Contents of issue 3 vol. LV

Engineering Transactions, T. **LV**, nr 3, str. 195–278, Warszawa 2007

ENGINEERING TRANSACTIONS • *Engng. Trans.* • **55**, 3, 197–216, 2007
Polish Academy of Sciences • Institute of Fundamental Technological Research

FINITE ELEMENT ANALYSIS OF STRAIN NON-UNIFORMITY IN TWO PROCESSES OF SEVERE PLASTIC DEFORMATION

J. M a c i e j e w s k i, H. K o p e ć, H. P e t r y k

Institute of Fundamental Technological Research
Polish Academy of Sciences
Warsaw, Poland

Two severe plastic deformation (SPD) processing techniques, namely equal-channel angular pressing (ECAP) and cyclic extrusion-compression (CEC), are investigated by using the finite element method. The major aspect examined is the non-uniformity of the accumulated, equivalent plastic strain after processing with the use of different shapes of the die. The quantitative effect of several parameters on the plastic flow is determined. It is found that the diameter ratio of the chambers and narrower channel in the CEC method, and also the inclination angle of connecting conical parts, can affect strongly the degree of strain non-uniformity. Comparison is made of distributions of equivalent strain after two passes of ECAP for two different routes and with two die profiles.

1. Introduction

In the last two decades many studies have been devoted to the material processing by applying severe plastic deformation (SPD). The principal aim of imposing extremely large plastic strains is to achieve ultra-fine grain sizes. Decrease of grain size to a sub-micrometer level is related to beneficial mechanical properties such as very high strength which may also be accompanied by low-temperature superplasticity. In contradistinction to traditional cold rolling or extrusion, the purpose of SPD is not to reduce a cross-section dimension to a given value. Several SPD processes have been designed so that the billet shape remains essentially unchanged after the processing by cyclic deformation. Two such processes are investigated here, namely, equal-channel angular pressing (ECAP) and cyclic extrusion-compression (CEC). ECAP is the most popular SPD technique, introduced in the former Soviet Union [1]. An extensive review of the vast literature and recent developments related to the use of ECAP for grain refinement can be found in Ref. [2]. CEC is the SPD technique developed in Poland [3] and subsequently studied in a series of papers, e.g. [4–7].

Majority of papers dealing with SPD techniques have concentrated on microstructural aspects of the grain refinement and on the changes in mechanical prop-

erties of the material. Modelling of the material flow was also performed by using an analytic approach [8–10], the slip-line field theory [11, 12], an upper-bound approach [13, 14], and the finite element method (FEM) [15–23], predominantly for ECAP. The present authors are aware of the only one finite element study of the CEC process [17].

The aim of this paper is to present selected results of FEM simulations of the above-mentioned two SPD processes. The axisymmetric FEM simulations of CEC and plane-strain simulations of ECAP have been performed by using the commercial computer code ADINA [24]. The modelling of CEC constitutes the major part of this work and provides new results. ECAP was thoroughly studied in the literature, nevertheless the present analysis of two passes for two different ECAP routes and two die profiles and for different friction conditions, appears to be novel. The major aspect examined below is the non-uniformity of the accumulated, equivalent plastic strain after processing with the use of different shapes of the die. The predictions of numerical simulations can help in optimal design of the tools for technological applications.

2. Simulations of cyclic extrusion-compression (CEC)

2.1. Description

The cyclic extrusion-compression method can be applied to achieve unlimited accumulated strains by a combination of the processes of extrusion and compression, with preservation of the initial shape of a material sample [3–7]. This is an important advantage of the CEC method. The sample (Fig. 1) is placed in the die consisting of left and right chambers of equal diameter d_0, connected by a channel of a smaller diameter d_m. The deformation proceeds by the cyclic flow of metal from one chamber to the other. For example, in a single cycle the sample is extruded from the left chamber, say, by the left active punch, with the backpressure exerted by the right passive punch which causes plastic compression of the material just after extrusion. In the next cycle, the right punch becomes active and the deformation proceeds by flow of the sample from the

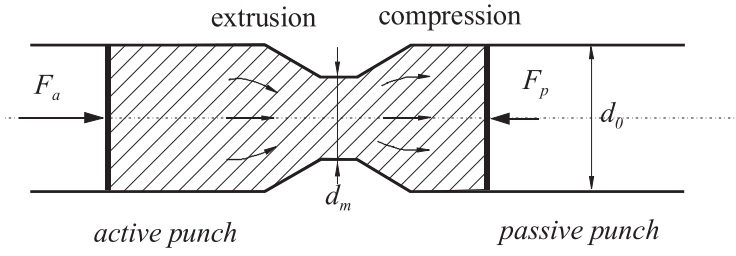

Fig. 1. Sketch of CEC deformation process.

right chamber to the left, and so on. In this way the metal sample is deformed cyclically to unlimited strains without removing the material sample from the apparatus. If the deformation is assumed to be uniform, then the magnitude of the accumulated equivalent (von Mises) strain after $n$deformation cycles can be estimated as follows [3]:

$$\varepsilon = 4n \ln\left(\frac{d_0}{d_m}\right). \tag{2.1}$$

The FEM simulations of the CEC process were performed using the ADINA 8.3 System, assuming axisymmetric deformation and large strain and large displacement kinematics under isothermal and quasistatic conditions. A finite element mesh covering the deformed material consisted of 2240 axisymmetric 9-node solid elements. The die and the punches were modelled as rigid bodies. Figure 2 presents the scheme of the FE mesh and of the die shape. In the numerical analysis the die shape was described by parameters: d_0, d_m, α inclination angle of the conical parts) and a (length of the central cylindrical channel). The transition from the cylindrical to conical parts of the die was modelled as a curve of radius $R = 1/2\ (d_0 - d_m)$.

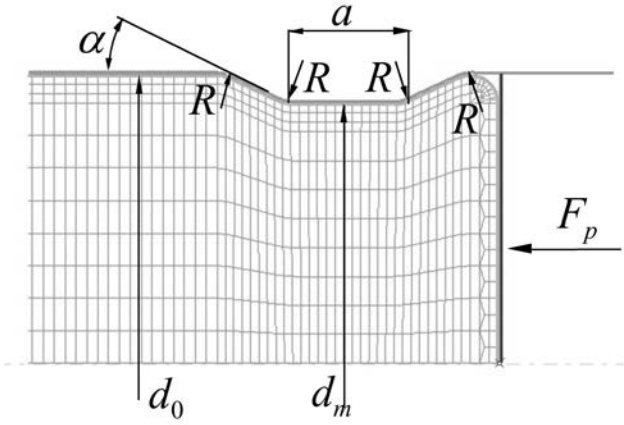

FIG. 2. Scheme of the finite element mesh and CEC channel geometry.

In the calculations the material was modelled as elastic-plastic with isotropic hardening described by a multilinear approximation of a true (logarithmic) strain – true (Cauchy) stress curve shown in Fig. 3. The experimental data points that correspond to polycrystalline pure aluminium (99.99%) deformed at room temperature (293 K) in a wide range of strain are taken from Refs. [25]. The experimental data for low strain level were obtained in the tension test, whereas for larger strains the data were reported for samples produced in the ECAP process [26, 27]. The elasticity parameters were taken as Young's modulus $E =$ 70 GPa and Poisson's ratio $\nu = 0.32$.

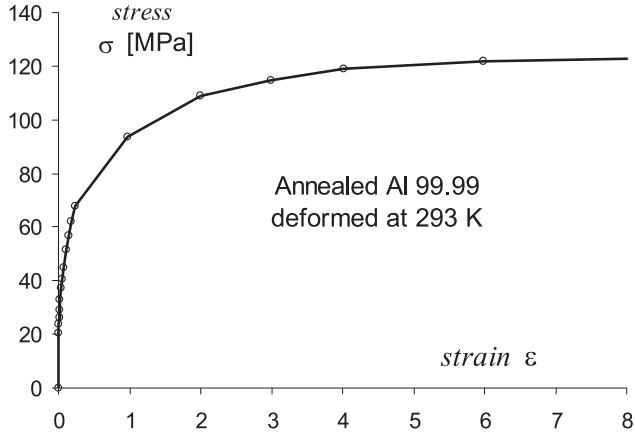

FIG. 3. Experimental data points [25–27] and a multilinear approximation of the true stress–true strain curve for a wide range of strain for pure Al deformed at room temperature (293 K).

In the real CEC process the samples were covered with a graphite lubricant before the deformation process [5]. In the calculation, the frictionless contact as well as the nonlinear friction model were used. The nonlinear friction law was assumed in the form

$$\tau_n = A_1 \left(1 - \exp(-\sigma_n A_2)\right), \tag{2.2}$$

where τ_n, σ_n are the shear and normal contact stress components, respectively, and A_1, A_2 are the friction parameters whose meaning is indicated in Fig. 4.

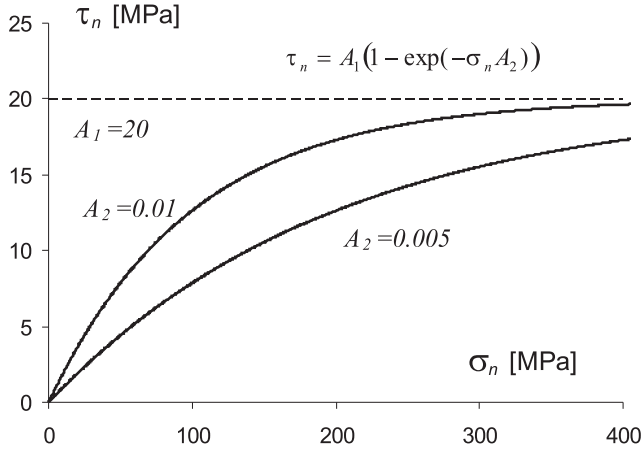

FIG. 4. The nonlinear friction law used in numerical simulations.

Two different sequences of calculation of the CEC process have been used (Fig. 5). In the first sequence illustrated in Fig. 5a, the material was initially in the left chamber only. As the left punch was kinematically forced to move, the forward extrusion (without the passive punch) was executed. Next the pressure was applied to the right punch, aimed at complete filling of the die space, while the active left punch was still moving. As a result, a roughly uniform deformation zone around the sample axis in the central part was obtained. In the end zone of the sample the strain distribution was non-uniform. In the second sequence of calculation, the initial material position was as shown in Fig. 5b. The right-hand side of the sample was positioned at once beyond the narrower channel, the hardening effect of compression by the right-hand punch being neglected. Then, along with the active left punch motion, the backpressure was exerted by the right punch. After completing the first cycle of deformation, the resulting strain distributions were found to be of steady-state type and practically the same for both calculation sequences, except in the vicinity of the right end of the sample. Therefore, in further calculations only the second sequence was used as a simpler approach.

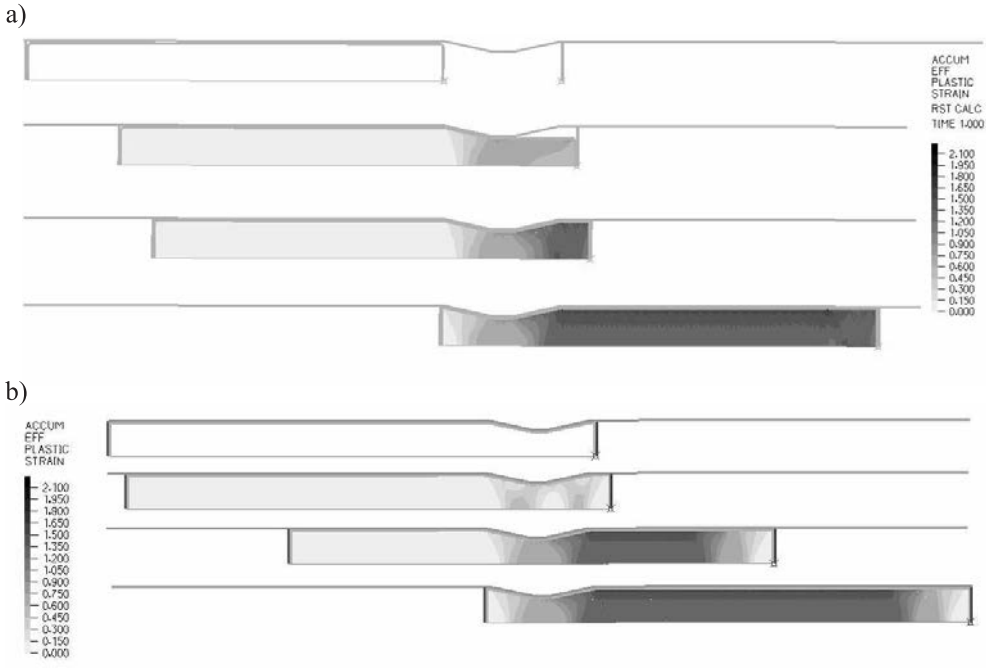

FIG. 5. Equivalent plastic strain field in the subsequent phases of the first pass in CEC process for two different sequences of calculation: a) initial position of the sample in the left chamber only, b) initial position of the sample covers the central channel.

2.2. Results and discussion

Figure 6 shows the distribution of accumulated equivalent plastic strain in the transverse cross-section of the sample after the first cycle of the CEC process for different diameter ratios d_0/d_m = 15/11, 10/8, 10/9, 10/9.5. The results were taken from for the major (central) part of the sample where the strain was uniform along the axis. In this calculation, a frictionless contact was assumed and the additional shape parameters were: inclination angle of conical parts $\alpha = 25°$ and channel length $a = 2$ mm. The dashed lines in the figure represent the theoretical magnitude of accumulated equivalent plastic strain after one cycle which, according to Eq. (2.1), is equal to 1.24, 0.982, 0.421, 0.20 for diameter ratio d_0/d_m =15/11, 10/8, 10/9, 10/9.5, respectively. It is clearly seen that the deformation is non-uniform with the peak near the outer radius of the sample, with a strong influence of the diameter ratio on the strain non-uniformity. For the dies with diameter ratio 10/9 and 10/9.5, the accumulated equivalent strain near the sample axis is less than the theoretical value. The strain distribution is fairly uniform for $d_0/d_m = 15/11$.

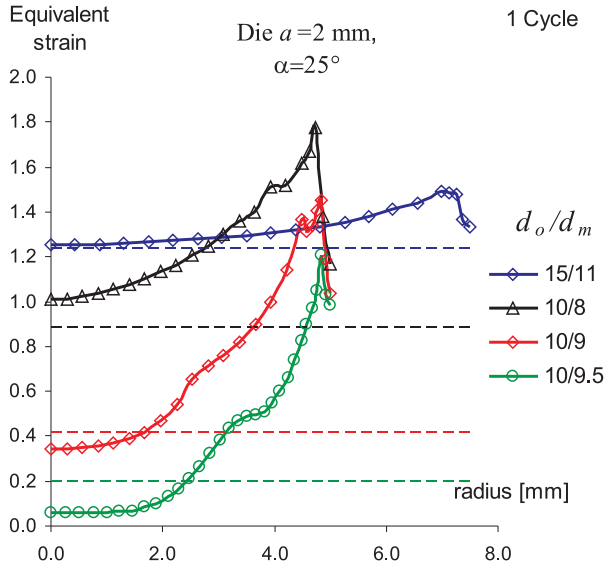

FIG. 6. Accumulated equivalent plastic strain versus the sample radius after one cycle of the CEC process, with different diameter ratios d_0/d_m = 15/11, 10/8, 10/9, 10/9.5. Channel length $a = 2$ mm, inclination angle of conical parts $\alpha = 25°$, frictionless conditions.

