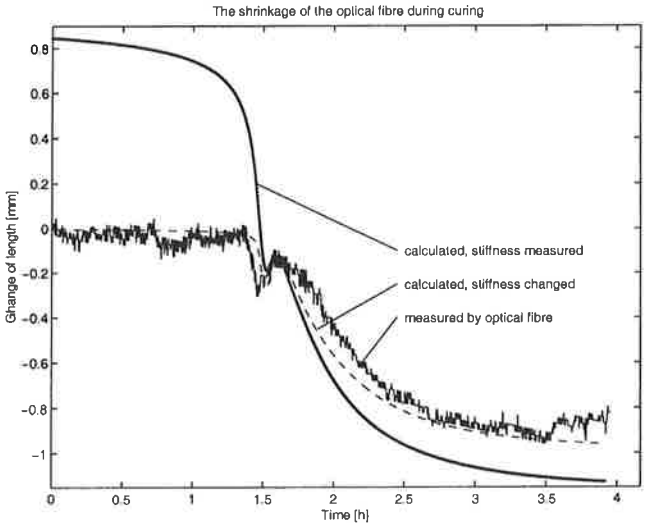
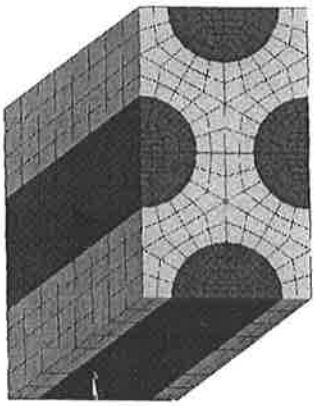


Fig. 8: The calculated and the measured change of length of the optical fibre in the specimen.

Next, a small representative volume element of polyester resin and glass fibres (see Fig. 9) was assumed to cure at 80 °C when the structure was heated and cooled at 5 °C/min. The structure was analysed by assuming that the reinforcing fibres are perfectly bonded to the resin and that the model boundaries remain straight planes, i.e. in the boundaries the displacement components perpendicular to the surface were coupled together. The development of maximum von Mises stress in the resin was calculated by using the stiffness properties measured for the resin and by using the changed stiffness properties used for the resin in the case of the optical fibre. In the Fig. 9 the comparison of the results is presented and as can be

seen the results differ remarkably from each other. Therefore, it is essential to know the actual stiffness properties of the resin during the manufacturing process when analysing the residual stresses of composite material.



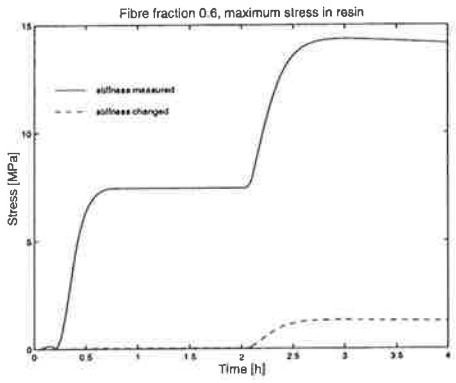


Fig. 9: The finite element model used for the representative volume element and the growth of maximum stresses in the resin during the cure.

CONCLUSIONS

A procedure for calculation of the chemical shrinkage of a polyester resin is presented. A simulation model of chemical reactions of a polyester resin is created and it was noticed accurate in the verifying measurements. The simulation model is coupled with thermal conduction analysis where the temperature field of the structure is solved during manufacturing process. The subsequent structural analysis is solved by using a linear viscoelastic material model.

In the finite element method program used it is not possible to change the viscoelastic material properties during the calculation. The restriction prevents a proper description of the polymer stiffness during the curing process. Also, when the change of thermal expansion coefficient during the analysis is not allowed, the chemical shrinkage and the thermal expansion cannot be simultaneous taken into account. However, the problem concerning the change of volume is bypassed by defining a new thermal history for the curing phase of the manufacturing process such that the new thermal history produces a volume change that is equal to the cooperative action of the chemical shrinkage and the thermal expansion.

The results obtained in the structural analyses of the cylindrical specimen and the representative volume element reveal the significance of the stiffness properties for the development of the residual stresses.

REFERENCES

1. Bogetti, T. A., Gillespie, J. W., "Process-Induced Stress and Deformation in Thick-Section Thermoset Composite Laminates". Journal of Composite Materials, Vol. 26, No. 5, s. 626 - 660, 1992.

2. White, S. R., Hahn, H. T., "Process Modeling of Composite Materials: Residual Stresses Development during Cure. Part I. Model Formulation", *Journal of Composite Materials*, 1992, **26**, No. 16, pp. 2402 - 2422.

3. Hilton, H. H., Yi, S., "Stochastic Delamination Simulations of Nonlinear Viscoelastic Composite during Cure". *Proceedings of the 4th International Conference on Stochastic Structural Dynamics*, 1998, South Bend, IN, USA.

4. Lahtinen, H., "Calculation of residual stresses of composite materials". *Proceedings of the 3th Nordic Meeting on Materials and Mechanics*, May 8 - 11, 2000, Rebild Bakker Conference Centre, Denmark.

5. Tabell, H., *"A study of curing of thermosetting resins by differential scanning calorimetry"* (in finnish), Laboratory of Chemistry, University of Oulu, Finland (2000).

6. Lam, P. W. K., "The Characterization of Thermoset Cure Behavior by Differential Scanning Calrorimetry. Part I: Isothermal and Dynamic Study", *Polymer Composites*, 1987, **8**, No.6, pp. 427-436.

JÄÄNNÖSJÄNNITYSTEN MITTAAMINEN

M. LINDGREN, T. LEPISTÖ
Materiaaliopin laitos
Tampereen teknillinen korkeakoulu
PL 589
33101 Tampere

TIIVISTELMÄ

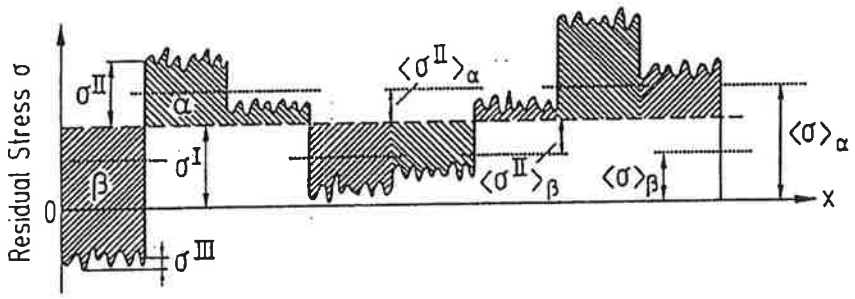
Useimmissa komponenteissa on valmistuksen jäljiltä jäännösjännityksiä, jotka voivat vaikuttaa ominaisuuksiin. Materiaali- ja valmistustekniikan kehitys on luonut uusia haasteita jäännösjännitysten mittaamiselle, koska komposiittimateriaalien ja pinnotteiden käyttö on yleistynyt. Nykyisten jäännösjännitysten mittausmenetelmien, joita on sekä ainettarikkovia että ainettarikkomattomia, käytettävyys on materiaali- ja sovelluskohtaista. Tässä esityksessä tarkastellaan erityyppisiä jäännösjännitysten mittausmenetelmiä. Yksityiskohtaisemmin esitellään reiänporaus ja röntgendiffraktio, jotka ovat yleisimmin käytettyjä jäännösjännitysten mittaamiseen.

JOHDANTO

Jäännös- eli sisäisiksi jännityksiksi kutsutaan jännityksiä, joita on kappaleessa, vaikka siihen ei vaikuta ulkoista kuormitusta ja lämpötilajakauma on tasainen. Määritelmän mukaan jäännösjännitysten aiheuttamien voimien ja momenttien summa kappaleessa on nolla. Jäännösjännitykset jaotellaan kolmeen tyyppiin riippuen etäisyydestä, jolla niiden aiheuttamien voimien ja momenttien kumoutuminen tapahtuu. Tyypin yksi jäännösjännitykset (σ^{I} = makrojännitykset) vaikuttavat laajalla alueella ja niiden laukaiseminen aiheuttaa kappaleen makroskooppisia mittamuutoksia. Tyypin kaksi jäännösjännitykset (σ^{II} = mikrojännitykset) vaikuttavat esimerkiksi materiaalien eri faasien, eli fysikaalisilta ja kemiallisilta ominaisuuksiltaan poikkeavien alueiden, välillä. σ^{II}:n suuruus on kunkin faasin keskimääräinen poikkeama vallitsevasta makrojännitystasosta. Tyypin kolme jäännösjännityksiä (σ^{III} = epähomogeeniset mikrojännitykset) esiintyy esimerkiksi hilavirheiden ympäristössä. Kuvassa 1 on havainnollistettu eri jäännösjännitystyyppejä ja niiden vaikutusalueita faaseista α ja β koostuvassa materiaalissa [1,2,3].

Jäännösjännitykset aiheutuvat joko materiaaliominaisuuksista, komponentin valmistusmenetelmistä tai ulkoisista olosuhteista, kuten kuormituksesta tai kemiallisesta ympäristöstä. Kaikki valmistusmenetelmät, joiden synnyttämät plastiset muodonmuutokset ovat epähomogeenisiä, aiheuttavat jäännösjännityksiä. Lisäksi monifaasimateriaaleilla erot

elastisuusvakioissa ja lämpölaajenemiskertoimissa aiheuttavat mikrojännityksiä. Kuvaan 2 on koottu joitakin jäännösjännityksiä aiheuttavia tekijöitä [3,4].



Kuva 1 Makrojännitykset σ^{I} ja mikrojännitykset σ^{II} ja σ^{III} faasien α ja β muodostamassa materiaalissa [3].

Materiaaliominaisuudet	Muodonanto ja liittäminen
• **Erot materiaaliominaisuuksissa** monifaasimateriaalit komposiittirakenteet pinnoitteet • **Faasimuutokset**	• **Valmistusmenetelmät** valaminen muokkaus ruiskuvalu • **Liittämismenetelmät** • **Muodonantomenetelmät**
Käsittelyt	**Kuormitus ja ympäristö**
• **Lämpökäsittelyt** • **Pintakäsittelyt** • **Pinnoitus**	• **Epätasainen kuormitus** • **Epätasainen lämpötilajakauma** • **Väsyttävä kuormitus** • **Veden absorptio**

Kuva 2 Jäännösjännityksiä aiheuttavia tekijöitä.

Jäännösjännitykset voivat joko huonontaa tai parantaa materiaalin ominaisuuksia. Jäännösjännityksiä hyödynnetään esimerkiksi aiheuttamalla väsyttävän kuormituksen alaisena olevan komponentin pintaan puristusjännityksiä, jolloin sen väsymiskestävyys paranee [5,6]. Vetojännitykset yleensä huonontavat väsymiskestävyyttä. Lisäksi ne voivat altistaa jännityskorroosiolle tai jännityssäröilylle. Jäännösjännitykset vaikuttavat myös kulumiseen ja pinnoitteen ja alustamateriaalin tai matriisin ja lujitekuidun väliseen tartuntaan [7]. Jäännösjännitysten laukeminen käytön aikana voi muuttaa kappaleen dimensioita [8].

Vaikka nykyään useiden valmistusmenetelmien aiheuttamia jäännösjännityksiä voidaan mallintaa numeerisesti, jäännösjännitysten mittaamisen tarve ei ole vähentynyt, koska numeeriset mallit tarvitsevat tuekseen vertailun kokeellisesti mitattuihin arvoihin [9-10]. Jäännösjännitysten mittaamiseen on olemassa useita menetelmiä, joilla on omat etunsa ja rajoituksensa, joiden ymmärtäminen antaa pohjan mittaustulosten arvioinnille.

1. JÄÄNNÖSJÄNNITYSTEN MITTAUSMENETELMÄT

Jäännösjännitysten mittaamisen periaate on mitata jotakin jännityksestä riippuvaa suuretta ja määrittää tämän avulla jäännösjännitystila. Menetelmät jaotellaan ainettarikkoviin ja ainettarikkomattomiin eli NDT-menetelmiin. Mekaaniset menetelmät ovat ainettarikkovia tai puolirikkovia, mikäli vaurio on korjattavissa, sillä niissä jäännösjännitysten tasapainotilaa muutetaan materiaalia poistamalla. Aiheutetusta vahingosta huolimatta mekaanisia menetelmiä käytetään paljon, koska niillä saavutettava mittaussyvyys on suurempi kuin yleisimmillä ainettarikkomattomilla menetelmillä.

Ainettarikkomattomat menetelmät perustuvat jännityksestä riippuvan suureen, kuten hilatasojen välisen etäisyyden tai Barkhausen-kohinan amplitudin, mittaamiseen. Yleensä jännityksen ja mitattavan suureen väliseen riippuvuuteen vaikuttaa myös materiaalin mikrorakenne. Yleisimmin käytetty NDT-menetelmä jäännösjännitysten mittaamiseen on röntgendiffraktio. NDT-menetelmien soveltuvuus jäännösjännitysten mittaamiseen on materiaalikohtaista, kun taas mekaaniset menetelmät ovat periaatteessa sovellettavissa kaikille materiaaleille [1,2].

1.1 Mekaaniset menetelmät

Mekaaniset menetelmät jäännösjännitysten mittaamiseen perustuvat siihen, että jäännösjännitysten tasapainotilaa muutetaan materiaalia poistamalla, jolloin jäännösjännitykset hakeutuvat uuteen tasapainotilaan. Tästä aiheutuva muodonmuutos mitataan ja ennen materiaalinpoistoa vaikuttanut jäännösjännitystila voidaan laskea. Mekaaniset menetelmät mittaavat ainoastaan makrojännityksiä [1,11].

2.1.1 Reiänporaus

Reiänporauksessa kappaleen pintaan kiinnitetyn venymäliuskarusetin keskelle porataan reikä, jolloin jäännösjännitykset laukeavat. Tämä aiheuttaa reiän lähistössä jäännösjännitysten uudelleenjärjestäytymistä, jonka aiheuttamat muodonmuutokset rekisteröidään venymäliuskoilla [12-13]. Mitatuista venymistä voidaan laskea ennen porausta vallinneet jäännösjännitykset:

$$\sigma_{max}, \sigma_{min} = \frac{\varepsilon_1 + \varepsilon_3}{4A} \pm \frac{\sqrt{(2\varepsilon_2 - \varepsilon_1 - \varepsilon_3)^2 + (\varepsilon_1 - \varepsilon_3)^2}}{4B} \qquad (1)$$

missä σ_{max} ja σ_{min} ovat pääjännitykset, ε_i mitatut venymät, A ja B kalibrointikertoimet.

Jäännösjännitysten laskeminen mitatuista venymistä edellyttää kalibrointikertoimien A ja B tuntemista. Mikäli porattava reikä lävistää materiaalin, kalibrointikertoimet voidaan ratkaista analyyttisesti [14]. Käytännössä jäännösjännityksiä mitataan usein paksuseinäisistä kappaleista, jolloin reikä ei mene seinämän läpi vaan on ns. blind-hole. Tällöin menetelmä on standardoitu, kun jäännösjännitykset eivät vaihtele syvyyssuunnassa, eivätkä ylitä 50 %:a materiaalin myötölujuudesta. Standardissa [14] on taulukoituna kalibrointikertoimet yleisimmille venymäliuskaruseteille. Standardin soveltamisen edellytyksenä on jäännösjännitysten tasaisuus syvyyssuuntaan, jonka arvioimiseksi mitattuja venymiä vertaillaan standardikuvaajiin. Vertailua on kritisoitu kirjallisuudessa [15], sillä poikkeamia standardikuvaajista voi tulla, vaikka jäännösjännitykset eivät vaihtelisi syvyyssuuntaan. Mikäli jäännösjännitykset vaihtelevat syvyyssuunnassa, kalibrointikertoimien laskemiseen on kirjallisuudessa esitetty useita menetelmiä [16-21].

Reiänporauksen mittaussyvyys on noin 1 mm riippuen venymäliuskarusetista ja porattavan reiän halkaisijasta [15, 19]. Koska jäännösjännitysten laskeminen perustuu lineaaris-elastiseen teoriaan, jäännösjännitykset eivät saa olla liian korkeita, jotta reiän ympäristön plastisoituminen ei aiheuta virhettä tuloksiin. Kirjallisuudessa on esitetty arvioita, että plastisoitumisesta johtuva virhe on mitätön, kun jäännösjännitykset ovat alle 50-70 %:a materiaalin myötörajasta, ja myötörajan suuruisten jäännösjännitysten aiheuttama virhe on luokkaa 10-20 %:a [12, 22-23]. Muita reiänporauksessa virhettä aiheuttavia tekijöitä ovat muun muassa reiän epäkeskisyys ja porauksessa tai venymäliuskojen kiinnityksessä mahdollisesti aiheutuneet jännitykset.

Reiänporausta käytetään eniten metallien jäännösjännitysten mittaamiseen, mutta myös jonkin verran muille materiaaliryhmille kuten muoveille. Menetelmä on paljon tutkittu ja luotettava. Haittapuolia ovat hitaus ja ainettarikkovuus.