The angle of inclination of conical parts of the die has also a significant influence on the resulting homogeneity of the sample. The variation of accumulated equivalent plastic strain with respect to the sample radius for different inclina-

tion angles $\alpha = 15°, 25°, 35°, 45°$ is presented in Fig. 7. These calculations were performed for diameter ratio $d_0/d_m = 10/8$ and channel length $a = 2$ mm. With the angle growth, the non-homogeneity increases.

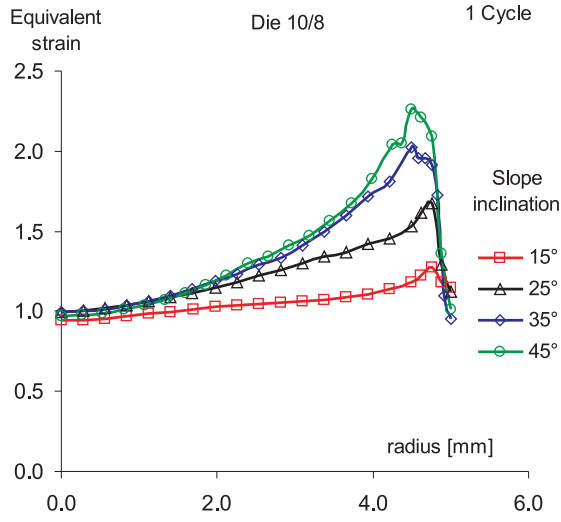

FIG. 7. Accumulated equivalent plastic strain versus the sample radius after one cycle of the CEC process, with $d_0/d_m = 10/8$ and channel length $a = 2$ mm, for different inclination angles $\alpha = 15°, 25°, 35°, 45°$ of conical parts of the die.

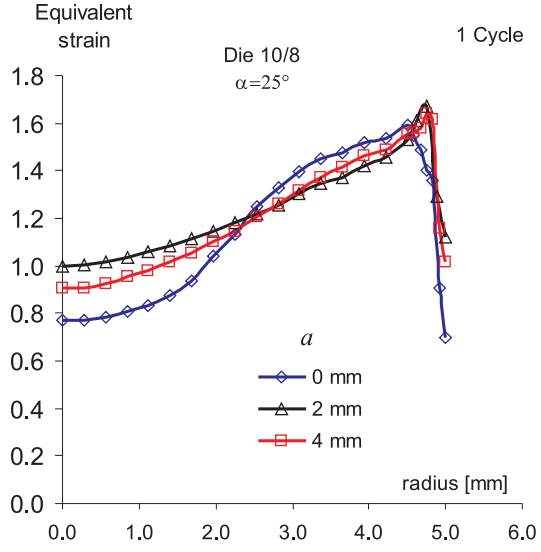

FIG. 8. Accumulated equivalent plastic strain versus the sample radius after 1 cycle of the CEC process, with $d_0/d_m = 10/8$ and $\alpha = 25°$, for different channel lengths $a = 0$, 2, 4 mm.

The influence of the length a of the cylindrical channel of the die is less substantial as that shown in Fig. 8. The non-homogeneity is the largest for channel length $a = 0$ when the conical parts of the die are in contact.

Figure 9 shows the effect of friction on the material/tool contact surface, on the radial distribution of accumulated equivalent plastic strain. In the case of diameter ratio 10/9 and inclination angle of conical parts $\alpha = 25°$, the influence of friction was significant only near the outer surface being in contact with the tool, whereas in the remaining part the deformation was only slightly affected by friction.

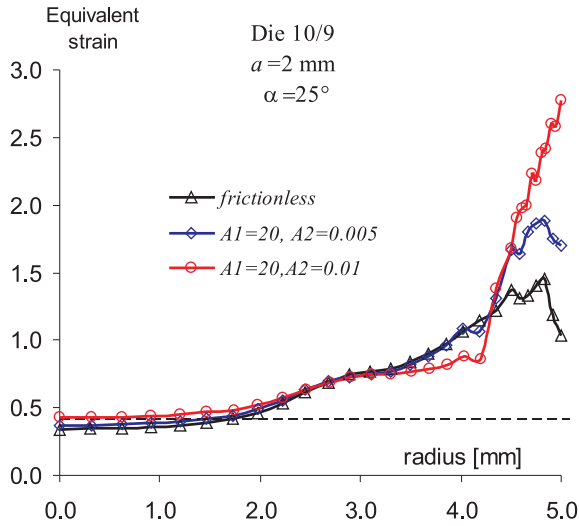

FIG. 9. Accumulated equivalent plastic strain distribution for different friction parameters.

The next three figures show perhaps the most important effect of the ratio of the chamber and channel diameters on the plastic strain non-uniformity. Each figure corresponds to the same values of channel length and inclination angle of conical parts, taken in this case as $a = 2$ mm and $\alpha = 25°$, respectively, while the diameter ratio is the only variable parameter. Frictionless contact has been assumed. In Fig. 10 the radial distribution of the accumulated equivalent plastic strain is shown after the first, second and third cycle of CEC, for diameter ratio 10/8. The distribution is clearly not uniform, but the degree of non-uniformity (ratio of maximum to minimum equivalent strains) equal to 1.7 does not increase after subsequent cycles. Therefore, in the range of severe plastic deformation where the strain hardening saturates, that strain non-uniformity does not need to lead to substantial macroscopic heterogeneity in the mechanical properties such as the yield stress or microhardness. If fact, the microhardness distributions

along the sample diameter, after 4 and 22 cycles of CEC with the die of diameter, ratio 10/8, were found to be fairly uniform and close to each other [3].

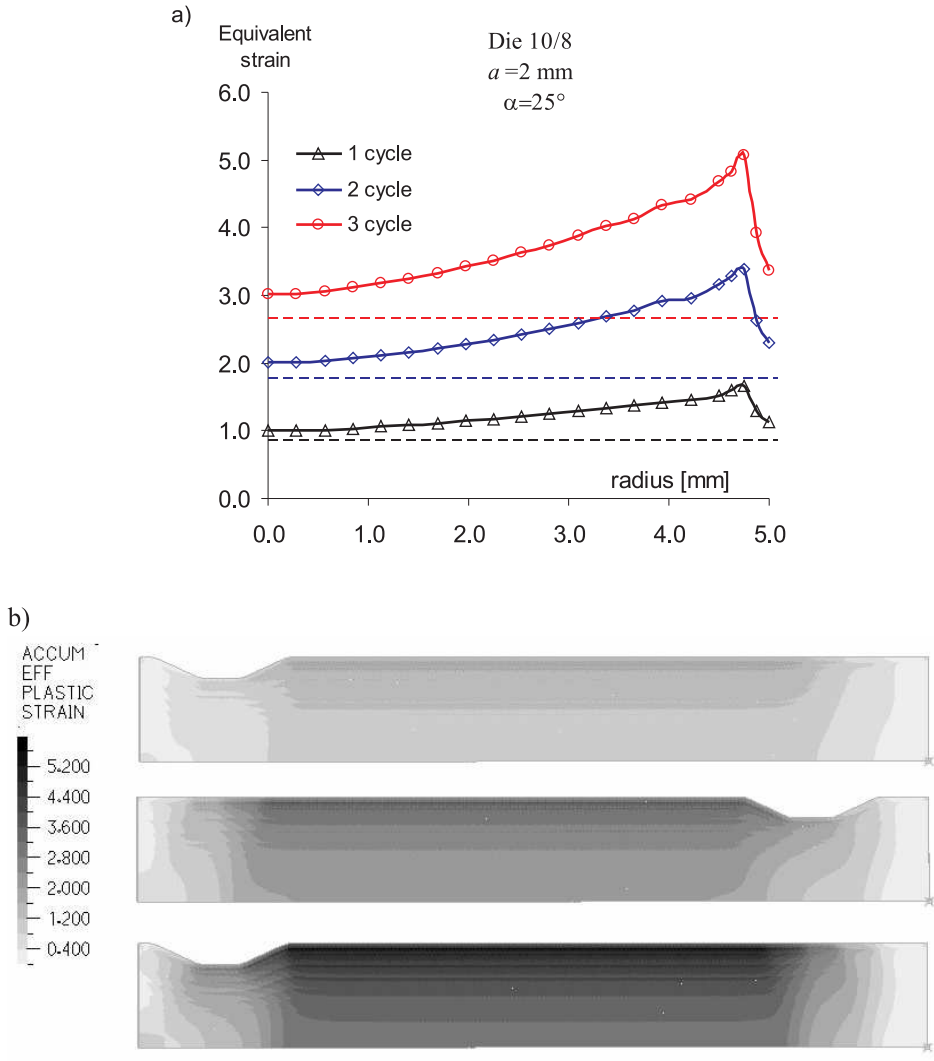

FIG. 10. Accumulated equivalent plastic strain after 1, 2 and 3 cycles of the CEC process, with $d_0/d_m = 10/8$, $a = 2$ mm and $\alpha = 25°$; a) distribution versus the sample radius, b) 2D distribution in the longitudinal cross-section.

The situation changes when the difference in diameter of the chambers and the connecting channel is too small. As shown in Figs. 11 and 12, in that case the degree of non-uniformity is large and increases with subsequent cycles to 9.7 and 40, for the diameter ratio d_0/d_m= 10/9 and 10/9.5, respectively. The

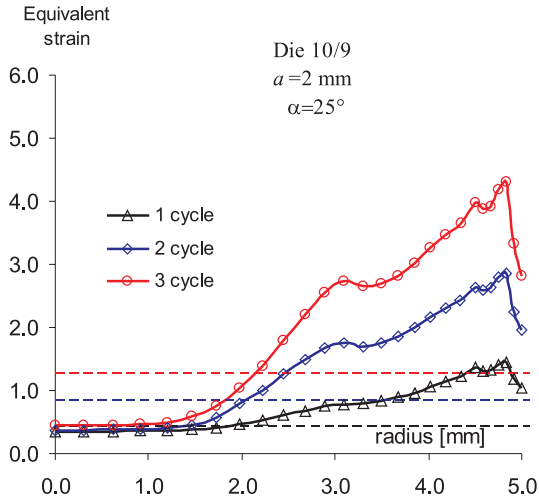

FIG. 11. Accumulated equivalent plastic strain versus the sample radius after 1, 2 and 3 cycles of the CEC process, with d_0/d_m =10/9.

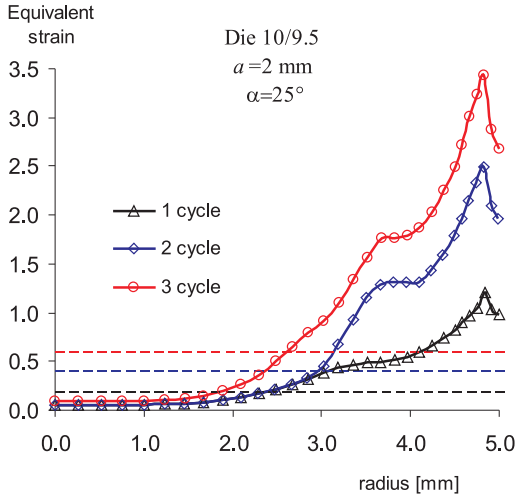

FIG. 12. Accumulated equivalent plastic strain versus the sample radius after 1, 2 and 3 cycles of CEC, with d_0/d_m =10/9.5.

equivalent strain in the inner part of the sample, deformed when using the die of diameter ratio d_0/d_m =10/9 is accumulated much slower than that predicted by the formula (2.1) for uniform distribution of strain, while in the outer part it becomes much larger. If the diameter ratio d_0/d_m is smaller, then the plastic strain accumulation in the inner part may even be suspended, cf. Fig. 12.

It may be noted that the above conclusion has been drawn from the numerical results obtained by using the conventional plasticity model, which does not describe the formation of macroscopic shear bands. Such bands running across the sample were observed experimentally [4, 6]. Presence of macroscopic shear bands, while introducing local flow non-uniformities, might nevertheless reduce the strain non-uniformity on the scale of the sample radius. This question requires further study which is, however, beyond the scope of the present paper.

3. Simulations of equal-channel angular pressing (ECAP)

3.1. Description

The ECAP process was analyzed much more frequently in the literature than CEC, therefore some results for ECAP are presented here to provide a comparison of the flow non-uniformity in both processes. As it is well known [2], the nature of the deformation imposed by ECAP is simple shear which occurs as the sample passes through an abruptly bent channel. The square cross-sectional dimensions do not experience any changes, therefore the sample may be pressed repeatedly to achieve very high accumulated strains. Different deformation routes are possible here. As illustrated in Fig. 13, the billet can be placed again in the die without any rotation about its axis (route A), or be rotated by 180° after each pass (route C). This is equivalent to pressing the sample through an U-shaped or S-shaped channel, respectively; this analogy has been used in the present 2D calculations. Other routes (B_A – the billet is rotated clockwise through 90° between each cycle, B_C – the billet is rotated through 90° clockwise

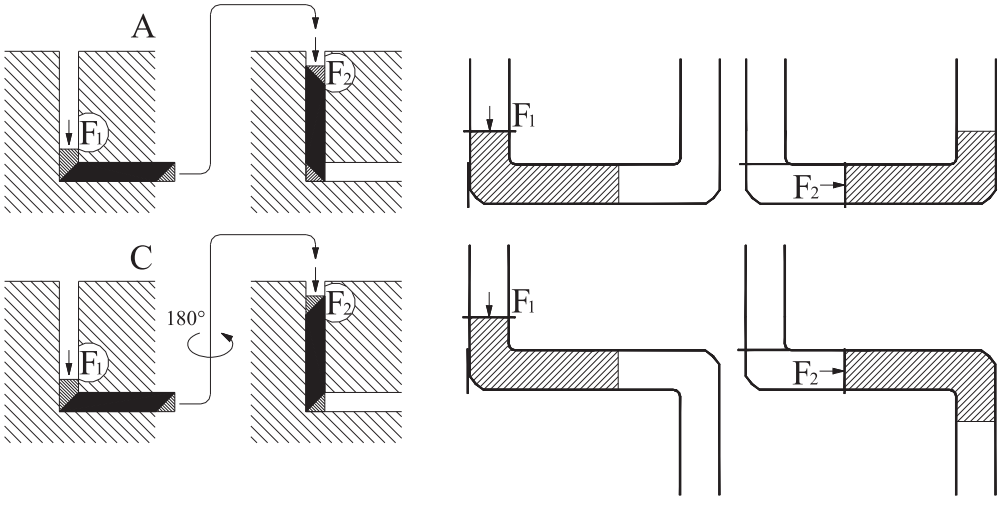

Fig. 13. Sketch of the examined routes of ECAP.

and then 90° anticlockwise each alternate cycle) are not examined here since they would require a more complex 3D analysis.

Details of die geometry have a certain influence of the material flow through the die. In this paper, two types of die geometry have been investigated, as shown in Fig. 14. Such dies with the angle of intersection of the channels equal to $\phi = 90°$ and with the angle of the outer arc of curvature where the two channels intersect, equal either to $\psi = 20°$ or $\psi = 0°$, are typical in applications [2]. The radius r of rounded-off corners is introduced here mainly to overcome numerical problems related to simulation of plastic flow around sharp corners. The finite element mesh consisted of 5472 plane-strain 9-node elements; its part is shown in Fig. 15.

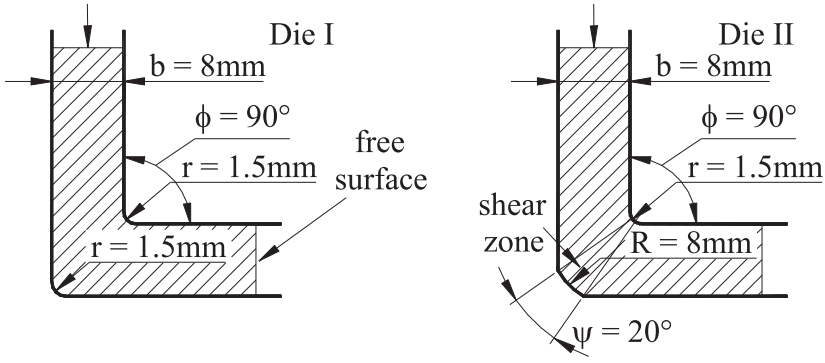

FIG. 14. Geometry of two ECAP dies used in numerical simulations.

The material model used in calculations corresponds to pure aluminium as in CEC calculations, however, with two modifications. First, hardening was described here in a more simple way by assuming linear hardening from the initial yield stress $\sigma_0 = 95$ [MPa] with the tangent modulus $E_T = 15$ [MPa]. Either pure isotropic or pure kinematic hardening were examined.

3.2. FEM results and discussion

Figure 15 shows the distribution of the accumulated equivalent plastic strain in the longitudinal cross-section of the sample in the vicinity of the basic shear zone, together with the distribution of contact surface tractions, for four values $\mu = 0.01$, 0.1, 0.2, 0.3 of the assumed Coulomb friction coefficient. It has been found that while the value of μ does not influence the strain distribution except in the surface layer, it affects the length of the outer corner gap identified as a zone where the surface tractions are zero.

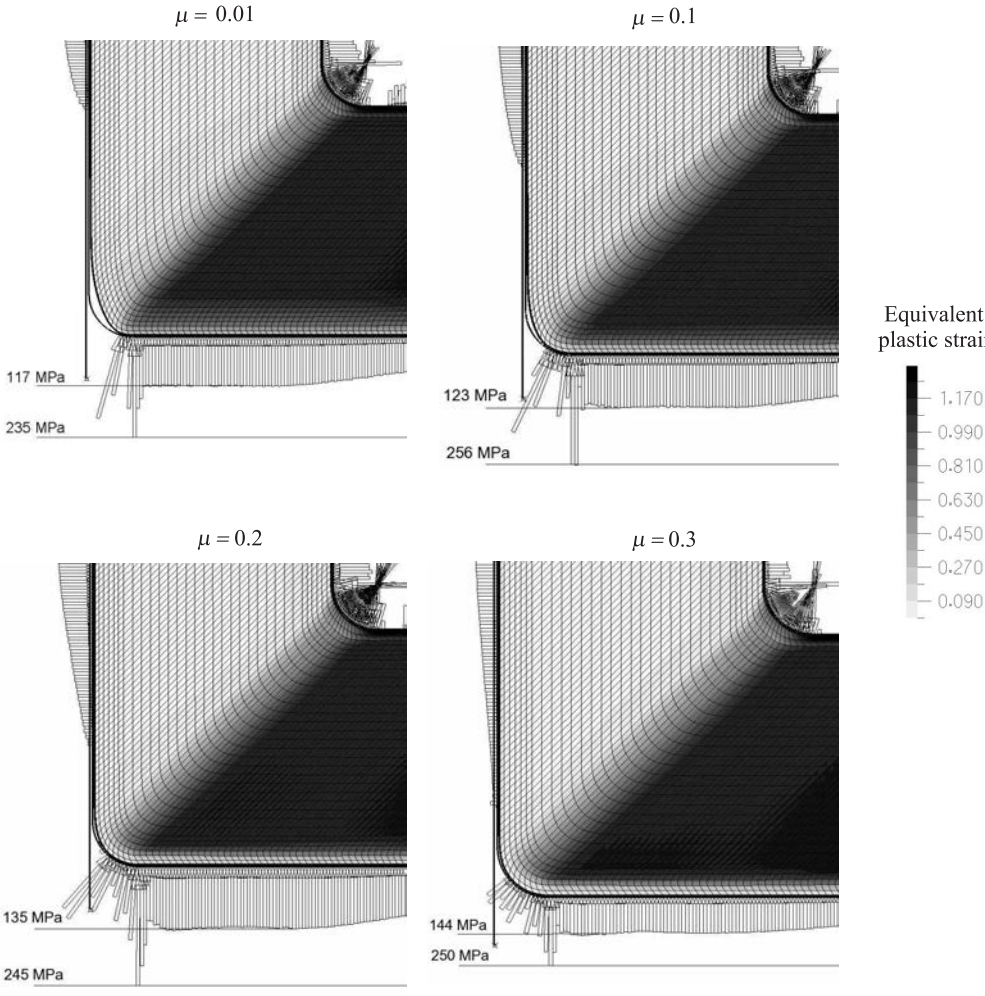

FIG. 15. Distribution of equivalent plastic strain and of surface tractions during the first pass through Die I, for different values $\mu = 0.01$, 0.1, 0.2, 0.3 of the Coulomb friction coefficient. The development of the corner gap is dependent on the friction conditions.

In Fig. 16a the force acting on the punch during the first and second pass, using Die I for route C, is plotted versus the punch displacement for different values of the Coulomb friction coefficient. In Fig. 16b the respective distributions of equivalent plastic strain across the billet are presented. The forming force grows with increasing friction coefficient as expected, whereas the distributions of equivalent plastic strain are very similar for different friction coefficients. Small differences in the sub-surface layers of the billet are only observed.

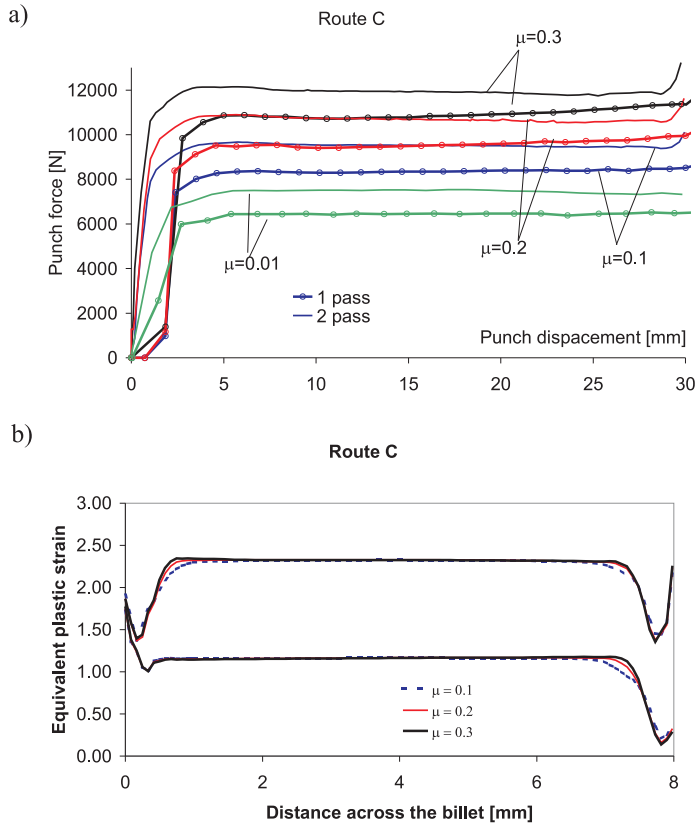

FIG. 16. Simulation results for ECAP Route C using Die I for different values $\mu = 0.1$, 0.2, 0.3 of the Coulomb friction coefficient, a) force acting on the punch versus the punch displacement, b) distribution of equivalent plastic strain across the billet.

In Fig. 17 the longitudinal distribution of equivalent plastic strain after the first pass through the ECAP Dies I and II is shown first, followed by similar pictures after the second pass according to routes A and C. It can be seen that the strain distribution is more non-uniform after the second pass for route A than that for C, which is more clearly visible for Die I. In turn, the distribution of equivalent plastic strain is found more uniform for Die I with $\psi = 0°$ than that for Die II with $\psi = 20°$. During a passage through Die II, a less deformed layer is formed in the lower part of the sample. After a closer look at the deformed mesh, a conclusion can be drawn that this is due to formation of a rotating and less deforming zone, in vicinity of the rounded outer corner of Die II. The obtained results are in qualitative agreement with previous studies [9, 11, 20, 28, 29]. A more detailed plots of the distributions of equivalent plastic strain across the material sample are shown in Fig. 18.

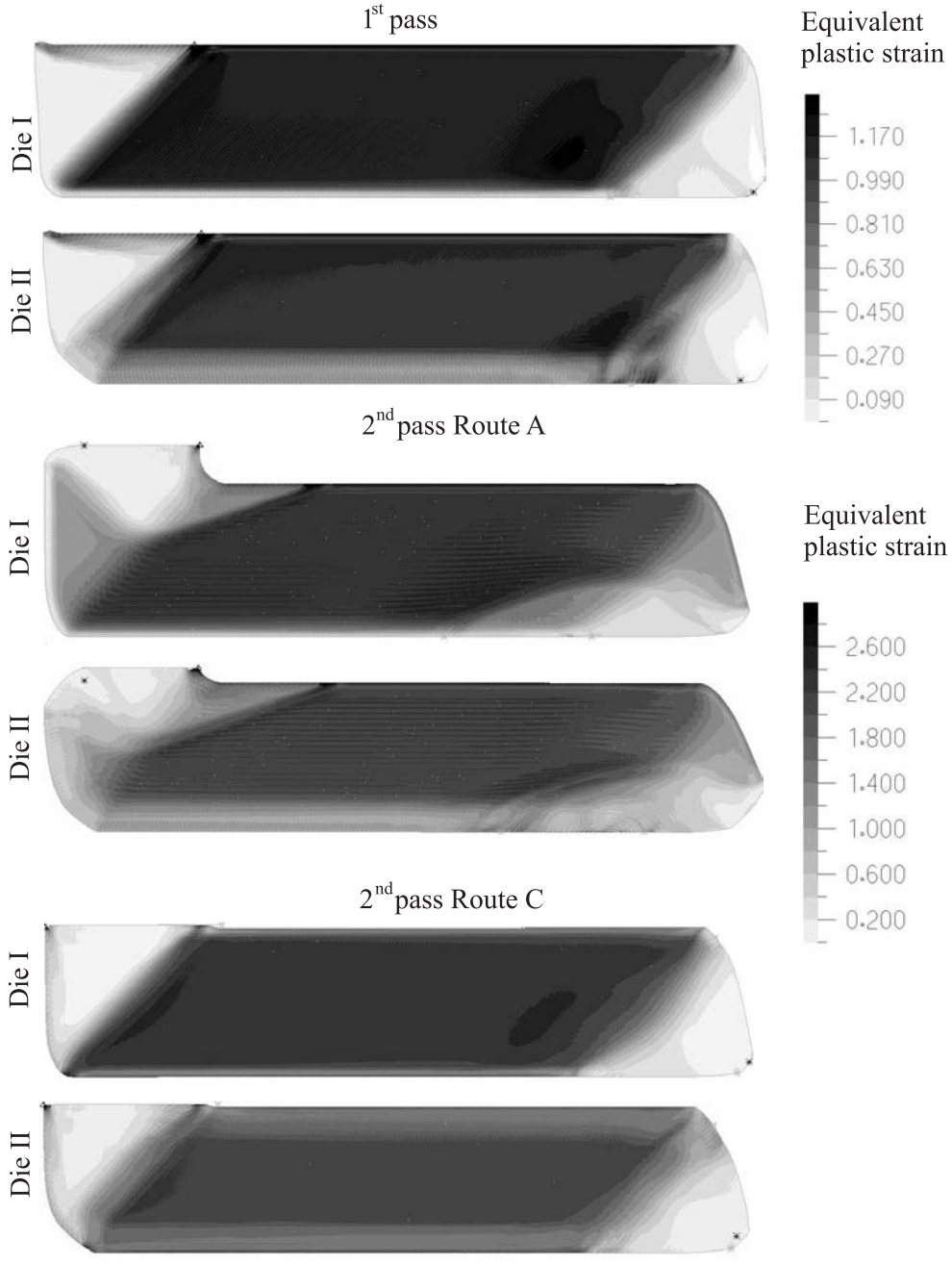

FIG. 17. The fields of equivalent plastic strain during ECAP, after the first and second pass through Die I and II for routes A and C.

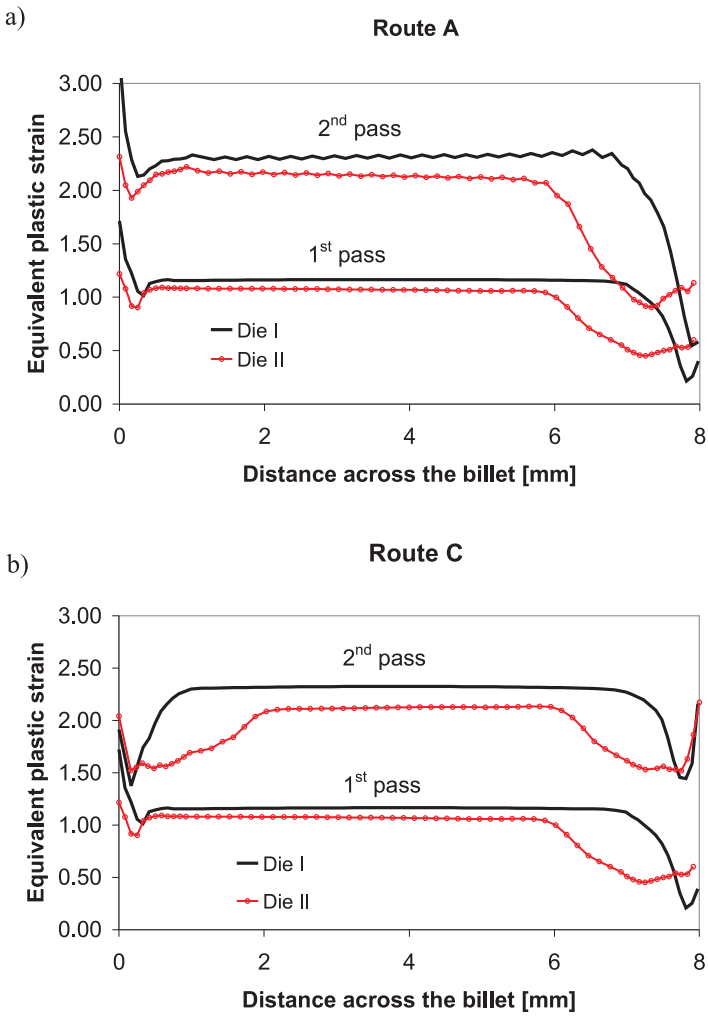

FIG. 18. Distribution of equivalent plastic strain across the billet during ECAP, after the first and second pass through Dies I and II, for routes A and C.