2.1.2 Kerroksen poisto

Jäännösjännitysten mittaaminen kerroksen poisto -menetelmällä perustuu siihen, että levymäisestä koekappaleesta poistetaan tasainen kerros materiaalia. Tällöin jäljelläolevat jäännösjännitykset hakeutuvat uuteen tasapainotilaan, minkä vaikutuksesta koekappale käyristyy. Kun koekappaleen käyryys mitataan, voidaan ennen kerroksen poistoa vaikuttanut jäännösjännitys laskea seuraavasti [24]:

$$\sigma_x(z_1) = \frac{-E}{6(1-\nu)}\left\{(z_0+z_1)^2\left[\frac{d\,\rho_x(z_1)}{dz_1}+\nu\frac{d\,\rho_y(z_1)}{dz_1}\right]+4(z_0+z_1)\left[\rho_x(z_1)+\nu\rho_y(z_1)\right] - 2\int_{z_0}^{z_1}\left[\rho_x(z_1)+\nu\rho_y(z_1)\right]dz\right\} \quad (2)$$

missä σ_x on jäännösjännitys x-suuntaan, E ja ν elastisuusvakiot, z_0 kappaleen pinnan alkuperäinen z koordinaatti, z_1 kerroksen poiston jälkeinen z-koordinaatti ja ρ_i kaarevuus i-suuntaan

Menetelmän käytössä oletetaan, että jäännösjännitykset vaihtelevat ainoastaan kappaleen paksuussuuntaan. Jakaumaa syvyyssuunnassa voidaan tarkastella poistamalla useita kerroksia peräkkäin. Kerrosten poisto voidaan tehdä kemiallisesti kiillottamalla tai mekaanisesti. Menetelmää sovelletaan esimerkiksi muoveille, muovikomposiiteille ja pinnotteille [25-27].

2.1.3 Muut mekaaniset menetelmät

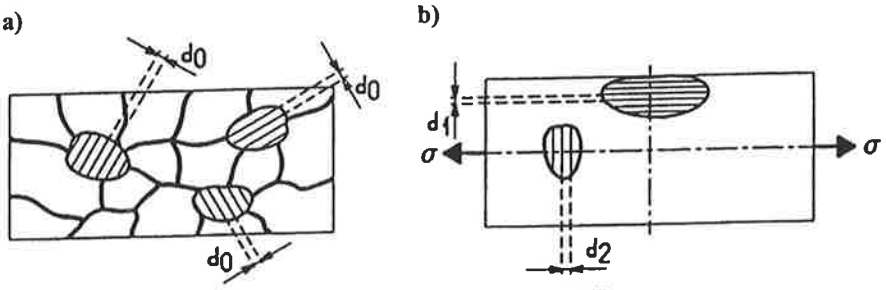
Mekaanisiin menetelmiin kuuluu esimerkiksi rengasuramenetelmä (englanniksi ring-core) [28], joka muistuttaa reiänporausta, mutta reiän asemasta työstetään rengasmainen ura venymäliuskarusetin ympärille. Menetelmän etuna reiänporaukseen verrattuna on suurempi mittaussyvyys, mutta myös kappaleelle aiheutettu vahinko on suurempi. Muita mekaanisia menetelmiä ovat esimerkiksi paloittelu (englanniksi sectioning) [29] ja renkaan katkaisu [30,31].

1.2 Diffraktiomenetelmät

Diffraktiomenetelmät käyttävät materiaalin hilatasojen välistä etäisyyttä d mikroskooppisena "venymäliuskana". Kun materiaali on jännityksettömässä tilassa, hilatasojen välinen etäisyys d_0 tietyille tasoille on sama kaikkiin suuntiin, mikä on esitetty kuvassa 3 a. Kun materiaalissa on jäännösjännityksiä, vetojännitys kasvattaa hilatasojen välistä etäisyyttä vaikutussuunnassaan ja puristusjännitys puolestaan pienentää sitä. Hilatasojen välinen etäisyys riippuu tarkastelusuunnasta. Kuvassa 3 b vetojännitys on kasvattanut vaikutussuunnassaan hilatasojen välisen etäisyyden arvoon d_2. Tällöin materiaalissa vaikuttava venymä on

$$\varepsilon = \frac{d_2 - d_0}{d_0} \tag{3}$$

missä ε on venymä, d_0 on hilatasojen välinen etäisyys jännityksettömässä tilassa, d_2 on hilatasojen välinen etäisyys jännityksen alaisena.



Kuva 3 a) Jännityksettömässä tilassa hilatasojen välinen etäisyys on sama kaikkiin suuntiin. b) Jännityksen alaisena hilatasojen välinen etäisyys riippuu tarkastelusuunnasta.

Hilatasojen välinen etäisyys saadaan laskettua mitatusta diffraktiokulmasta. Braggin lain mukaan materiaali diffraktoi säteilyä vain tietyillä kulmilla, jotka saadaan Braggin laista:

$$2\,d_{hkl}\,\sin\theta = \lambda \tag{4}$$

missä d_{hkl} on hilatasojen välinen etäisyys tasolle hkl, θ on diffraktiokulma, ja λ on käytetyn säteilyn aallonpituus

Jäännösjännitysten mittaamiseen diffraktion avulla on kaksi vaihtoehtoa: röntgendiffraktio ja neutronidiffraktio. Molemmat menetelmät ovat ainettarikkomattomia, ja niiden suurin ero on tunkeumasyvyydessä. Röntgensäteilyä käytettäessä mittaussyvyys teräksellä on tyypillisesti 10 μm:n luokkaa, kun taas neutronidiffraktiolla se on luokkaa 10-25 mm [1,2].

2.2.1 Röntgendiffraktio

Röntgendiffraktio on pintaherkkä menetelmä, jolloin jäännösjännitysmittauksissa oletetaan, että mitattavassa pintakerroksessa vaikuttaa tasojännitystila. Tällöin hilatasojen välinen etäisyys jännityksettömässä tilassa d_0, on sama kuin pinnan suuntaisten hilatasojen välinen etäisyys $d_{\psi=0}$. Kun hilatasojen välinen etäisyys mitataan useille kallistuskulmille ψ, voidaan jäännösjännitys mittaussuuntaan ϕ ratkaista kaavasta [32]:

$$\frac{d_{\phi\psi}-d_{\psi=0}}{d_{\psi=0}} = \frac{1}{2}S_2\left(\sigma_{11}\cos^2\phi+\sigma_{12}\sin 2\phi+\sigma_{22}\sin^2\phi\right)\sin^2\psi - S_1\left(\sigma_{11}+\sigma_{22}\right) + \frac{1}{2}S_2\left(\sigma_{13}\cos\phi+\sigma_{23}\sin\phi\right)\sin 2\psi \tag{5}$$

missä $d_{\phi\psi}$ on hilatasojen välinen etäisyys suuntaan ϕ kallistuskulmalla ψ, $d_{\psi=0}$ on hilatasojen välinen etäisyys kallistuskulmalla 0, S_1 ja $1/2S_2$ ovat elastisuusvakiot

Jotta mitatut venymät voidaan muuttaa jännitykseksi, tarvitaan elastisuusvakioiden S_1 ja $1/2S_2$ arvot. Jos suhteelliset jäännösjännityksen arvot riittävät, elastisuusvakiot saadaan materiaalin elastisuusvakioista E ja ν [2]. Todellisuudessa kiteisen materiaalin elastisuusvakiot riippuvat tarkastelusuunnasta eli ovat anisotrooppisia. Röntgendiffraktiolla mitattaessa vain pieni osa rakeista toteuttaa Braggin lain (kaava 4) ja vaikuttaa jäännösjännitysmittaukseen. Koska näiden diffraktoivien rakeiden elastinen käyttäytyminen voi poiketa keskimääräisestä käyttäytymisestä, elastisuusvakioina täytyisi käyttää mitatuille hilatasoille ominaisia arvoja, joita kutsutaan röntgenelastisuusvakioiksi. Röntgenelastisuusvakiot voidaan määrittää laskennallisesti lähtien erilliskiteiden elastisuusvakioista tai kokeellisesti kuormittamalla kappaletta tunnetulla jännityksellä [32].

Röntgendiffraktiolla jäännösjännitys mitataan yhdestä faasista kerrallaan. Tällöin mittaustulos on keskiarvo ko. faasille röntgensäteilyn tunkeumasyvyydellä vaikuttavista makro- ja mikrojännityksistä, joka α-faasille on merkitty kuvaan 1 symbolilla $<\sigma>_\alpha$. Jos materiaalissa on ainoastaan faasia α, röntgendiffraktiolla mitattu jäännösjännitys vastaa

makrojännitystä. Mikäli jäännösjännitykset pystytään mittaamaan monifaasirakenteessa kaikista faaseista, voidaan makro- ja mikrojännitykset erotella. Määritelmän mukaan mikrojännitykset eri faaseissa kumoavat toisensa, jolloin makrojännityksen arvo saadaan faasiosuuksilla painotettuna keskiarvona röntgendiffraktiomittauksista:

$$\sum_{f=1}^{n} c^f < \sigma >^f = \sigma^I \qquad (6)$$

missä c on faasiosuus, f merkitsee faasia, $<\sigma>^f$ on röntgendiffraktiolla mitattu jäännösjännitys ko. faasille ja σ^I on makrojännitys.

Edellä esitetty jäännösjännitysten mittaaminen perustuu d vs. $\sin^2\psi$ -suhteen lineaarisuuteen, mikä yleensä pitää paikkansa. Epälineaarisuuksia d vs. $\sin^2\psi$ -suhteeseen voivat aiheuttaa esimerkiksi merkittävät leikkausjännitykset, voimakkaat jäännösjännitys- tai koostumusgradientit sekä voimakas tekstuuri eli kiderakenteen suuntautuneisuus. Jäännösjännitysten määrittämistä näissä tapauksissa on käsitelty lähteissä [33-36].

Röntgendiffraktiota voidaan soveltaa jäännösjännitysten mittaamiseen lähes kaikille kiteisille materiaaleille. Yhteenveto röntgendiffraktion uusimmista materiaaliteknisistä sovellutuksista on lähteessä [37].

2.2.2 Neutronidiffraktio

Neutronidiffraktion käyttö jäännösjännitysten mittaamiseen perustuu samaan periaatteeseen kuin röntgendiffraktio eli mitataan materiaalin hilatasojen välistä etäisyyttä d. Koska neutronien tunkeumasyvyys metalleihin on useita senttimetrejä, hilatasojen välinen etäisyys jännityksettömässä tilassa d_0 on tunnettava, koska sama olettamus tasojännitystilasta, joka tehtiin röntgendiffraktiossa, ei tällöin päde [1,2,38,39].

Kun d_0 tunnetaan, venymät voidaan laskea kaavasta 3 ja tämän jälkeen jännitys saadaan laskettua, kun mitatut venymät ovat pääkoordinaatiston mukaisia:

$$\sigma_{ij} = \frac{E}{1+\nu}\left(\varepsilon_{ij} + \frac{\nu}{1-2\nu}\delta_{ij}\varepsilon_{kk}\right) \qquad (7)$$

Kun jäännösjännityksiä mitataan neutronidiffraktiolla, täytyy d_0 tuntea mahdollisimman tarkkaan. Virheet d_0:ssä aiheuttavat systemaattista virhettä jännitystilan hydrostaattiseen osaan, kun taas jännitystilan deviatoriseen osaan sillä ei ole vaikutusta. Hilatasojen välinen etäisyys jännityksettömässä tilassa voidaan määrittää esimerkiksi 1) mittauksella jännityksenpoistohehkutuksen jälkeen tai 2) mittauksella pulverinäytteestä tai 3) mittauksella komponentista irroitetuista pienistä koekappaleista, joista makrojännitykset ovat lauenneet [2,40]. Hilatasojen välisen etäisyyden määrittäminen voi tulla ongelmaksi esimerkiksi hitsisaumoissa, joissa se voi vaihdella paikan funktiona johtuen eroista mikrorakenteessa. Väärästä kohtaa mitattua d_0:ttä käyttämällä vetojännitykset voidaan virheellisesti määrittää puristusjännityksiksi ja päinvastoin [41].

Periaatteessa kaikkien kiteisten materiaalien jäännösjännityksiä voidaan määrittää neutronidiffraktion avulla, muutamaa poikkeusta lukuunottamatta. Poikkeuksia ovat esimerkiksi B, Cd ja Al-Ti –seokset [2]. Ensisijainen ongelma neutronidiffraktion soveltamisessa on se, että menetelmä vaatii neutronilähteen. Tämän vuoksi mittaukset ovat kalliita. Toinen rajoittava tekijä on menetelmällä saavutettava resoluutio. Yleensä mitattava tilavuus on luokkaa 1-100 mm^3 [1,2].

Neutronidiffraktiota on sovellettu muun muassa hitsien jäännösjännitysten mittaamiseen [1, 42, 43], metallimatriisikomposiiteille [40] ja $MoSi_2$-ruostumaton teräs juotoksille [44].

1.3 Muut menetelmät

Jäännösjännitysten mittaamiseen on olemassa myös muita, vain tietylle materiaalille sopivia mittausmenetelmiä. Tällöin mitataan jotakin jännityksestä riippuvaa suuretta, jonka arvo muutetaan kalibroinnin avulla jäännösjännitykseksi. Tälläisiä menetelmiä ovat esimerkiksi:

1) Barkhausen-kohina. Menetelmässä mitataan ferromagneettisilla materiaaleilla (esimerkiksi raudalla) Barkhausen-kohinan amplitudia, joka riippuu jännityksestä ja materiaalin mikrorakenteesta [1,2,45-46]
2) Ultraääni. Ultraäänen nopeus riippuu jännityksestä ja materiaalin mikrorakenteesta [1,2]
3) Kemialliset altistukset. Amorfisia muoveja voidaan altistaa tietyillä kemikaaleilla, jolloin säröily kertoo jäännösjännitysten ylittävän tietyn raja-arvon [47].

3. YHTEENVETO

Jäännösjännitysten mittaamiseen on useita menetelmiä, joilla on omat etunsa ja rajoittavat tekijänsä. Mittausmenetelmän valinnassa täytyy huomioida mitattavan materiaalin lisäksi millaista tietoa jäännösjännityksistä halutaan (suhteellisia vai absoluuttisia arvoja), onko jäännösjännityksissä gradientteja, onko menetelmän oltava ainettarikkomaton, millaiselta mittaussyvyydeltä tietoa tarvitaan jne. Käytetty mittausmenetelmä määrittää sen, mitataanko makro- ja/tai mikrojännityksiä eli millaista tietoa jäännösjännityksistä saadaan. Eri menetelmillä mitattujen jäännösjännitysarvojen vertaileminen edellyttää mittausmenetelmien fysikaalisten periaatteiden ymmärtämistä.

LÄHTEET

1. Handbook of measurement of residual stresses, J. Lu (toim.) Lilburn, Fairmont Press, 1996, 238 s.
2. Hauk, V., Structural and residual stress analysis by nondestructive methods, Amsterdam, Elsevier, 1997, 640 s.
3. Macherauch, E. Kloos, K.H., Teoksessa: Residual Stresses in Science and Technology, vol 1, E. Macherauch (toim.), 1987, s 3-27

4. Kloos, K.H., Kaiser, B., Teoksessa: Residual Stress: measurement, calculation, evaluation, V. Hauk et al. (toim.), Verlag, 1991, s 205-226
5. Desvignes, M., Gentil, B., Castex, L., Teoksessa: Proceedings of International Conference on Residual Stresses-ICRS-2, G. Beck et al. (toim.), 1989, s 791-796
6. Diepart, C.P., Teoksessa: Residual Stresses in Science and Technology, vol 1, E. Macherauch (toim.), 1987, s 457-466
7. Bauer, G., Opel, A., Rüsselsheim, A.G., Residual Stress, V Hauk et al (toim.), Verlag, 1993, s 495-504
8. Fujiwara, H., Residual Stresses in Science and Technology, vol 1, E. Macherauch (toim.), 1987, s 79-98
9. Rasouli, S., Retraint, D., Lu, J., Journal of Strain Analysis, vol 33, No. 6, 1998, s 449-458
10. Troive, L., Näsström, M., Jonsson, M., Journal of Pressure Vessel Technology, vol 120, No. 3, 1998, s 244-251
11. Kockelmann, H., Residual Stress: measurement, calculation, evaluation, V. Hauk et al. (toim.), Verlag, 1991, s 37-52
12. Beaney, E.M., Procter, E., Strain, January 1974, s 7-14
13. Beaney, E.M., Strain, July 1976, s 99-106
14. ASTM E 837-89
15. Schajer, G.S., Teoksessa: Proceedings of International Conference on Residual Stresses-ICRS-2, G. Beck et al. (toim.), 1989, s 71-77
16. Schajer, G.S., Journal of Engineering Materials and Technology, vol 103, 1981, s 157-163
17. Flaman, M.T., Mannig, B.H., Experimental Mechanics, vol 25, No. 9, 1985, s 205-207
18. Schajer, G.S., Journal of Engineering Materials and Technology, vol 110, 1988, s 338-343
19. Schajer, G.S., Journal of Engineering Materials and Technology, vol 110, 1988, s 344-349
20. Niku-Lari, A., Lu, J., Flavenot, J.F., Experimental Mechanics vol. 25, no. 2, 1985, s 175-185
21. Petrucci, G., Zuccarello, B., Journal of Strain Analysis, vol 33, No. 1, 1998, s 27-37
22. Scaramangas, A.A., Porter Goff, R.F.D., Leggatt, R.H., Strain, August 1982, s 88-97
23. Beghini, M., Bertini, L., Raffaelli, P., Journal of Testing and Evaluation, JTEVA, vol 22, No. 6, 1994, s 522-529
24. Treuting, R.G., Reed, W.T.Jr., Journal of Applied Physics, vol 22, No. 2, 1951, s 130-134
25. Turnbull, A., Maxwell, A.S., Pillai, S., Journal of Materials Science, vol 34, 1999, s 451-459
26. Wilkinson, S.B., White, J.R., Journal of Materials Science, vol 33, 1998, s 3101-3108
27. McGrann, R.T.R., Rybicki, E.F., Shadley, J.R., Teoksessa: Proceedings of International Conference on Residual Stresses ICRS-5, T. Ericsson et al. (toim.) 1997, s 994-999
28. Keil, S., Experimental Techniques, September/October 1992, s 17-24
29. Tebedge, N., Alpsten, G., Tall, L., Experimental Mechanics, vol 13, No. 2, 1973, s 88-96
30. Clutton, E.Q., Williams, J.G., Polymer Engineering and Science, vol 35, No. 17, 1995, s 1381-1386