The difference between the deformation fields calculated for isotropic hardening and for kinematic hardening, corresponding to the same stress-strain curve for monotonic loading and for the same ECAP conditions, has not been found to be significant. A more essential and expected difference appears between the plots of the total force acting on the punch during the second pass according to route C, cf. Fig. 19. The reason for the difference is clear as for this route the direction of shear is reversed, which corresponds to the decrease of the shear stress when kinematic rather than isotropic hardening is assumed.

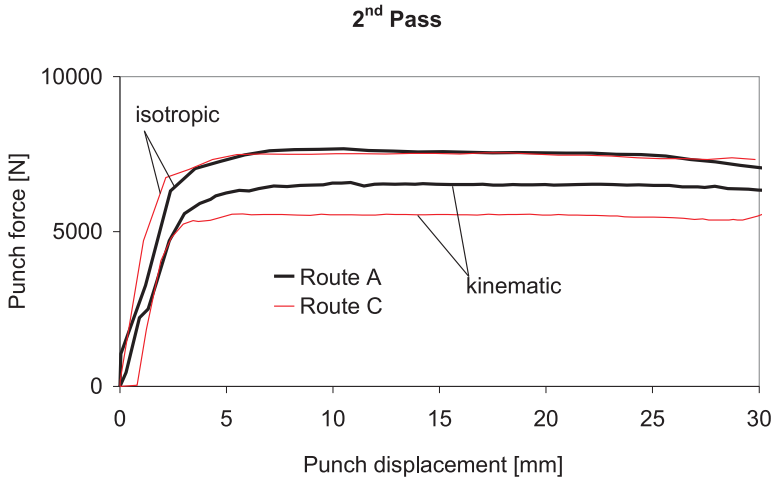

FIG. 19. Force acting on the punch versus the punch displacement during the second pass of ECAP, using Die II with negligible friction ($\mu = 0.01$). The results are for routes A and C and for either isotropic or kinematic hardening.

4. CONCLUSIONS

Finite Element Method simulations of the CEC and ECAP processes have been carried out for the elastoplastic material with strain hardening under isothermal conditions. The model parameters were calibrated for polycrystalline pure aluminium (99.99%) at room temperature.

The CEC process was simulated for various die shape parameters: ratio d_0/d_m (of diameters of the chambers and central cylindrical channel of the die, respectively), α (inclination angle of conical parts that connect chambers with the channel) and a (length of the cylindrical channel). The diameter ratio d_0/d_m is the most important parameter which determines the overall magnitude of strain in a single pass. The simulations performed have shown that the deformation is always non-uniform, with a peak near the outer radius of the sample. However, the degree of non-uniformity, defined as the ratio of maximum to minimum equivalent strain along the sample radius, is found to depend strongly on the value of d_0/d_m.

For dies with diameter ratio 10/9.5 and 10/9 the accumulated equivalent strain near the sample axis is much less than the theoretical value for uniform deformation, and the degree of non-uniformity is large and increases with subsequent cycles. If the diameter ratio d_0/d_m is too small then the plastic strain accumulation in the inner part may even be suspended, as it happened for the die 10/9.5 with the other parameters provided in Fig. 12. With the increase of the

diameter ratio the deformation becomes more uniform, and for diameter ratio 10/8 the degree of non-uniformity equal to 1.7 does not increase after subsequent cycles. Equivalent plastic strain distribution is found to be fairly uniform for somewhat higher diameter ratio (15/11), with the degree of non-uniformity reduced to 1.2. The above conclusions have been drawn using the conventional plasticity model which does not describe formation of macroscopic shear bands observed experimentally.

The angle of inclination of conical parts of the die has also a significant influence on the resulting homogeneity of the sample, and with the angle growth the non-homogeneity increases. By decreasing the inclination angle at a fixed value of d_0/d_m, homogeneity of the deformation can be substantially improved. The influence of length a of the central cylindrical channel of the die is less substantial.

Friction on the material/tool contact surface has a considerable influence on the strain peak near the outer surface being in contact with the tool, whereas in central part the deformation it is found to be only slightly affected by friction.

Finite Element simulations of two passes of the ECAP process were carried out for routes A and C using two die shapes and for different friction conditions.

It is found that the value of the Coulomb friction coefficient μ does not influence the strain distribution except in the sub-surface layers, although the value of μ affects the length of the outer corner gap identified as a zone where the surface tractions are zero.

A uniform region of equivalent plastic strain accumulated in the first and second pass of ECAP is obtained in the central part of the billet. The strain distribution after the second pass is more non-uniform for route A than for route C, especially for Die I. In turn, the distribution of equivalent plastic strain is found to be more uniform for Die I with $\psi = 0^\circ$ than for Die II with $\psi = 20^\circ$. During a passage through Die II, a less deformed layer is formed near the outer corner of the die.

Acknowledgment

This work has been supported by the State Committee for Scientific Research (KBN) in Poland through Grant No. PBZ-KBN-096/T08/2003.

References

1. V.M. Segal, *Materials processing by simple shear*, Mater. Sci. Eng. A, **197**, 157–164, 1995.

2. R.Z. Valiev, T.G. Langdon, *Principles of equal-channel angular pressing as a processing tool for grain refinement*, Progress in Materials Science, **51**, 881–981, 2006.

3. J. Richert, M. Richert, *A new method for unlimited deformation of metals and alloys*, Aluminium, **8**, 604–607, 1986.

4. M. Richert, A. Korbel, *The Effect of strain localization on mechanical properties of Al99,992 in the range of large deformations*, J. Mater. Proc. Technol., **53**, 331–340, 1995.

5. M. Richert, J. Richert, J. Zasadzinski, H. Dybiec, *The boundary strain hardening of aluminium with unlimited cumulation of large deformation*, Z. Metallkd., **79**, 741, 1988.

6. M. Richert, Q. Liu, N. Hansen, *Microstructural evolution over a large strain range in aluminium deformed by cyclic extrusion–compression*, Materials Science and Engineering A, **260**, 275–283, 1999.

7. M. Richert, H.P. Stuwe, M.J. Zehetbauer, J. Richert, R. Pippan, Ch. Motz, E. Schafler, *Work hardening and microstructure of AlMg5 after severe plastic deformation by cyclic extrusion and compression*, Materials Science and Engineering A, **355,** 180–185, 2003.

8. L. Dupuy, E.F. Rauch, *Deformation paths related to equal channel angular extrusion*, Materials Science and Engineering A, **337**, 241–247, 2002.

9. I. J. Beyerlein, C. N. Tome, *Analytical modeling of material flow in equal channel angular extrusion (ECAE)*, Materials Science and Engineering A, **380**, 171–190, 2004.

10. C. Liu, D.J. Alexander, *A kinematic analysis of the angular extrusion (AE) operations* Materials Science and Engineering A, **391**, 198–209, 2005.

11. V.M. Segal, *Equal channel angular extrusion: from macromechanics to structure formation*, Materials Science and Engineering A, **271**, 322–333, 1999.

12. V.M. Segal, *Slip line solutions, deformation mode and loading history during equal channel angular extrusion*, Materials Science and Engineering A **345**, 36–46, 2003.

13. J. Alkorta, J.G. Sevillano, *A comparison of FEM and upper-bound type analysis of equal-channel angular pressing (ECAP)*, Journal of Materials Processing Technology, **141**, 313–318, 2003.

14. B.S. Altan, G. Purcek, I. Miskioglu, *An upper-bound analysis for equal-channel angular extrusion*, Journal of Materials Processing Technology, **168**, 137–146, 2005.

15. P.B. Prangnell, C. Harris, S.M. Roberts, *Finite element modelling of equal-channel angular extrusion*, Scripta Mater., **37**, 983–989, 1997.

16. H.S. Kim, M.H. Seo, S.I. Hong, *On the die corner gap formation in equal channel angular pressing*, Materials Science and Engineering A, **291**, 86–90, 2000.

17. A. Rosochowski, R. Rodiet, P. Lipiński, *Finite element simulation of cyclic extrusion-compression*, Metal Forming 2000, Pietrzyk *et al.* [Eds.], 253–259, 2000.

18. A. Rosochowski, L. Olejnik, *Numerical and physical modelling of plastic deformation in 2-turn equal channel angular extrusion*, Journal of Materials Processing Technology, **125–126**, 309–31, 2002.

19. C.J. Luis Pérez, P. González, Y. Garcés, *Equal channel angular extrusion in a commercial Al–Mn alloy*, Journal of Materials Processing Technology, **143–144**, 506–511, 2003.

20. S. Li, M.A.M. Bourke, I.J. Beyerlein, D.J. Alexander, B. Clausen, *Finite element analysis of the plastic deformation zone and working load in equal channel angular extrusion*, Materials Science and Engineering A, **382**, 217–236, 2004.

21. F. Yang, A. Saran, K. Okazaki, *Finite element simulation of equal channel angular extrusion*, Journal of Materials Processing Technology, **166**, 71–78, 2005.

22. R.K. Oruganti, P.R. Subramanian, J.S. Marte, M.F. Gigliotti, S. Amancherla, *Effect of friction, backpressure and strain rate sensitivity on material flow during equal channel angular extrusion*, Materials Science and Engineering A, **406**, 102–109, 2005.

23. A.V. Nagasekhar, Yip Tick-Hon, S. Li, H.P. Seow, *Stress and strain histories in equal channel angular extrusion/pressing*, Materials Science and Engineering A, **423**, 143–147, 2006.

24. ADINA R&D Inc. ADINA, *Theory and Modeling Guides*, Reports ARD 05–7, 2005.

25. N.Q. Chinh, J. Illy, Z. Horita, T. G. Langdon, *Using the stress–strain relationships to propose regions of low and high temperature plastic deformation in aluminum*, Materials Science and Engineering A, **410–411**, 234–238, 2005.

26. Y. Iwahashi, Z. Horita, M. Nemoto, T.G. Langdon, *Factors influencing the equilibrium grain size in equal-channel angular pressing: Role of Mg additions to aluminum*, Metall. Mater. Trans., **29A**, 2503–2510, 1998.

27. S. Komura, Z. Horita, M. Nemoto, T.G. Langdon, *Influence of stacking fault energy on the development of microstructure in equal-channel angular pressing*, Journal of Materials Research, **14**, 4044–4050, 1999.

28. A.V. Nagasekhar, T.H. Yip, *Optimal tool angles for equal channel angular extrusion of strain hardening materials by finite element analysis*, Computat Mater. Sci., 489–495, 2004.

29. J.R. Bowen, A. Gholinia, S.M. Roberts, P.B. Prangnell, *Analysis of the billet deformation behaviour in equal channel angular extrusion*, Materials Science and Engineering A, **287**, 87–97, 2000.

Received October 25, 2006; revised version January 16, 2007.

ENGINEERING TRANSACTIONS • *Engng. Trans.* • **55**, 3, 217–227, 2007
Polish Academy of Sciences • Institute of Fundamental Technological Research

DYNAMIC BUCKLING OF A SANDWICH BAR COMPRESSED BY PERIODICALLY VARIABLE FORCE

W. Morzuch

Wrocław University of Technology
Wrocław, Poland

The paper presents an analysis of dynamic buckling of a sandwich bar compressed by a periodically variable force. In order to determine the stability of the bar transverse motion equations of its transverse vibration were formulated. From the equations of motion, differential equations interrelating of the bar dynamic deflection with space and time were derived. The partial differential equations were solved using the method of separation of variables (Fourier's method). Then an ordinary differential equation (Hill's equation) describing the bar vibration was solved. An analysis of the solution became the basis for determining the regions of sandwich bar motion instability. Finally, the critical damping coefficient values at which parametric resonance occurs have been calculated.

Key words: sandwich bars, stability.

1. Introduction

Sandwich constructions are characterized by light weight and high strength. Such features are highly valuable in aviation, building engineering and automotive applications. The primary aim of using sandwich constructions is to obtain properly strong and rigid structures with vibration damping capacity and good insulating properties. Figure 1 shows a scheme of a sandwich construction which is composed of two thin facing plates and a relatively thick core [4, 5]. The core, made of plastic and metal sheet or foil, transfers transverse forces and maintains a constant distance between the plates. Sandwich constructions are classified into bars, plates and beams. A major problem in the design of sandwich constructions is the assessment of their stability under axial loads causing their buckling or folding. The existing methods of calculating such structures are limited to the assessment of their stability under loads constant in time [3, 5].

There are no studies dealing with the analysis of parametric vibration and dynamic stability (dynamic buckling). This paper presents a dynamic analysis of a sandwich bar compressed by a periodically variable force, assuming that the core is linearly viscoelastic. Differential equations describing the dynamic flexural buckling of bars are derived and regions of instability are identified. The

dynamic analysis of sandwich constructions is of great importance for automotive vehicles and aeroplanes, since most of the loads which occur in them have the form of time-dependent forces.

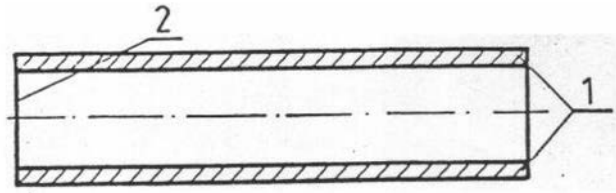

Fig. 1. Scheme of sandwich construction 1 – plates , 2 – core.

2. Dynamic buckling of a sandwich bar

A simply-supported sandwich bar compressed by time-dependent force F is shown in Fig. 2. Force F can be expressed as follows:

$$F = F_1 + F_2 \cos(pt), \tag{2.1}$$

where F_1 – constant component of the compressive force, F_2 – amplitude of the variable component of the compressive force, p – frequency of variable component F_2, t – time.

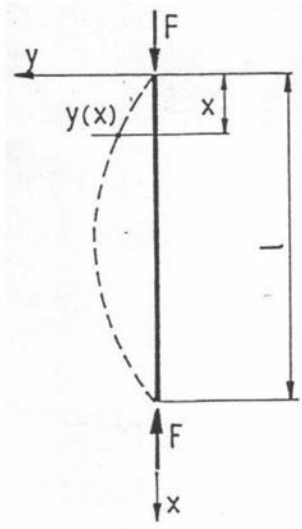

Fig. 2. Sandwich bar compressed by force F.

The cross-section of the sandwich bar is shown in Fig. 3. The basis for describing the dynamic buckling of a sandwich bar is the differential equation of sandwich beam centre line. The equation can be written as

$$B\frac{\partial^4 y}{\partial x^4} = q - k\frac{B}{S}\frac{\partial^2 q}{\partial x^2}, \tag{2.2}$$

where B – flexural rigidity of the bar, q – load intensity, k – a coefficient representing the influence of the transverse force on the deflection of the bar, S – transverse rigidity of the bar.

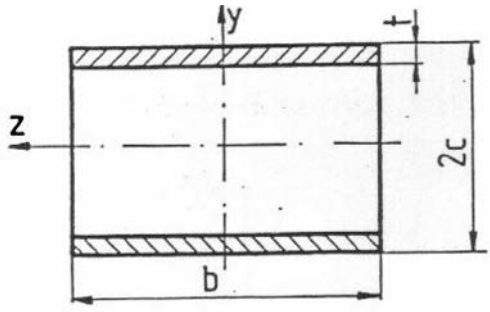

Fig. 3. Cross section of sandwich bar.