31. Kazakov, A., Polymer Testing, vol 17, 1998, s 443-450
32. Noyan, I.C., Cohen, J.B., Residual Stress: measurement by diffraction and interpretation, Springer-Verlag, New York, 1987, 276 s
33. Hauk, V., Teoksessa: Advances in X-ray Analysis, vol 27, Cohen et al. (toim.), Plenum Press, 1984, s 101-120
34. Noyan, I.C., Metallurgical Transactions A, vol 14A, 1983, s 249-258
35. Dölle, H., Cohen, B., Metallurgical Transactions A, vol 11A, 1980, s 831-836
36. Van Houtte, P., De Buyser, L., Acta Metallurgica and Materialia, vol 41, No. 2, 1993, s 323-336
37. Lu, J., Retraint, D., Journal of Strain Analysis, vol 33, No. 2, 1998, s 127-136
38. Stacey, A., MacGillivary, H.J., Webster, G.A., Webster, P.J., Ziebeck, K.R.A., Journal of Strain Analysis, vol 20, No. 2, 1985, s 93-100
39. Allen, A.J., Hutchings, M.T., Windsor, C.G., Advances in Physics, vol 34, No. 4, 1985, s 445-473
40. Rustichelli, F., Metallurgical Science and Technology, vol 14, No 1, 1996, s 21-36
41. Krawitz, A.D., Winholtz, R.A., Materials Scince and Engineering A185, 1994, s 123-130
42. Albertini, G., Bruno, G., Carrodò, A., Measuring. Science and Technology. Vol 10, 1999 s R56-R73
43. Root, J.H., Holden, T.M., Schröder, J., Hubbard, C.R., Spooner, S., Dodson, T.A., David, S.A., Materials Science and Technology, vol 9, 1993, s 754-759
44. Vaidya, R.U., Rangaswamy, P., Bourke, M.A.M., Butt, D.P., Acta Materialia, Vol 46, No. 6, 1998, s 2047-2061
45. Jagadish, C., Clapham, L., Atherton, D., IEEE Transactions on Magnetics, vol 26, No. 3, 1990, s 1160-1163
46. Shaw, B.A., Hyde, T.R., Evans, J.T., Teoksessa: Proceedings of 1st International Conference on Barkhausen Noise and Micromagnetic Testing, 1999, s 187-196
47. Gupta, V.B., Radhakrishnan, J., Sett, S.K., Journal of Applied Polymer Science, vol 55, 1995, s 247-254

PAPERIKONEEN TELOJEN NIPPIKONTAKTIN YKSINKERTAISTETTU FEM-MALLINNUS

VELI-MATTI JÄRVENPÄÄ
Konedynamiikan laboratorio
Tampereen Teknillinen Korkeakoulu
PL 589, 33101 Tampere
03 365 2749
vmj@ruuvi.me.tut.fi

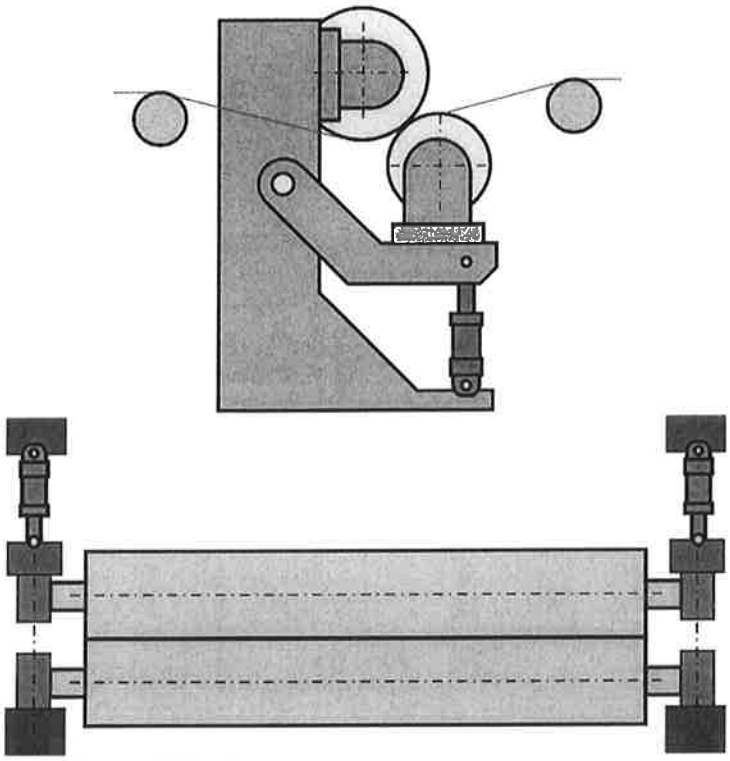
TIIVISTELMÄ

Tämän paperin tavoitteena on esittää yksinkertainen mallinnustapa kahden nippikontaktissa olevan paperikoneen telan laskentaa varten käyttämällä elementtimenetelmää ja rajoiteyhtälöfromulaatiota. Telojen väliin määritellään yhden kontaktipisteen esipuristettu ja kitkaton rajoiteyhtälö matriisimuunnoksella, ja telojen liikeyhtälöt kirjoitetaan niiden mukana jäykän kappaleen pyörimisliikkeessä olevassa koordinaatistossa. Tavoitteena on aikaansaada laskentamalli suhteellisen pitkiä aikatason simulointeja varten telan pintavärähtely- ja viiveyhtälötutkimusta varten. Mallinnustapaa testataan tasoesimerkeillä.

1. JOHDANTO

Paperinvalmistuksessa nippiyksiköitä käytetään pääasiassa paperin kalanterointi- ja päällystysprosesseissa. Ne sijaitsevat yleensä paperikonelinjan lopussa itsenäisinä yksiköinä. Kuvassa 1 on esitetty paperikoneen nippiyksikön periaatekuva. Itse nipillä tarkoitetaan kahden yhteenpuristetun telan välistä kontaktia ja nippiyksikkö koostuu telojen lisäksi telojen laakeroinnista, runkorakenteesta, kuormitusmekanismista sekä kuormittavasta hydrauliikkapiiristä, jolla nippi suljetaan. Käytön aikana paperirata kulkee yhteenpuristetun nipin lävitse, milloin paperin pinta saa halutun käsittelyn eli kalanteroinnissa paperin pinnankarheus siloitetaan ja päällystyksessä paperin pintaan lisätään haluttu pinnoite. Näiden prosessien laadun kannalta nippikuormituksen (painejakauman) suuruus on oleellinen parametri ja paperinvalmistuksessa asiaa kuvataan nipin viivakuormitus termillä.

Kalanteritelat ovat yleensä molemmat polymeeripäällysteisiä tai toinen tela on terästela ja toinen päällystetty. Kovissa nipeissä pinnoitemateriaalit nimen mukaisesti ovat suhteellisen kovia, milloin kontaktipinta-ala jää pieneksi ja viivakuorma-käsite kuvaa kontaktia hyvin.



Kuva 1. Periaatekuva nippiyksiköstä.

Nykyisin käytetään yleisemmin kuitenkin ns. pehmeitä nippejä, joissa pinnoitemateriaali on huomattavasti pehmeämpää, ja kontaktipinta-ala on vastaavasti suurempi. Käytön syy on saavutettava parempi paperin painolaatu.

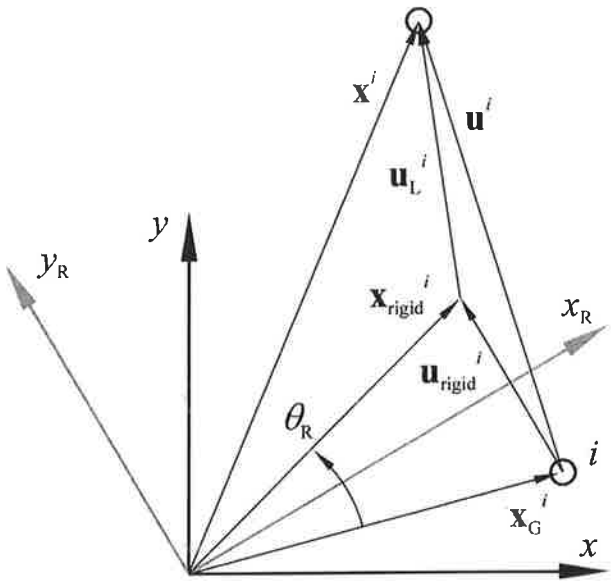
Tässä yhteydessä rajoitutaan käsittelemään ainoastaan kovan nipin kontaktiongelman mallinnusta pitäen tavoitteena yksikertaista mallinnustapaa ja nopeaa laskentaa. Aluksi oletetaan, että kontakti ei aukea ja kontaktipinta-alan muutos värähtelyjen takia oletetaan merkityksettömäksi. Telojen välinen kontakti on kuvattu sitomalla niiden elementtiverkkojen vastakkaisten solmukohtapisteiden asemat yhteen pistemäisten rajoiteyhtälöiden avulla normaalinsuunnassa, milloin telojen pyöriminen ei esty. Telojen elementtiverkot on määritelty yhdenmukaisiksi. Tämän paperin esimerkissä tarkastelu on rajattu vain tasotapaukseen, mutta periaate on yleistettävissä spin rotaatiossa olevalle 3D-telaparille, kun käytetään lineaarisia elementtejä. Kontaktimallinnuksessa ei ole otettu huomioon kitkaa.

Telat on mallinnettu roottoridynaamisena systeeminä seuraavasti. Telan ainoaksi suuriamplitudiseksi liikkeeksi on oletettu spin rotaatio eli telan pyörimisliike. 3D-hyrrävoimat on oletettu pieniksi, joten ne on jätetty pois laskentamallista. Telan jäykän kappaleen rotaatioliike on erotettu telan muodonmuutoskentästä (värähtelyliikkeestä), joka oletetaan lineaarisen kimmoteorian mukaiseksi, ja joka kuvataan käyttämällä elementtimenetelmää ja

lineaarista jäykkyysmatriisia. Hitausvoimat kuvataan käyttämällä keskitettyä massamatriisia eli elementtiverkon solmumassat ovat pistemassoja. Liikeyhtälöt on kirjoitettu telan jäykän kappaleen liikkeen mukana pyörivässä koordinaatistossa ja telan on oletettu pyörivän vakionopeudella. Tavoitteena on aikaansaada nopea laskentamalli, jota voidaan käyttää telojen pintavärähtelyjen mallinnuksessa ja pinnoitteiden viskoelastisten viiveilmiöiden tutkimuksessa.

2. TELAN ELEMENTTIMALLI

Perusoletus on siis, että jäykän kappaleen spin rotaatio θ_R voidaan erottaa muodonmuutoksista ja muodonmuutokset lasketaan käyttämällä lineaarista kimmoteoriaa. Lokaali siirtymävektori $\mathbf{u}_{RL}$ on lausuttu pyörivässä koordinaatistossa (x_R, y_R) kuvan 2 mukaisesti (vrt. [1],[2],[3] ja [4]).



Kuva 2. Solmupisteen asema- ja siirtymävektorit.

Solmupisteiden asemavektori on muotoa

$$\mathbf{x} = \mathbf{R}^T(\mathbf{u}_{RL} + \mathbf{x}_G) \tag{1}$$

missä koordinaatistonmuunnosmatriisi on

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_n & & & \\ & \mathbf{R}_n & & \\ & & \ddots & \\ & & & \mathbf{R}_n \end{bmatrix} \quad \mathbf{R}_n = \begin{bmatrix} \cos\theta_R & \sin\theta_R & 0 \\ -\sin\theta_R & \cos\theta_R & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2}$$

ja alkugeometria(koodinaatti)vektori

$$\mathbf{x}_G = \begin{bmatrix} \dots & x_G^{\,i} & y_G^{\,i} & 0 & \dots \end{bmatrix}^T \tag{3}$$

Systeemin liike-energia on tällöin muotoa

$$T = \frac{1}{2}\dot{\mathbf{x}}^T \mathbf{M}\dot{\mathbf{x}} \tag{4}$$

missä $\mathbf{M}$ on keskitetty massamatriisi ja derivointi

$$\dot{\mathbf{x}} = \mathbf{R}^T \dot{\mathbf{u}}_{RL} + \dot{\mathbf{R}}^T \mathbf{u}_{RL} + \dot{\mathbf{R}}^T \mathbf{x}_G \tag{5}$$

Kimmoenergia on muotoa

$$U = \frac{1}{2}\mathbf{u}_{RL}^T \mathbf{K}\mathbf{u}_{RL} \tag{6}$$

ja kun liike-energia ja kimmoenergia sijoitetaan Lagrangen yhtälöön

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{\mathbf{y}}}\right) - \frac{\partial T}{\partial \mathbf{y}} + \frac{\partial U}{\partial \mathbf{y}} = \mathbf{Q} \tag{7}$$

saadaan liikeyhtälöt (lopulta) muodossa

$$\begin{bmatrix} \mathbf{M} & \mathbf{M}_s(\mathbf{u}_{RL} + \mathbf{x}_G) \\ (\mathbf{u}_{RL} + \mathbf{x}_G)^T \mathbf{M}_s & (\mathbf{u}_{RL} + \mathbf{x}_G)^T \mathbf{M}(\mathbf{u}_{RL} + \mathbf{x}_G) \end{bmatrix} \begin{bmatrix} \ddot{\mathbf{u}}_{RL} \\ \ddot{\theta}_R \end{bmatrix} + \begin{bmatrix} 2\mathbf{M}_s\dot{\theta}_R & \mathbf{0} \\ \mathbf{0} & 2(\mathbf{u}_{RL} + \mathbf{x}_G)^T \mathbf{M}\dot{\mathbf{u}}_{RL} \end{bmatrix} \begin{bmatrix} \dot{\mathbf{u}}_{RL} \\ \dot{\theta}_R \end{bmatrix} + \begin{bmatrix} \mathbf{K} & \mathbf{0} \\ \mathbf{0} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{u}_{RL} \\ \theta_R \end{bmatrix} = \begin{bmatrix} \mathbf{M}\dot{\theta}_R{}^2(\mathbf{u}_{RL} + \mathbf{x}_G) \\ 0 \end{bmatrix} + \mathbf{Q} \tag{8}$$

missä $\mathbf{M}_s$ on vinosymmetrinen massamatriisi. Kun valitaan vakio pyörimisnopeus asettamalla $\ddot{\theta}_R = 0$ ja $\theta_R = \omega$, saadaan yhtälö (8) muotoon

$$\mathbf{M}\ddot{\mathbf{u}}_{RL} + 2\mathbf{M}_s\omega\,\dot{\mathbf{u}}_{RL} + \mathbf{K}\mathbf{u}_{RL} = \mathbf{M}\omega^2(\mathbf{u}_{RL} + \mathbf{x}_G) + \mathbf{Q} \tag{9}$$

Yhtälön (9) tapauksessa ulkoinen kuormitusvektori **Q** pitää valita siten, että sillä ei ole vaikutusta telan pyörimisnopeuteen. Tämän paperin kontaktimallinnuksessa ulkoista kuormitusta ei ole käytetty. Kun käytetään lokaalia asemavektoria lokaalin siirtymävektorin sijaan ja jätetään siirtymäkomponentti pois yhtälön (9) keskipakotermistä saadaan yhtälö muotoon

$$\mathbf{M}\ddot{\mathbf{x}}_{RL} + \mathbf{C}\dot{\mathbf{x}}_{RL} + \mathbf{K}\mathbf{x}_{RL} = \mathbf{f}_c + \mathbf{K}\mathbf{x}_G \tag{10}$$

missä $\mathbf{C} = 2\mathbf{M}_s\omega$ ja $\mathbf{f}_c = \mathbf{M}\omega^2\mathbf{x}_G$.