In sandwich constructions the core is merely sheared and does not transfer normal stresses whereby coefficient k is equal to one ($k = 1$).

$$S = 2bcG_c, \tag{2.3}$$

where b, c – dimensions of the core (Fig. 3), G_c – modulus of rigidity of the core material.

Load intensity q can be written in the form:

$$q = q_1 + q_2 + q_3, \tag{2.4}$$

$$q_1 = -F\frac{\partial^2 y}{\partial x^2}, \qquad q_2 = -\mu\frac{\partial^2 y}{\partial t^2}, \qquad q_3 = -\eta_r\frac{\partial y}{\partial t}, \tag{2.5}$$

where μ – unit mass of the sandwich bar, η_r – damping coefficient of the core material.

After substituting Eqs. (2.5) into differential Eq. (2.2), the following differential equation is obtained:

$$B\left(1 - \frac{F}{S}\right)\frac{\partial^4 y}{\partial x^4} + F\frac{\partial^2 y}{\partial x^2} - \frac{B}{S}\mu\frac{\partial^4 y}{\partial x^2 \partial t^2} + \mu\frac{\partial^2 y}{\partial t^2} + \eta_r\frac{\partial y}{\partial t} - \frac{B}{S}\eta_r\frac{\partial^3 y}{\partial x^2 \partial t} = 0. \tag{2.6}$$

The above equation is a fourth-order homogeneous equation with time-dependent coefficients. It was solved by the method of separation of variables (Fourier's method). The solution can be presented in the form of an infinite series:

$$y = \sum_{n=1}^{\infty} X_n(x) T_n(t). \tag{2.7}$$

Eigenfunctions $X_n(x)$, satisfying the boundary conditions at the supports of the bar at its ends, have the following form:

$$X_n(x) = A_n \sin\left(\frac{\pi n x}{l}\right). \tag{2.8}$$

Having substituted Eqs. (2.7) and (2.8) into the differential Eq. (2.6), one gets the following ordinary differential equation describing functions $T_n(t)$:

$$\ddot{T}_n + 2h\dot{T}_n + \omega_{on}^2 (1 - 2\psi_n \cos pt) T_n = 0, \tag{2.9}$$

where

$$2h = \frac{\eta_r}{\mu}, \qquad 2\psi_n = \frac{F_2 \left(\frac{\pi n}{l}\right)^2}{\mu \omega_{on}^2}. \tag{2.10}$$

The square of frequency ω_{on} can be expressed as follows:

$$\omega_{on}^2 = \omega_o^2 - \frac{F_1 \left(\frac{\pi n}{l}\right)^2}{\mu}, \tag{2.11}$$

where ω_o – natural frequency of vibration of the bar when $F_1 = 0$, $\eta_r = 0$.

The square of frequency ω_o can be expressed as follows:

$$\omega_o^2 = \frac{B \left(\frac{\pi n}{l}\right)^2}{\mu \left[1 + \frac{B}{S} \left(\frac{\pi n}{l}\right)^2\right]}. \tag{2.12}$$

Differential equation (2.9) is Hill's equation in the form [1, 2]:

$$\ddot{T}_n + 2h\, \dot{T}_n + \Omega_n^2 [1 - f(t)] T = 0. \tag{2.13}$$

If there is no damping in the core ($h = 0$) and assuming $f(t) = 2\psi \cos pt$, one gets the following classical Mathieu equation:

$$\ddot{T}_n + \omega_{on}^2 (1 - 2\psi_n \cos pt) T_n = 0. \tag{2.14}$$

In order to solve Eq. (2.13), a change of variable was made and the solution was expressed in the form:

$$T_n(t) = e^{-ht}\varphi_n(t). \tag{2.15}$$

In this way, a new differential equation for function $\phi_n(t)$ was obtained:

$$\ddot{\varphi_n} + \omega_n^2 [1 - f_1(t)]\varphi_n = 0, \tag{2.16}$$

where

$$\omega_n^2 = \Omega_n^2 - h^2, \tag{2.17}$$

$$f_1(t) = \frac{\Omega_n^2}{\omega_n^2} f(t). \tag{2.18}$$

Equation (2.16) is the Mathieu equation without damping. Therefore for the analysis of this equation one can use the solution of Eq. (2.14), substituting $f_1(t)$ for $f(t)$ and $\Omega_n^2 - h^2$ for ω_n^2.

Let us now analyze the stability of solutions of the differential equation (2.16), limiting the analysis to the first (most important) region of instability. First the Mathieu equation without damping, i.e. Eq. (2.14), should be solved. Here the author's original method has been used for this purpose. The solution of Eq. (2.14), containing the first region of instability, can be presented in the form:

$$T_n(t) = A(t)\cos\frac{pt}{2} + B(t)\sin\frac{pt}{2}, \tag{2.19}$$

where $A(t), B(t)$ – slowly variable functions of time t, satisfying the following conditions:

$$\ddot{A} \ll \dot{A} \ll A, \qquad \ddot{B} \ll \dot{B} \ll B. \tag{2.20}$$

Expression (2.19) was differentiated and substituted into (2.14), whereby the following differential equation describing function $A(t)$ has been obtained:

$$\ddot{A}(t) + \frac{1}{p^2} A(t) \left(\frac{1}{4}p^2 - \omega_{on}^2 - \psi_n\omega_{on}^2\right)\left(\frac{1}{4}p^2 - \omega_{on}^2 + \psi_n\omega_{on}^2\right) = 0. \tag{2.21}$$

The solution of the above equation has the form:

$$A(t) = C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t} \tag{2.22}$$

where λ_1, λ_2 – the roots of the characteristic equation:

$$\lambda^2 + \frac{1}{p^2}\left(\frac{1}{4}p^2 - \omega_{on}^2 - \psi_n\omega_{on}^2\right)\left(\frac{1}{4}p^2 - \omega_{on}^2 + \psi_n\omega_{on}^2\right) = 0. \tag{2.23}$$

Unstable solutions are obtained when $\lambda^2>0$, i.e. when

$$\left(\frac{1}{4}p^2-\omega_{on}^2-\psi_n\omega_{on}^2\right)\left(-\frac{1}{4}p^2+\omega_{on}^2-\psi_n\omega_{on}^2\right)>0. \tag{2.24}$$

Introducing

$$\left(\frac{\omega_{on}}{p}\right)^2=z, \tag{2.25}$$

one obtains condition (2.24) in the following form:

$$\left(\frac{1}{4}-z-\psi_n z\right)\left(-\frac{1}{4}+z-\psi_n z\right)>0. \tag{2.26}$$

The relevant quadratic equation which results from the above inequality has the following roots:

$$z_1=\frac{1}{4\,(1+\psi_n)},\qquad z_2=\frac{1}{4\,(1-\psi_n)}. \tag{2.27}$$

The solution of inequality (2.26) has been presented graphically in Fig. 4 and expressed as:

$$z_1<z<z_2. \tag{2.28}$$

When substitution (2.25) was taken into account, the following relations de scribing the boundary lines of the first region of instability have been obtained:

$$2\sqrt{1-\psi_n}<\frac{p}{\omega_{on}}<2\sqrt{1+\psi_n}. \tag{2.29}$$

The above relations are identical with the generally known expressions which can be found, for example, in [1].

Relation (2.29) is illustrated graphically in Fig. 4. On the basis of the solution of differential Eq. (2.14), differential equation (2.13) was solved and reduced to Eq. (2.16). Finally, the following condition of instability was obtained:

$$\lambda^2>h^2. \tag{2.30}$$

When solving the above inequality, the relation for λ^2 (formula (2.23)) was used, substituting $\Omega_n^2-h^2$ for ω_{on}^2, and the following inequality has been obtained:

$$\frac{1}{p^2}\left[\frac{1}{4}p^2-\left(\Omega_n^2-h^2\right)\left(1+\frac{\Omega_n^2}{\Omega_n^2-h^2}\psi_n\right)\right]\cdot\left[-\frac{1}{4}p^2-\left(\Omega_n^2-h^2\right)\left(\frac{\Omega_n^2}{\Omega_n^2-h^2}\psi_n-1\right)\right]>h^2. \tag{2.31}$$

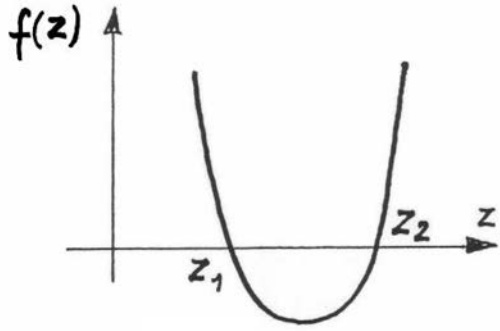

FIG. 4. Solutions z_1 and z_2 of inequality (2.28)

By solving the above inequality, the following condition for the occurrence of parametric resonance for a system with damping has been obtained:

$$\psi_n > 2\sqrt{\xi_n - 2\xi_n^2}, \tag{2.32}$$

where

$$\xi_n = \left(\frac{h}{\Omega_n}\right)^2.$$

Also the following condition for the relative damping coefficient ξ_n at which parametric resonance occurs was derived:

$$0 < \xi_n < \frac{1}{3}. \tag{2.33}$$

The roots of the appropriate quadratic equation which can be derived from inequality (2.32) were expressed as follows:

$$z_1 = \frac{1 - 3\xi_n - \sqrt{\psi_n^2 - 4\xi_n + 8\xi_n^2}}{4\left[(1-\xi_n)^2 - \psi_n^2\right]}, \tag{2.34}$$

$$z_2 = \frac{1 - 3\xi_n + \sqrt{\psi_n^2 - 4\xi_n + 8\xi_n^2}}{4\left[(1-\xi_n)^2 - \psi_n^2\right]}. \tag{2.35}$$

When $z_1 = z_2$, the coordinates of a 'wedge' of instability in system $\left(\psi_n, \frac{p}{\Omega_n}\right)$ are obtained. The equality of roots z_1 and z_2 occurs when the discriminant of inequality (2.31) is equal to zero.

After performing appropriate transformations, the following equation is obtained:

$$z = z_1 = z_2 = \frac{1}{4(1-\xi_n)}. \tag{2.36}$$

Substituting

$$z = \left(\frac{\Omega_n}{p}\right)^2 , \tag{2.37}$$

the following result is obtained:

$$\frac{p}{\Omega_n} = 2\sqrt{1 - 3\xi_n} \, . \tag{2.38}$$

Hence the 'wedge' of the first region of instability has the coordinates:

$$\psi_n = 2\sqrt{\xi_n - 2\xi_n^2} \, , \qquad \frac{p}{\Omega_n} = 2\sqrt{1 - 3\xi_n} \, . \tag{2.39}$$

The boundary lines of the first region of instability are shown in Fig. 5. In a similar way as in the case without damping, the following relations for the boundary lines of the first region of instability are obtained:

$$\frac{p}{\Omega_n} < 2\sqrt{\frac{(1-\xi_n)^2 - \psi_n^2}{1 - 3\xi_n - \sqrt{\psi_n^2 - 4\xi_n + 8\xi_n^2}}} \, , \tag{2.40}$$

$$\frac{p}{\Omega_n} > 2\sqrt{\frac{(1-\xi_n)^2 - \psi_n^2}{1 - 3\xi_n + \sqrt{\psi_n^2 - 4\xi_n + 8\xi_n^2}}} \, . \tag{2.41}$$

Relations (2.40) and (2.41) describe the upper and lower boundary line, respectively.

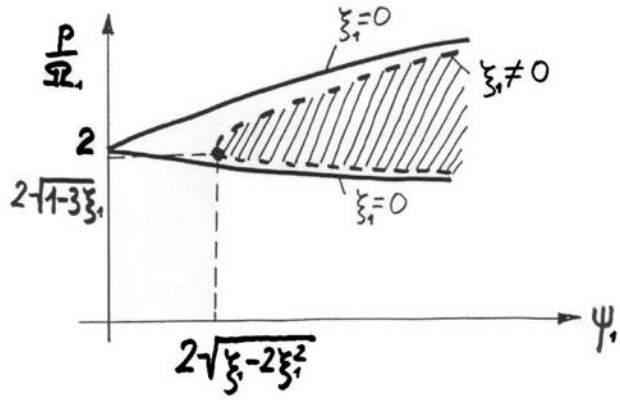

FIG. 5. First region of instability ($\xi_n = 0$, without damping; $\xi_n \neq 0$, with damping)

3. Example of calculations

This section presents calculations of a simply-supported sandwich bar compressed by variable axial force F. The bar dimensions: $b = 25 \cdot 10^{-3}$ m, $c = 7.5 \cdot 10^{-3}$ m, $t = 0.5 \cdot 10^{-3}$ m, $l = 0.5$ m and the following data describing the physical properties of the plate and the core material were assumed.:

$$E_t = 68.67 \cdot 10^3 \text{ MPa}, \qquad G_c = 6.867 \cdot 10^3 \text{ MPa}, \qquad \mu = 14.5 \cdot 10^{-2} \frac{\text{Ns}^2}{\text{m}^2}.$$

The flexural rigidity of the plates is expressed by

$$B = E_t \cdot I = 96.6 \text{ Nm}^2,$$

where I – moment of inertia of the plates' cross-section with respect to the z axis (Fig. 3):

$$I = 2btc^2 = 14062.5 \cdot 10^{-13} \text{ m}^4,$$

The transverse rigidity of the core was calculated from relation (2.3):

$$S = 2bcG_c = 25751 \cdot 10^3 \text{ N}.$$

From relation (2.11), after equating ω_o to zero, the first critical force F_{1kr} was calculated:

$$F_{1kr} = 3809 \text{ N}.$$

It was assumed that constant compressive force component $F_1 = 2000$ N and the variable component amplitude $F_2 = 900$ N. From relations (2.11) and (2.12), ω_o and ω_{on} were calculated:

$$\omega_o = 1018 \text{ s}^{-1}, \quad \omega_{on} = 701.5 \text{ s}^{-1}. \quad (\Omega_n)$$

The corresponding damping coefficient $\xi_n = 0.01$ was assumed. From relation (2.10)

$$\psi_n = 0.25$$

was calculated.

From formula (2.40), the boundary value of coefficient ψ_n at which the parametric resonance occurs was calculated:

$$\psi_{ngr} = 0.198.$$

If $\Psi_n < \Psi_{\text{ngr}}$, no parametric resonance arises. It follows from the above that there exist compressive force components F_1 and F_2 at which the bar does not lose stability.

Assuming that there is no damping ($\zeta_n = 0$), from relation (2.30) the frequencies of the exciting force at which the bar loses stability are calculated:

$$p' = 1567 \text{ s}^{-1}, \qquad p'' = 1215 \text{ s}^{-1}, \qquad (\Delta p = 352 \text{ s}^{-1}),$$

where p' and p'' are the frequencies corresponding to respectively the upper and lower boundary line.