3. KONTAKTIMALLI

Nippikontakti on siis mallinnettu pistemäistä rajoiteyhtälöä käyttämällä. Rajoiteyhtälö on kirjoitettu vastinpisteiden asemille eikä siirtymille, jotta olisi mahdollista mallintaa nipin esipuristus. Rajoiteyhtälöt annetaan liikeyhtälöille käyttämällä ei-neliömäistä rajoiteyhtälömatriisia **H**, joka sitoo vapaat ja rajoitetut vapausasteet toisiinsa [5].

$$\mathbf{x}_{unconst} = \mathbf{H}\mathbf{x}_{const} \tag{11}$$

Perusajatus on määritellä kahden eri elementtiverkon (telan) systeemimatriisit globaalimatriiseissa kuten

$$\mathbf{M}_{global} = \begin{bmatrix} \mathbf{M}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_2 \end{bmatrix}, \mathbf{K}_{global} = \begin{bmatrix} \mathbf{K}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_2 \end{bmatrix} \tag{12}$$

Elementtiverkoilla on vastakkaiset synkrooniset pyörimisnopeudet. Kun globaalit matriisit ja rajoiteyhtälöt sijoitetaan yhtälön (10) saadaan liikeyhtälöt muotoon

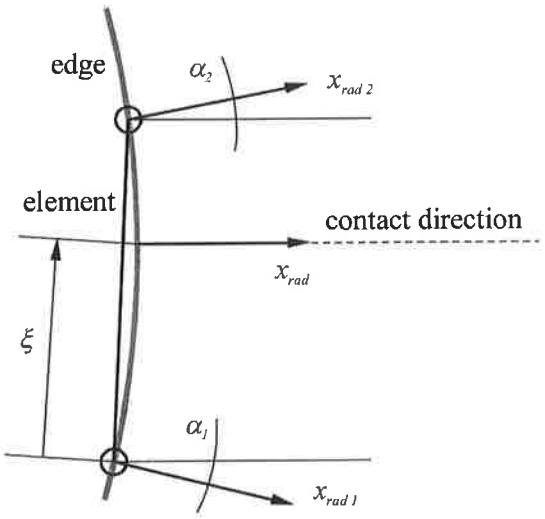
$$\mathbf{H}^T\mathbf{M}_{global}\mathbf{H}\ddot{\mathbf{x}}_{RLconst} + \mathbf{H}^T\mathbf{C}_{global}\mathbf{H}\dot{\mathbf{x}}_{RLconst} + \mathbf{H}^T\mathbf{K}_{global}\mathbf{H}\mathbf{x}_{RLconst} = \mathbf{H}^T(\mathbf{f}_c + \mathbf{K}_{global}\mathbf{x}_G) \tag{13}$$

Yhtälö (13) pitää päivittää joka aika-askeleen alussa koska kontaktipiste liikkuu elementtiverkon reunaa pitkin ja rajoiteyhtälö muuttuu. Esipuristettu kontaktitilanne aikaansaadaan asettamalla kappaleiden alkugeometriat päällekkäin, ja tämä menettelytapa vaatii esianalyysin, jossa ratkaistaan systeemin dynaaminen tasapaino pyörimättömille kappaleille. Esimerkeissä on käytetty lyhyttä simulointia, jossa on käytetty suhteellisen suurta Rayleight-vaimennusta.

Varsinainen rajoiteyhtälö kahden kappaleen välillä on muotoa (vrt. kuva 3)

$$x_{rad} = N_1 x_{rad\,1} + N_2 x_{rad\,2} \tag{14}$$

missä x_{rad} edustaa vastakkaisen kappaleen kontaktipisteen asemaa, ja N_1 ja N_2 ovat reunan interpolointifunktiot. Kun yhtälö (14) ilmaistaan solmuasemien avulla, se saadaan muotoon



Kuva 3. Kontaktipisteen interpolointi.

$$x_{rad} = N_1(\cos\alpha_1\, x_1 + \sin\alpha_1\, y_1) + N_2(\cos\alpha_2\, x_2 + \sin\alpha_2\, y_2) \tag{15}$$

Rajoiteyhtälö voidaan nyt kirjoittaa matriisiin **H**. Kulmat α_1 ja α_2 saadaan laskettua alkugeometriasta. Rajoiteyhtälö (15) edellyttää koordinaatin ξ määrittämistä rotaatioliikkeen aikana ja tämä toteutetaan laskemalla kontaktipisteen sijainti jäykän kappaleen rotaation perusteella suhteellisen suoraviivaisella kirjanpitoprosessilla.

4. AIKATASON SIMULOINTI

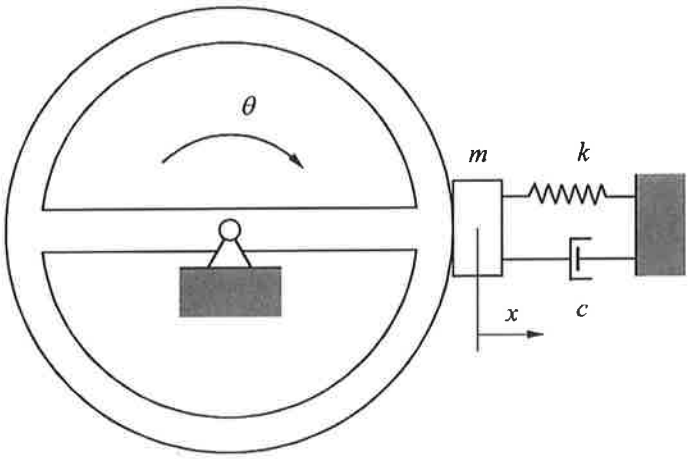
Aikatason ratkaisu on toteutettu käyttämällä epälineaarista Newmarkin aikaintegrointimenetelmää [6]. Aikaintegrointimenetelmä sisältää Newton-Raphson iterointisilmukan ja on siis sopiva myös epälineaarisille tapauksille. Menetelmän käyttämä iterointimatriisi on muotoa

$$\mathbf{S} = \mathbf{K} + \frac{\gamma}{\beta h}\mathbf{C} + \frac{1}{\beta h^2}\mathbf{M} \tag{16}$$

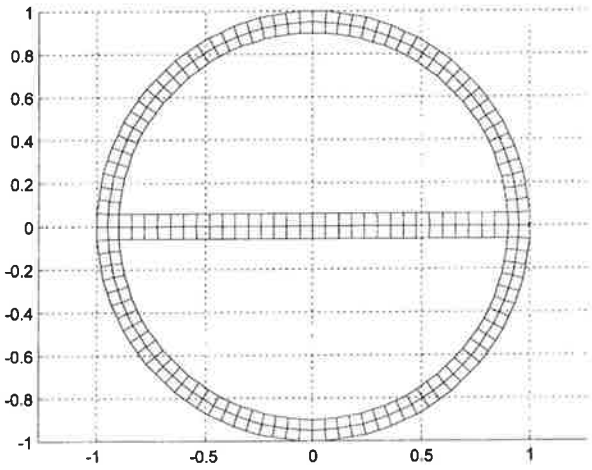
missä β ja γ ovat vakioita ja h on aika-askel. Tämä matriisi **S** pitää päivittää joka aika-askeleen alussa, koska rajoiteyhtälö muuttuu. Laskenta toteutettiin MATLAB ympäristössä käyttämällä hyväksi sen sisältämää harvamatriisiominaisuutta (sparse matrix).

5. TESTIESIMERKKI

Testiesimerkiksi valittiin elastinen tasopyörä, jossa on poikkipuola. Vastinkappaleeksi valittiin yhden vapausasteen värähtelijä ja rakenne on esitetty kuvassa 4. Rakenteen ele-



Kuva 4. Testiesimerkki.



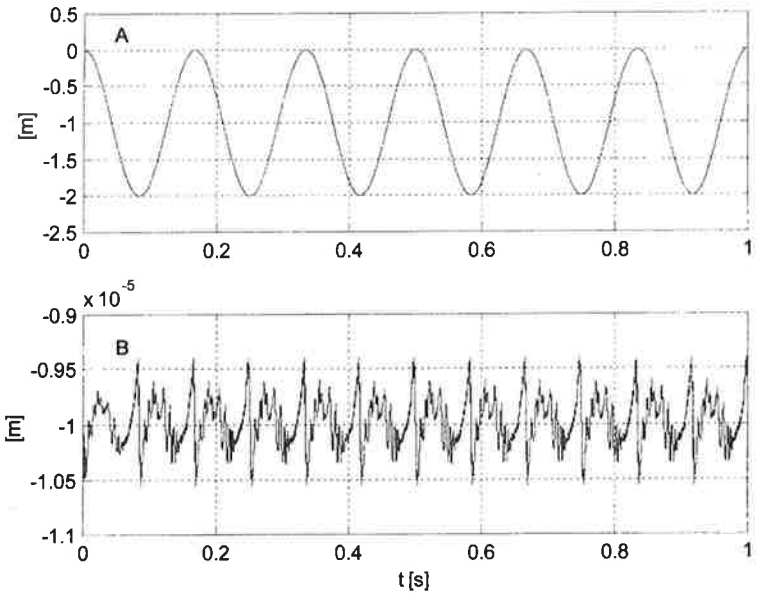
Kuva 5. Testiesimerkin elementtiverkko.

menttiverkko on esitetty kuvassa 5. Elementtityyppinä käytettiin nelisolmuista tasovenymäelementtiä. Pyörän ulkosäde on 1 m ja sen paksuus 0.01 m. Kimmomoduli 20e9 N/m^2, Poissoin luku 0.3 ja tiheys 1000 kg/m^3. Yhden vapausasteen systeemin tiedot ovat m = 10 kg, k = 1e10 N/m ja c = 1e6 kg/s. Aluksi suoritettiin siis esianalyysi dynaamisen lähtöaseman saavuttamiseksi. Rakenteeseen lisättiin Rayleight-vaimennusta vaimennusmatriisilla

$$\mathbf{D} = \alpha\mathbf{K} \tag{17}$$

Tätä vaimennusta käytettiin myös varsinaisessa simuloinnissa. Pyörimisnopeudeksi asetettiin 6 Hz ja alkuasema on kuvan 5 mukainen.

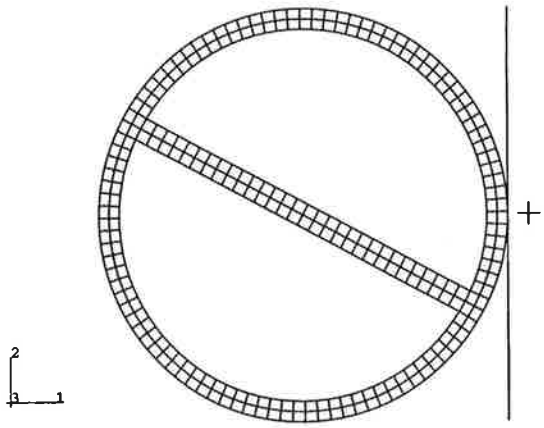
Ensimmäisessä analyysissä kontaktipisteeseen annettiin esipuristusta 1e-5 m verran. Analyysin kestoaika Pentium II 400 MHz PC-koneella oli 38 min. Pyörän rotaatioliike ja kontaktipisteen värähtely on esitetty kuvassa 6. Kontaktipisteen värähtelykäyttäytyminen vaikuttaisi rekisteröivän ortotrooppisen poikkipuolan periodisen iskuvärähtelyn varsin hyvin.



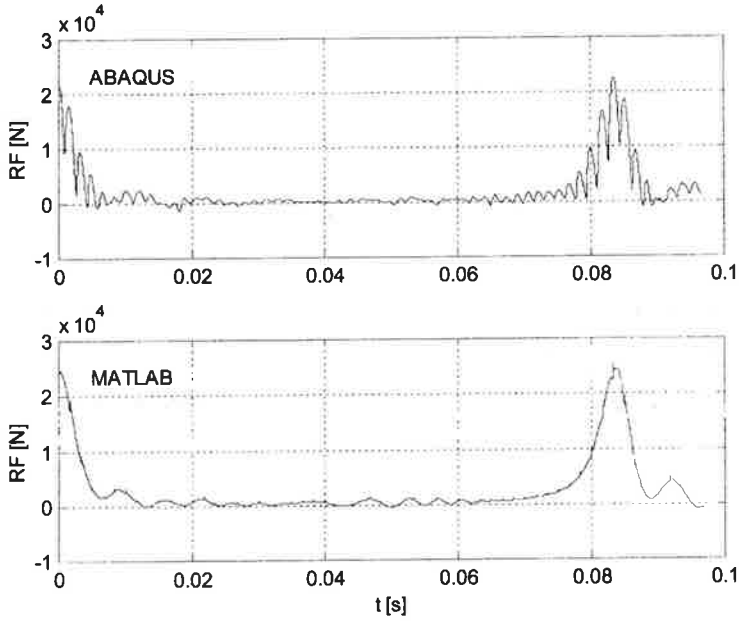
Kuva 6. Rotaatioliikkeen x-siirtymä (A) ja kontaktipisteen värähtely radiaalisuunnassa (B).

Ensimmäisen analyysin perusteella on vaikea sanoa kuinka hyviä tuloksia malli antaa. Tämän vuoksi suoritettiin toinen analyysi, jota verrattiin ABAQUS ohjelman kontaktianalyysiin. ABAQUS mallinnuksessa käytettiin samaa elementtiverkkoa, mutta vastinkappaleena ei ole jousi-massa -systeemi, vaan täysin jäykkä pinta (rigid surface), joka ei värähtele (vrt. kuva 7). Aluksi jäykkä pinta ajettiin kiinni elementtiverkkoon epälineaarisen statiikan ajona ja esipuristumaksi asetettiin 1e-3. Varsinaisessa epälineaarisessa transienttidynamiikan ajossa elementtiverkolle annettiin vakio alkupyörimisnopeus ja laskentaa suoritettiin noin reilun puolen kierroksen verran. MATLAB mallissa puristuma asetettiin samaksi ja massa m ja vaimennus c asetettiin nolliksi.

Analyysien kestoajat olivat ABAQUS-ajo 136 min ja MATLAB-ajo 40 min samalla aika-askeleella h = 1e-5 s. Vertailu ei tässäkään tapauksessa ole kovin tarkka, koska tehtävien reunaehdot ovat täysin erilaiset, mutta tulokset ovat todennäköisesti kohtuullisen suuntaa-antavia. Kuvassa 8 on esitetty molemmilla analyyseillä laskettu kontaktipisteen tukireaktio.



Kuva 7. ABAQUS elementtimalli ja rigid surface.



Kuva 8. ABAQUS- ja MATLAB-analyysien tulosten vertailu. Kuvan suure on kontaktipisteen kontaktisuunnan tukireaktio.

Käyrien yleinen muoto näyttäisi olevan suhteellisesti ottaen sama eli poikkipuola näyttäisi antavan samaa suuruusluokkaa olevan periodisen iskun. Erona tulosten välillä kuitenkin on, että ABAQUS-mallissa esiintyy tämän lisäksi korkeampitaajuista värähtelyä MATLAB-mallin käyttäytyessä huomattavan rauhallisemmin.

6. JOHTOPÄÄTÖKSET

Rajoiteyhtälöllä tapahtuva kahden elastisen kappaleen välinen kontaktimallinnus näyttää varsin lupaavalta lähestymistavalta. Käytetty mallinnustapa ei ole vielä täysin valmis ja se vaatii lisää testausta. Testiesimerkkien antamat tulokset ovat oikean suuntaisia ja mallinnusmenetelmä on varsin kompaktissa muodossa. Aikatason laskenta näyttää varsin nopealta ja se soveltunee pitkiin aikatason simulointeihin. Ainoa hidas osa on iteraatiomatriisin **S** päivittäminen joka aika-askeleen alussa, mutta mahdollisuus käyttää harvamatriiseja helpottaa tilannetta. Tämän paperin testiesimerkit eivät varsinaisessa mielessä ole kiinnostavia, ja tätä rajoiteyhtälöformulointia pitää testata tulevaisuudessa kahden yhteenpuristetun telan 3D-mallilla.

LÄHDELUETTELO

[1] **Ahmed A. Shabana**, *Dynamics of Multibody Systems*, Cambridge University Press, 1998

[2] **Crisfield M. A.** and **Moita G. F.**, *A co-rotational formulation for 2-D continua including incompatible modes*, International journal for numerical methods in engineering, Vol. 39, pp. 2619-2633, 1996

[3] **Crisfield M. A.**, *Non-linear Finite Element Analysis of Solid and Structures Volume 2*, John Wiley & Sons Ltd., 1997

[4] **V.-M. Järvenpää** and **Erno K. Keskinen**, *Finite element model for rotating paper machine roll*, IMAC-XVIII, San Antonio, 2000

[5] **O. C. Zienkiewicz** and **R. L. Taylor**, *The Finite Element Method Volume 1*, McGraw-Hill International (UK) Limited, 1997

[6] **Géradin M.** and **Rixen D.**, *Mechanical vibrations*, John Wiley & Sons Ltd., 1997