For damping described by coefficient $\zeta_n = 0.01$ the following values were obtained from relations (2.41) and (2.42):

$$p' = 1487 \text{ s}^{-1}, \qquad p'' = 1270 \text{ s}^{-1}, \qquad (\Delta p = 217 \text{ s}^{-1}).$$

The above calculations show that the frequency range Δp in which instability of the bar occurs is smaller when damping is present. Then the calculations for $F_1 = 2000$ N and $F_2 = 1500$ N were performed. From formula (2.10)

$$\psi_n = 0.4$$

was obtained

The following boundary frequency values were obtained:

$$p' = 1660 \text{ s}^{-1}, \qquad p'' = 1086 \text{ s}^{-1} \qquad (\text{if } \zeta_n = 0),$$

$$p' = 1606 \text{ s}^{-1}, \qquad p'' = 1108 \text{ s}^{-1} \qquad (\text{if } \zeta_n = 0.01).$$

The following frequency ranges Δp were obtained:

$$\Delta p = 574 \text{ s}^{-1} \qquad \text{for } \xi_n = 0,$$

$$\Delta p = 498 \text{ s}^{-1} \qquad \text{for } \xi_n = 0.01.$$

4. Conclusions

If the sandwich bar is compressed by a time-dependent force, there are several ranges of the frequency of its variation in which the bar loses stability.

Damping reduces the force variation frequency range in which instability of the sandwich bar occurs.

Owing to damping, there are certain values of compressive force components (F_1, F_2) at which the sandwich bar does not lose stability.

References

1. Z. Dżygadło, S. Kaliski, L. Solarz, E. Włodarczyk, *Vibratiions and waves*, WAT, Warsaw 1965.
2. N.W. McLachlan, *Theory and application of Mathieu functions*, Oxford 1947.
3. W. Morzuch, *Dynamic stability of a sandwich bar compressed by a time-dependent force* [in Polish], Rozprawy Inżynierskie, **37**, 2, 1989.
4. J.F. Plantema, *Sandwich construction*, Stanford University, London 1966.
5. F. Romanów, L. Stricker, J. Teisseyre, *Stability of sandwich structures* [in Polish], Skrypt Politechniki Wrocławskiej, Wrocław 1972.

Received October 20, 2005; revised version July 6, 2006.

ENGINEERING TRANSACTIONS • *Engng. Trans.* • **55**, 3, 229–238, 2007
Polish Academy of Sciences • Institute of Fundamental Technological Research

EVALUATION OF STEEL CLAD PLATE WELDABILITY USING RAM TENSILE TEST METHOD

Z. Zatorski

Naval University of Gdynia
Faculty of Mechanical and Electrical Engineering
ul. Śmidowicza 69, 81–103 Gdynia 3, Poland

The weldability evaluation of steel clad plate is performed using the ram tensile test method. Numerical analysis of the considered problem is performed using the finite element method. The author introduces and verifies a critical ram tension strength R_0 above the yield strength R_e^{II} – as the criterion of the weldability evaluation of clad steel plate, where: $R_e^{II}/R_e^I < 0.707$, R_e^{II}, R_e^I – yield strengths of soft and hard materials.

1. Introduction

Cladding may be performed by rolling or explosive welding or by a combination of the two methods. A manufacturer should demonstrate that the required properties of the base material are preserved after cladding [2]. The evaluation of welded connections can be performed by nondestructive methods, for example ultrasonic testing as well as some destructive methods. The destructive methods of the weldability evaluation are introduced by societies of mechanics, especially Requirements of ASTM and GOST. According to them, shear tests and ram tensile tests are performed. Steel-plated sheets are used mainly in the chemical industry, power industry and chemical tankers, where thin contacting plate layers show high corrosion resistance. The next layers of bearing construction are produced from constructional steel. In view of structural inhomogeneity and the similarity of tensile strength of comprised layers, the tearing often runs in weldable zones. It finds also the reflection in the preparation of ram tensile test specimens according to Requirements of ASTM, Fig. 1 [5, 7, 9]. The Requirements of IACS, LRS, GL and PRS prefer the shear test and side bend test [2, 6, 8]. Cracking in weldable zone of clad plate is anticipated in Requirements of ASTM.

The bonding between the base material and clad material should be adequate to ensure that clad material can not separate from the base material when the manufacturing process or service load are applied.

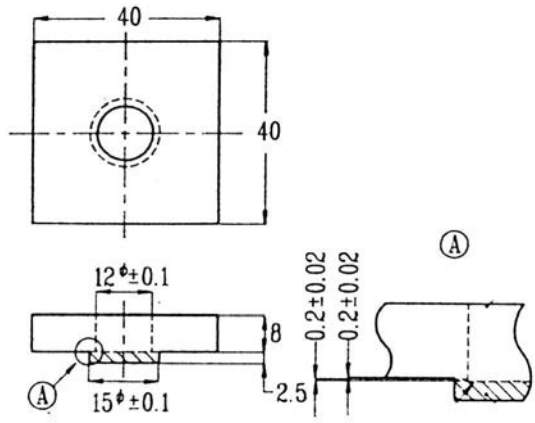

Fig. 1. Dimensions of ram tensile test specimen according to ASTM [5].

The separation processes are localized in the material that has a considerably lower tensile strength. The aim of this paper is elaboration of the ram tensile strength criterion for the above-mentioned clad plates.

2. Experimental procedure

The 30GHMVNb/15G2ANb steel clad plates that have the yield strength $R_{0.2} = 615$ MPa and $R_{0.2} = 325$ MPa, respectively, are supplied inrolled condition for ram tensile tests. The test specimens are cut out in the form of hollow discs. The ram tensile tests are performed on a special test stand as shown in Fig. 2.

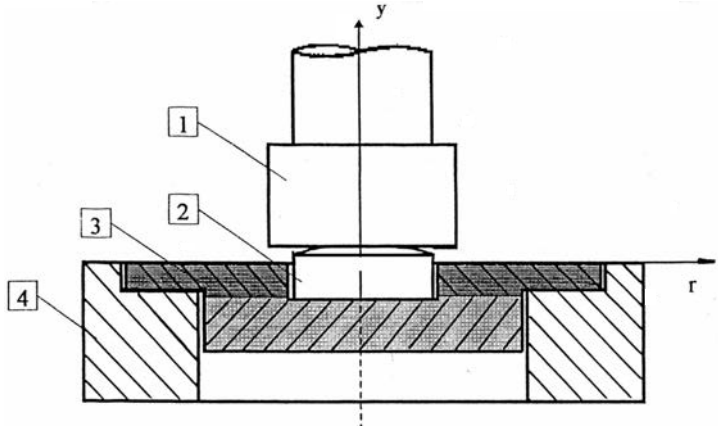

Fig. 2. Test specimen and diagram of the ram tensile test stand, where: 1 – pusher, 2 – stamp, 3 – test specimen, 4 – base.

Characteristic dimensions of the test specimen are shown in Fig. 3.

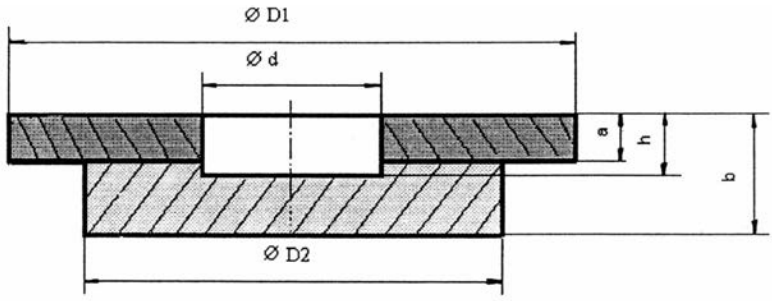

Fig. 3. Characteristic dimensions of the ram tensile test specimen.

The border example of the upper layer is a perfectly rigid material and the bottom layer is a perfectly plastic material. The numerical calculations and experimental verifications will be presented in the following section.

3. Numerical procedures

Experimental observations of the bimaterial system fracture reveal interfacial cracking. The fracture toughness of bimaterial specimen is strongly dependent on the difference in the yield strength and hardening properties of bonded materials. Softening of the material is mainly due to the nucleation and growth processes of microcracks. The coalescence of microcracks leads to the fragmentation process and cracking growth along the interface due to higher stress triaxiality in the softer material. The cracking growth along shear bands in the direction of maximum shear strains reveals away from the interface and into the softer material. From the continuity condition over the interface it follows that when R_e^{II}/R_e^I is less than about 0.7, the harder material remains in the elastic regime [3]. The maximum hoop stress does not occur in harder material just ahead of the crack tip, but away from the interface. The harder phase above the interface constrains the region of plastic deformations and induces high stress triaxiality. Then the maximum mean stress equals

$$\sigma_m = \frac{R_e^{II}}{\sqrt{3}} \left(1 + \frac{3\,\pi}{2}\right), \tag{3.1}$$

if $R_e^{II}/R_e^I < 0.707$, where: R_e^{II}, R_e^I – the yielding stresses of soft and hard materials [3].

Since the voids' nucleation, growth and coalescence result in the growth of damages along the interface, they are under control of mean stresses σ_m and stress triaxiality ξ. Since the maximum shear strains induce the cracking growth

along the shear bands, they are under control of the shear stresses τ and the equivalent stresses σ_H. Both effects should be presented in the numerical analysis of the fracture initiation, especially at geometrical and structural notches of the ram tensile test specimen. The Drucker–Prager flowing condition was adopted by Aifantis [1] and Zhu [10] for metals in view of dependence of the flow stresses K on mean stresses σ_m and expressed by the following equation:

$$f(\sigma_{ij}) = \sqrt{J_2} + \alpha\,\sigma_m - K = 0, \tag{3.2}$$

where: $f(\sigma_{ij})$ – the yielding function, $J_2 = 0.5\ S_{ij} \cdot S_{ij}$ – the second invariant of deviatoric stresses S_{ij}, $S_{ij} = \sigma_{ij} - \sigma_m \cdot \delta_{ij}$, α – the parameter defining the influence of mean stresses $\sigma_m = \sigma_{kk}/3$, $k = 1, 2, 3$ on the flowing stresses K. The ram tensile test method joins the loading force of the test specimen and average stress on the fracture surface. The criterion of average stresses is used to evaluate the clad steel weldability. Finite element method is used to calculate the stresses and strains during simulation of the ram tensile test. The packet of WAT–KAM programmes is applied to calculations. The test specimen is modelled as an axisymmetrical body. The geometry of arrangement with exact qualification of the load and defomation as well as boundary conditions of sub-domains is shown in Fig. 4. The physical model of two-layer steel clad plate is employed for the solution of this problem. Upper layer made of the 30GHMVNb steel and bottom layer of the 15G2ANb steel are used.

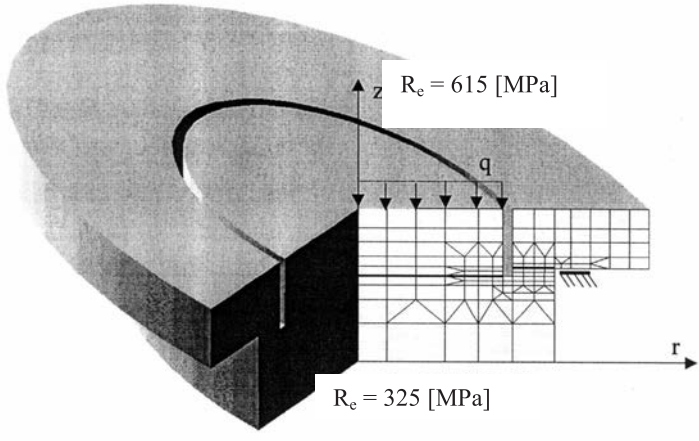

Fig. 4. Geometry of computational arrangement of test specimen for simulation of the ram tensile test.

The stamp is made of the 15G2ANb steel, similarly as the basis. The constant pressure loading is transferred from the pusher to the stamp and then to bottom layer of the test specimen. The support conditions of upper layer of the specimen on the basis are frictionless contact and they are limited to a small surface only.

The boundary conditions determine the global, kinematically admissible displacement field. The boundary conditions on free surface of the specimen

$$\sigma_{ij} \cdot n_j = 0 \qquad (i,j = 1,2,3) \tag{3.3}$$

and the continuity conditions on the interface of the specimen are expressed by

$$u_i^{(I)} = u_i^{(II)}, \tag{3.4}$$

$$\sigma_{ij}^{(I)} \cdot n_j = \sigma_{ij}^{(II)} \cdot n_j, \tag{3.5}$$

where: σ_{ij} – Cauchy stress tensor; u_i – displacement coordinate; n_j – coordinate of the unit normal vector with respect to the material interface.

The tensile testing of round – bar tensile specimens is conducted at a constant crosshead velocity, which corresponds to equivalent strain rate $\dot{\varepsilon} = 10^{-3}$ s^{-1} at the room temperature. Physical report to date from static tensile test of the 15G2ANb – EH 32 steel is defined by the hardening law σ_0

$$\sigma_0(\varepsilon_P) = A\,(\varepsilon_0 + \varepsilon_p)^n\,, \tag{3.6}$$

where: σ_0 – yield stress ($\sigma_0(0) = 292$ MPa), A – hardening multiplier ($A = 683.6$ MPa), n – hardening coefficient ($n = 0.1325$), ε_0 – relative elastic deformation ($\varepsilon_0 = 0.00162$), ε_p – relative plastic deformation.

4. Numerical calculations

The results of calculations: equivalent stresses, mean stresses and stress – triaxiality ratios are presented in a graphical form for equivalent stresses σ_H appointed according to the Huber hypothesis. The distributions of equivalent stresses σ_H as well as stress – triaxiality ratios ξ are shown in Figs. 5 – 7 for characteristic specimens thresholds 0.4 mm and 0.2 mm deep. The characteristic values of equivalent stresses σ_H as well as stress – triaxiality ratios ξ are elaborated by means of numerical analysis of ram tensile tests of the 30GHMVNb/15G2ANb steel plate specimens and they are presented in Table 1. The effect of stress – triaxiality ratios distribution ξ on the cracking processes, is especially essential in the ductile cracking analysis in high stress intensity ranges.

Two clear places of crack initiation step out for the test specimen threshold 0.4 mm deep:

- geometrical notch at internal side, and
- structural notch at external side of area of cracking with the continuity of plastic flow.

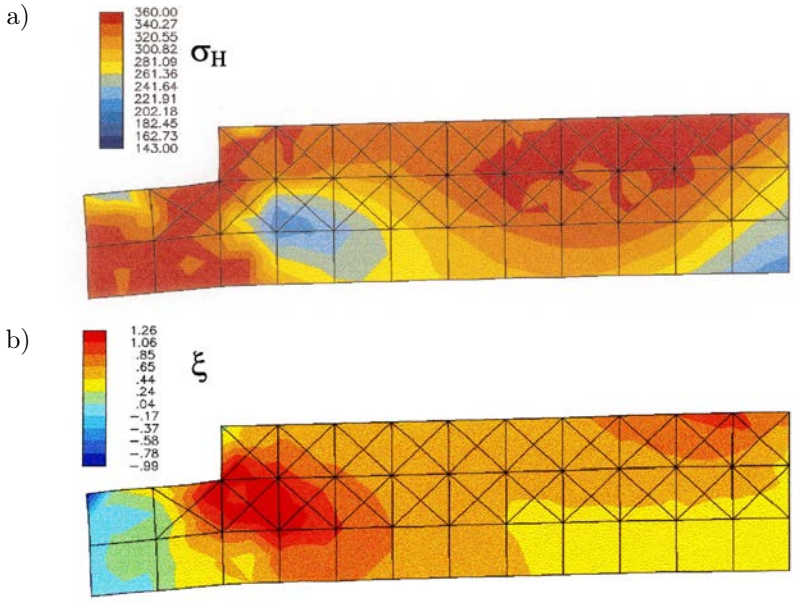

Fig. 5. Equivalent stress distribution σ_H for the 15G2ANb steel layer threshold 0.2 mm deep, of the 30GHMNbV/15G2ANb steel clad plate (a) and stress – triaxiality ratio distribution $\xi = (\sigma_r + \sigma_z + \sigma_t)/3\sigma_H$ (b), for ram tensile test with the load of $F = 38.85$ kN.

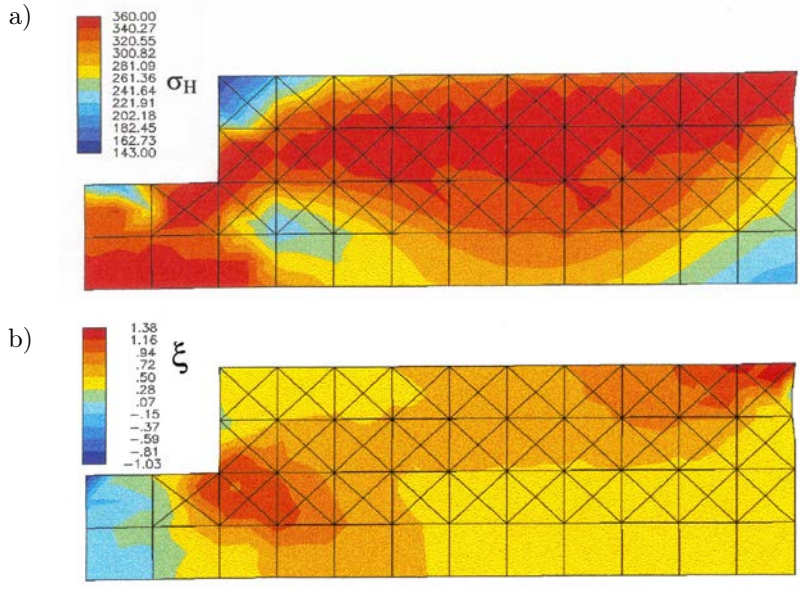

Fig. 6. Equivalent stress distribution σ_H for the 15G2ANb steel layer threshold 0.4 mm deep, of the 30GHMNbV/15G2ANb steel clad plate (a) and stress – triaxiality ratio distribution $\xi = (\sigma_r + \sigma_z + \sigma_t)/3\sigma_H$ (b), for ram tensile test with the load of $F = 38.85$ kN.

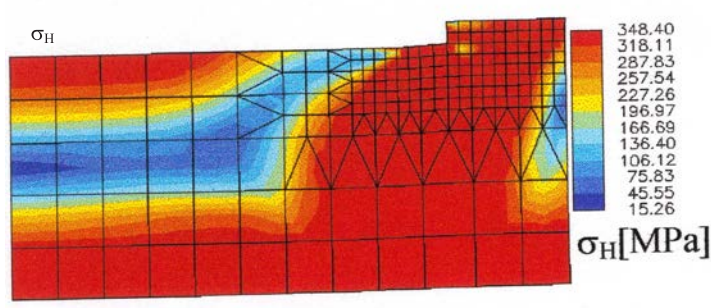

FIG. 7. Equivalent stress distribution σ_H for the 15G2ANb steel layer threshold 0.4mm deep, of the 30GHMNbV/15G2ANb steel clad plate, for ram tensile test with the load of $F = 42.92$ kN.

For the test specimen threshold 0.2 mm deep, the effect of geometrical and structural notches at internal side of cracking area on the stress – triaxiality ratios distribution ξ and equivalent stresses σ_H is clearer with the development of two independent zones of increased plastic flow.

The essential dimensions of the ram tensile test specimen are in approximate agreement with Requirements of the ASTM. It is the reason that good dimensional occuracy of the ram tensile test specimen should be ensured.

The Drucker–Prager flow condition (Eq. (3.2)) is adopted to the following expression

$$K = \sigma_H(1 + \alpha \cdot \xi), \tag{4.1}$$

where $\xi = \sigma_m/\sigma_H$, and the above numerical results are verified according to the experimental data. The highest values of equivalent stresses σ_H as well as stress – triaxiality ratios ξ occur at the specimen threshold 0.4 mm deep as the structural and geometrical notches (1, 3), Table 1.

Table 1. Characteristic values of equivalent stresses σ_H and stress – triaxiality ratios ξ obtained as a result of numerical analysis of ram tensile tests of the 30GHMVNb/15G2ANb steel clade plate.

Designation		Equivalent stress σ_H and stress – triaxiality ratio ξ			
No	specimen threshold	0.2 [mm]		0.4 [mm]	
–	specimen property	σ_H, [MPa]	ξ, [–]	σ_H, [MPa]	ξ, [–]
1	geometr. notch at internal side	340 – 360	1.08 – 1.28	340 – 360	1.10 – 1.30
2	structural notch at internal side	320 – 340	0.24 – 0.44	143 – 162	0.28 – 0.51
3	structural and geometrical notch at external side	320 – 340	0.65 – 0.85	340 – 360	1.10 – 1.30

The load of $F = 42.92$ kN (Fig. 7) produces equivalent stresses σ_H within the range (318–348) MPa with the plastic flow of material in the fracture's area. The highest values of equivalent stresses σ_H occure at the specimen threshold 0.2 mm deep and structural notch at the internal side of specimen (2) at equivalent stresses σ_H within the range (320–340) MPa, Table 1.

5. Experimental Verification

The characteristic dimensions of specimens of the 30GHMVNb/15G2ANb steel clad plate are obtained by means of numerical analysis of the ram tensile tests, Table 2.

The average values of ram tensile strength R_0 are lower than the tensile strength $R_m = 510$ MPa and higher than the yield strength $R_e = 325$ MPa for the 15G2ANb steel of the clad plate.

Table 2. Characteristic dimensions of specimens of the 30GHMVNb/15G2ANb steel clade plate and results of the ram tensile tests.

Designation	Dimensions $D_1 \times D_2 \times d$, [mm]	Maximum load F_0, [kN]	Ram tensile strength R_0, [MPa]
1	$30 \times 20.3 \times 16$	42.5	375
2	$30.1 \times 20.4 \times 16$	41.5	361

The test specimen fracture after the ram tensile test shows effect of the rolled texture and is connected with anisotropy, Fig. 8. The point of crack initiation is placed at geometrical notches of softer layer of clad plate what is confirmed by the numerical calculations.

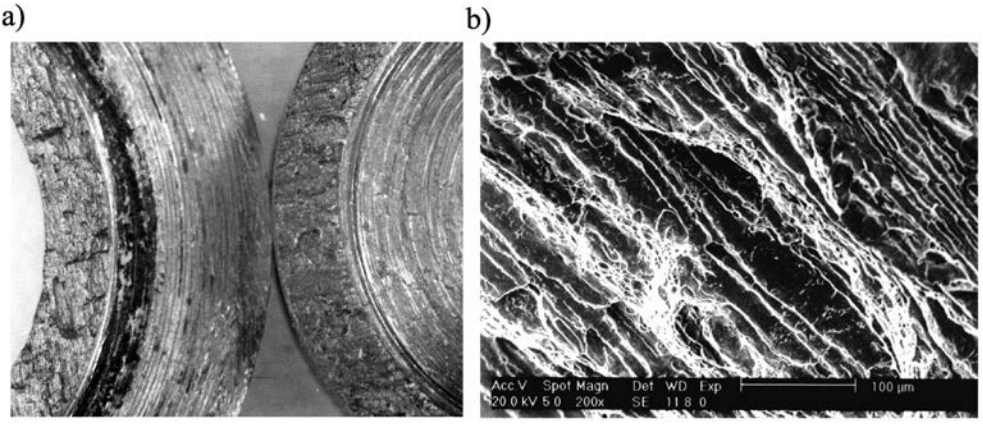

Fig. 8. The ram tensile test specimen fracture (a) and characteristic fracture of the 15G2ANb steel layer of a clad plate, SEM (b).

The plastic flow of material in the fracture area allows to specify possible places of initiation and directions of cracking that has been confirmed by fractographic investigations.

The criterion of weldability evaluation of the clad plate is performed using the ram tensile test. The highest values of equivalent stresses σ_H and stress – triaxiality ratios ξ occur at the the specimen ends as well as at structural and geometrical notches, for the equivalent stresses σ_H above 320 MPa, Table 1. At the same time, the yield strength for the 15G2ANb steel is $R_e = 325$ MPa.

The results presented above show that the plastic flow of soft material is decisive actor in the area of real fracture. On this basis, the author introduces the critical ram tensile strength R_0 higher than the yield strength of soft layer material R_e^{II} as the criterion to evaluate the clad steel plate weldability according to the following inequality

$$R_0 \geq R_e^{II}, \tag{5.1}$$

where $R_e^{II}/R_e^{I} < 0.707$.

The new criterion differs from the hitherto used, assumed a priori values of critical stresses to evaluate the clad plates weldability according to Requirements of the ASTM.

The parameters of explosive welding conditions can be properly chosen using the ram tensile test and the above criterion [3, 4, 5].

6. Conclusions

1. The criterion to evaluate the clad steel plate weldability is performed using the ram tensile test.
2. The finite element method is used to calculate the stresses and strains during the simulation of the ram tensile test.
3. The author introduces and verifies the critical ram tensile strength R_0, higher than the yield strength of soft layer material R_e^{II}, as a criterion to evaluate the steel clad plate weldability, using the ram tensile test method according to the following inequality

$$R_0 \geq R_e^{II},$$

where: $R_e^{II}/R_e^{I} < 0.707$, R_e^{II}, R_e^{I} – the yield strength of soft and hard layer material.

References

1. E.C. Aifantis, International Journal of Plasticity, **3**, 211–217, 1987.
2. Germanischer Lloyd of Shipping, *Rules for classification and construction*, Ch. II, *Material and welding technology*, P.1, Sec.1L, *Clad plates*, 28–30, 1992.

3. S. Hao, W.K. Liu, *Bimaterial interfacial crack growth with strain gradient theory*, Journal of Engineering Materials and Technology, **121**, 10, 413–442, 1999.

4. T. Izuma, K. Hokamoto, M. Fujita, M. Aoyagi, *Single – shot explosive welding of hard – to – weld JIS A5083/SUS304 clad using SUS304 intermediate plate*, Transactions of the Japan Welding Society, **23**, 1, 46–51,1992.

5. K. Hokamoto, T. Izuma, M. Fujita, *New explosive welding technique to weld aluminium alloy and stainless steel plates using a stainless steel intermediate plate*, Metallurgical Transactions A, **24A**, 10, 2289–2297, 1993.

6. Loyd's Register of Shipping, *Rules for the manufacture testing and certification of materials*, P.2., Ch.3, Sec. 7.7, *Clad plates*, 1991.

7. M. Nishida, Z. Murakami, *Behavior of bonded interface of explosive clad steel*, Transactions of the Japan Welding Society, **23**, 1, 9–16, 1992.

8. Polish Register of Shipping, *Rules for Classification and Shipbuilding* [in Polish], 1995, afterwards changes P. IX, Ch.9.5, *Materials and welding, Clad steels,* Gdańsk 1998.

9. W. Walczak, *Explosive welding of metals* [in Polish], WNT, Warszawa 1989.

10. X.H. Zhu, J.E.C. Arsley, W.M. Milligan, E.C. Aifantis, *On the failure of pressure – sensitive plastic materials, Part 1. Models of yield and shear band behavior*, Scripta Materialia, **36**, 6, 721–726, 1997.

Received March 23, 2006; revised version November 2, 2006.

ENGINEERING TRANSACTIONS • *Engng. Trans.* • **55**, 3, 239–275, 2007
Polish Academy of Sciences • Institute of Fundamental Technological Research

STRUCTURAL HEALTH MONITORING – A REVIEW WITH THE EMPHASIS ON LOW-FREQUENCY METHODS

P. K o ł a k o w s k i

Smart-Tech Centre
Institute of Fundamental Technological Research, PAN
Warsaw, Poland

Structural Health Monitoring (SHM) is a fast-developing, interdisciplinary field of research having its roots in vibroacoustics and non-destructive testing and evaluation. Fast development of the area is due to the fact that SHM is heavily stimulated by the engineering problems of maintenance and safe operation of technical infrastructure. The use of SHM is slowly becoming a standard in high-cost, modern infrastructure. Therefore, the possibility of application should always be on the horizon of any related research work. Thus far, the majority of SHM applications have been demonstrated in civil, aerospace and mechanical engineering. This paper reviews the main achievements and points out the current trends in this field with the emphasis on low-frequency methods.

1. Introduction

The point of this publication is to make the Reader acquainted with the historical milestones and up-to-date trends in the field of Structural Health Monitoring. A concise and informative definition of Structural Health Monitoring can be found in Sohn *et al.* [72]: "The process of implementing a damage detection strategy for aerospace, civil, and mechanical engineering infrastructure is referred to as *Structural Health Monitoring* (SHM). The SHM process involves the observation of a system over time, using periodically sampled dynamic response measurements from an array of sensors, the extraction of damage-sensitive features from these measurements, and the statistical analysis of these features to determine the current state of the system's health." The damage state of a system can be considered as a five-step process, as discussed in Rytter [70]. It can be described by answering the following questions:

1. Is there any damage in the system (existence)?
2. Where is the damage in the system (location)?
3. What kind of damage has occurred (type)?
4. How severe is the damage (extent)?
5. How much useful life remains (prognosis)?

The importance of diagnostics was probably first appreciated by the community of mechanical engineers dealing with *rotating machinery*, where damage involves high risk to staff and high cost of repair e.g. in turbines. NATKE, CEMPEL [61] present *vibroacoustics* as a tool for machine diagnostics, which still remains a challenge in practice as shown in EISENMANN, EISENMANN [19]. The specific feature of rotating machinery is that it is self-exciting and relatively compact (not of huge dimensions), which is often not the case in SHM. The machine diagnostics is usually a low-frequency problem.

Many SHM methods originate from the *Non-Destructive Testing and Evaluation* (NDT/E) methods i.e. ultrasonic testing, radiographic testing, acoustic emission, penetrant testing, magnetic particle inspection, eddy currents, and optical holography, which are successfully applied in the industry for local detection of flaws in structural components. Some NDT/E methods require external excitation, for instance ultrasonic testing, others do not, for instance acoustic emission. The NDT/E methods usually operate in high frequencies.

The objective of SHM is to create a monitoring system (possibly for the whole structure, which is sometimes of complicated topology and considerable dimensions), able to track changes in structural condition continually and raise appropriate alerts if a defect is detected. As a consequence of the evolution from the two major streams, i.e. machine diagnostics and NDT/E, the SHM methods for identifying structural damage can be roughly split into low-frequency methods (non-ultrasonic) and high-frequency methods (ultrasonic), respectively. As examples, low-frequency methods are used in civil engineering for examining stiffness degradation of a bridge, whilst high-frequency methods are used in aerospace engineering for crack identification in a wing.

As the author deals with low-frequency methods in his research, the focus of the article is naturally moved towards these methods. Special attention is paid to *system identification* as the basis for most low-frequency SHM methods currently used in engineering practice. Subsequent *damage identification* widely relies on the system identification methods.

The organization of the article is the following: Sec. 2 describes the measuring devices and techniques commonly used in SHM. Section 3 takes up the problem of system identification, which is a crucial issue in low-frequency SHM, because a well-recognized reference structure is needed for tracking subsequent changes in behaviour due to damage. Section 4 briefly characterizes the low-frequency (vibration-based) methods. Section 5 describes the NDT/E methods, which gave an impetus for the development of high-frequency (ultrasonic) SHM methods, the quick review of which is presented in subsequent Sec. 6. Section 7 reports the applications of artificial intelligence to perform *signal processing* or damage identification. Section 8 mentions examples of testing the SHM methods on real structures. Finally, Sec. 9 highlights modern trends in the field. A selection

of the essential references, including the periodically-held, key events on SHM [51, 25, 21, 74], the related events [3, 29, 42] and the relevant journals in the field [7, 17, 67] is enclosed.

2. Measuring devices and techniques in SHM

Almost all the SHM methods analyze a structural response due to excitation with an actuator. In order to capture the response, various sensors are mounted on the structure. This chapter briefly describes the most frequently used devices and techniques for making accurate measurements. The quality of the measurements is absolutely essential in SHM.

2.1. Piezoelectric transducers

Probably the most common measuring devices used in SHM are piezo-electric transducers because of their outstanding electromechanical properties, relatively low price, and both actuating and sensing capabilities. They have been used for years in classical NDT/E methods like ultrasonic testing or acoustic emission. An in-depth presentation of piezo-electric sensors is given in Gautschi [27].

The direct piezoelectric effect, utilized in sensors, is present when a mechanical deformation of the piezoelectric material produces a proportional change in the electric polarization of that material (electric charge appears on opposite faces of the material). The converse piezoelectric effect, utilized in actuators, means that an acting external electric field induces proportional mechanical stress in the piezoelectric material (the material is deformed when an electric voltage is applied).

A rather restricted number of piezoelectric materials have been found suitable for transduction elements in piezoelectric sensors. Basically, natural (e.g. quartz, tourmaline) and synthetic single crystals (e.g. gallium orthophosphate, crystals of the CGG group), piezoelectric ceramics (e.g. lead-zirconate-titanate – denoted as PZT), and thin films (e.g. polyvinylidene fluoride – denoted as PVDF) can be used. Piezoelectric materials used in sensors combine excellent mechanical properties with a high piezoelectric sensitivity at a low production cost. They have a number of advantages, which makes them particularly suitable for dynamic measurements. Piezoelectric sensors have extremely high stiffness (their deflections are usually in the μm range), high natural frequency (hundreds of kHz), wide measuring range and wide operating temperature range. The crystal-based sensors have very high stability whereas the ceramic-based ones can be produced in commercial quantities. Quasi-static measurements are possible with sensors having single crystals as transduction elements.

Piezoelectric sensors can directly measure the force, strain, acceleration, and pressure.

2.2. Fibre optics

Another widespread sensor with a growing number of applications in SHM is an optical fibre. An extensive review of various types of optical fibre sensors is presented in Udd [82] and Briley [9].

The phenomenon of guiding light by a transparent cylinder has been known since the antique era (Egyptians). The early light-guiding materials (glass) had very high optical loss parameters i.e. hundreds or thousands decibel per kilometre. A real technology breakthrough occurred in 1966, when Kao and Hockham [44] envisaged the fabrication of a glass fibre with optical loss lower than 20 dB/km and its potential application in telecommunications.

Modern optical fibres are composed of high-silica glass doped with some oxides to achieve a required refraction index. An optical fibre consists of a core encapsulated in a cladding with a smaller refraction index. This enables total internal refraction at some incidence angle of entering light. Optical fibres can be generally divided into two types: multi-mode and single-mode. Of special importance in strain sensor technology are single-mode optical fibres. Single-mode fibre limits its guidance capability for a chosen wavelength to one mode thanks to small diameter of the core and small difference between the core and cladding refraction indices. In practice however, instead of just one mode, two orthogonally-polarized, strongly-coupled modes are carried by most commercially produced fibres. This effect, called *birefringence*, is often preferable since it helps the fibre to maintain the polarization of a guided wave and to transmit it long distance. Light launching into a single-mode fibre of small core is difficult and favours a light source with highly directional output, e.g. light emitting diode or injection laser diode.

The major division of the fibre optic sensors distinguishes the interferometric sensors for outside application and fibre Bragg grating sensors for inside application (embedded in the structure).

The most widespread interferometric sensors interrogate a measurand-induced change of phase in the light propagating along a single-mode optical fibre. Several different layouts of interferometric sensors may be used, depending on arrangement of optical paths. The most common configurations, based on phase change analysis, are the Michelson and Fabry-Perot interferometers. Interferometric optical fibre sensors provide high-sensitivity measurements. They require a special signal recovery technique (demodulation) to perform absolute measurements. Development of low-cost fabrication methods that do not compromise the strength and fatigue life of the optical fibres should facilitate the wide use of the sensors.

Fibre Bragg grating (FBG) sensors are highly sensitive devices as well. Their manufacturing process is automated and ensures no strength loss of the optical fibre. The intracore Bragg grating fibre optic sensor relies on the narrow-band

reflection from a fibre segment of periodic variations (gratings) in the core index of refraction of a single-mode fibre. If FBG sensors are used with the ratiometric demodulation system, they can build robust, absolute-measurement, low-cost sensing system, which can be integrated on an optoelectronic chip, easily interconnected with the structure.

Interferometric optical fibre sensors and intracore FBG sensors have a great potential to become widely used instruments for strain-like measurements.

2.3. Other sensors

Other sensors used in SHM are: electro-magnetic acoustic transducers (EMAT) widely used in ultrasonic testing (see Sec. 5), micro-electro mechanical systems (MEMS), and laser interferometers – used in optical holography (see Sec. 5).

3. System identification

System identification is the process of developing a faithful mathematical model of a physical system using experimental data. The point of vibration-based system identification is determination of modal characteristics of a structure using either known or unknown excitation. It is a *sine qua non* stage of the vast majority of low-frequency SHM analyses, which is always carried out before the commencement of the damage identification procedure. As a result of system identification a reference model for SHM is determined, thereby enabling the tracing of subsequent states of damage.

System identification theory stems basically from the control theory. In this work a brief summary of system identification is given, emphasizing the methods used in SHM. Simple peak picking method was the origin of modal testing in the frequency domain. However, some subsequent and more sophisticated methods turned into the time domain, using only the output measurements for extracting modal parameters. Thorough treatment of the subject can be found in Juang [41], Ljung [49] and an overview in Peeters [64].

3.1. Useful models

The equations of motion of a linear-dynamic mechanical system of n_2 independent variables are second-order differential equations of the form:

$$M\ddot{w} + \zeta\dot{w} + Kw = f(w,t). \tag{3.1}$$

For the sake of performing efficient computational analysis, it is convenient to reshape Eq. (3.1) into the following compact form:

$$\dot{x} = A_c x + B_c u, \tag{3.2}$$

where:

$$A_c = \begin{bmatrix} 0 & I \\ -M^{-1}K & -M^{-1}\zeta \end{bmatrix}, \qquad B_c = \begin{bmatrix} 0 \\ M^{-1}B_2 \end{bmatrix},$$

$$x = \begin{bmatrix} w \\ \dot{w} \end{bmatrix}, \qquad f(w,t) = B_2 u\,(t),$$

A_c $(n \times n)$, $n = 2n_2$, is the state matrix and B_c $(n \times r)$ is the input matrix, with r being the number of inputs.

The response of the dynamic system can be measured by various sensors e.g. accelerometers, velocity sensors, strain gauges, and stored in the output vector $y(t)$ of q entries:

$$y = C_a\ddot{w} + C_v\dot{w} + C_d w, \tag{3.3}$$

where C_a, C_v, C_d are output matrices corresponding to acceleration, velocity and displacement, respectively. The matrices contain conversion factors between the measured units (e.g. meters) and electrical units (e.g. volts) indicated by the measuring equipment.

Solving (3.1) for $\ddot{w}$ and substituting the result into (3.3) yields:

$$y = Cx + Du, \tag{3.4}$$

where: $C = \left[C_d - C_a M^{-1}K \;\; C_v - C_a M^{-1}\zeta\right]$, $D = C_a M^{-1}B_2$, $C(q \times n)$ is the output matrix and $D(q \times r)$ is the direct transmission (throughput) matrix.

Equations (3.2) and (3.4) constitute a continuous-time, deterministic state-space model of a dynamical system. In practice however, the measurements are available at discrete time instants. Thus the discrete model, sampled in equally spaced instants 0, Δt, $(k+1)\Delta t, \ldots$, yields:

$$x_{k+1} = Ax_k + Bu_k, \tag{3.5}$$

$$y_k = Cx_k + Du_k, \tag{3.6}$$

where: $A = e^{A_c \Delta t}$, $B = B_c \int\limits_0^{\Delta t} e^{A_c \tau} d\tau$, $\tau = (k+1)\,\Delta t - \tau'$.

Assuming a system to be in a steady state and solving for the output y_k in terms of the previous inputs u_i $(i = 0, 1, \ldots, k)$ with zero initial condition, produces:

$$y_k = \sum_{i=1}^{k} Y_i u_{k-i} + Du_k, \tag{3.7}$$

where Y_k are the *Markov parameters*:

$$Y_k \equiv CA^{k-1}B, \quad Y_0 \equiv D, \tag{3.8}$$

Note that $Y_0 = D$. Equation (3.7) is called the *weighting sequence description* because the contribution to the output at time step k is made by the input at time step k and all previous steps $k-1$, $k-2$, $\ldots, 1, 0$, weighted by the *pulse response sequence* (Markov parameters).

Another useful model can be obtained by adding and subtracting the same term Gy_k in Eq. (3.5). Then, substituting y_k from (3.6), gives:

$$x_{k+1} = \bar{A}x_k + \bar{B}\bar{u}_k, \tag{3.9}$$

where: $\bar{A} = A + GC$, $\quad \bar{B} = \begin{bmatrix} B + GD & -G \end{bmatrix}$, $\quad \bar{u}_k = \begin{bmatrix} u_k \\ y_k \end{bmatrix}$.

Equations (3.9) and (3.6) constitute a discrete-time, deterministic *state-space observer* model with the matrix G being the observer gain matrix.

For a steady state and zero initial conditions, the state-space observer model can be expressed as:

$$y_k = \sum_{i=1}^{k} \bar{Y}_i \bar{u}_{k-i} + Du_k, \tag{3.10}$$

where:

$$\begin{aligned} \bar{Y}_k &\equiv C\bar{A}^{k-1}\bar{B} = \left[C\left(A + GC\right)^{k-1}\left(B + GD\right) \;\; -C\left(A + GC\right)^{k-1}G \right], \\ \bar{Y}_0 &\equiv D, \end{aligned} \tag{3.11}$$

$\bar{Y}_k$ are the *observer Markov parameters*. Note that $\bar{Y}_0 = D$. Equation (3.10) is commonly called the linear difference model or ARX *model* where AR refers to the AutoRegressive part (output data) and X refers to the eXogeneous part (input data). The ARX model corresponding to the state-space observer model is analogous to the weighting sequence description corresponding to the state-space model (cf. (3.7)).

If we decide to account for noise in modelling and measurements, the following discrete-time, deterministic-stochastic state-space model has to be considered:

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + w_k, \\ y_k &= Cx_k + Du_k + v_k, \end{aligned} \tag{3.12}$$

where w_k is the process noise due to modelling imperfections and v_k is the measurement noise due to equipment imperfections. If the input u_k is unknown

e.g. in the case of ambient vibration, the corresponding terms disappear from Eqs. (3.12), and the discrete-time, stochastic state-space model yields:

$$\begin{aligned} x_{k+1} &= Ax_k + w_k, \\ y_k &= Cx_k + v_k. \end{aligned} \tag{3.13}$$

Both noise components are assumed to be zero mean $E[w_k] = E[v_k] = 0$, white, independent of the actual state $E[x_k\ w_k^T] = E[x_k\ v_k^T] = 0$, with the following covariance matrices:

$$E\left[\begin{pmatrix} w_m \\ v_m \end{pmatrix} \begin{pmatrix} w_n^T & v_n^T \end{pmatrix} \right] = \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix} \delta_{mn}. \tag{3.14}$$

The stochastic process is also zero mean $E[x_k] = 0$ and stationary $E[x_k\ x_k^T] = \Sigma$, meaning that the state covariance matrix Σ is constant in time. The set of i output covariance matrices and the next state-output covariance matrix are defined as:

$$R_i = E\left[y_{k+i}\, y_k^T\right], \qquad \Gamma = E\left[x_{k+1}\, y_k^T\right]. \tag{3.15}$$

With the definitions (3.15), the following relations can be derived:

$$\Sigma = A\Sigma A^T + Q, \qquad \Gamma = A\Sigma C^T + S, \tag{3.16}$$

$$R_i = CA^{i-1}\Gamma, \qquad R_0 = C\Sigma C^T + R. \tag{3.17}$$

Equation (3.17) resembles the definition of Markov parameters in Eqs. (3.8) or (3.11). This formal similarity is crucial as it allows to use the system realization theory of deterministic time-invariant model in stochastic considerations. Therefore, it is the basis for developing stochastic, output-only system identification methods.

By applying a steady-state *Kalman filter* to the stochastic state-space model, another representation called the *forward innovation* model can be obtained:

$$\begin{aligned} z_{k+1} &= Az_k + Ke_k, \\ y_k &= Cx_k + e_k, \end{aligned} \tag{3.18}$$

where the white noise sequence e_k are called innovations and K is the so-called Kalman filter gain matrix. Note that the state vector z_k is in a different basis than x_k (cf. (3.13)). The corresponding general model (non-state-space) is the *Auto Regressive Moving Average* (ARMA) model with the AR term related to outputs and the MA term related to white noise input:

$$y_k + \alpha_1 y_{k-1} + ... + \alpha_{n_\alpha} y_{k-n_\alpha} = e_k + \gamma_1 e_{k-1} + ... + \gamma_{n_\gamma} e_{k-n_\gamma}. \tag{3.19}$$

For an ARMA model deduced from a state-space model, the AR order n_α is equal to the MA order n_γ.

3.2. Fundamentals of system identification

In order to transform time-dependent linear differential equations (the equations of motion) into the algebraic domain (frequency domain), we need the *Laplace transform* defined as:

$$L\left[y\left(t\right)\right] = \int_0^\infty y\left(t\right) e^{-st} dt, \tag{3.20}$$

where the scalar $s = \sigma + j\omega$ is a complex variable. The Laplace transform applies to the continuous time analysis and is customarily called the *s-transform*. Two-sided Laplace transform covering the whole range $(-\infty, \infty)$ is called *continuous Fourier transform*.

In practice, the data for system identification are sampled at discrete time instants $k\Delta t$ $(k = 0, 1, 2, \ldots)$, and an analogous *z-transform* in the discrete time is defined as:

$$Z\left[y\left(k\Delta t\right)\right] = \sum_0^\infty y\left(k\Delta t\right) z^{-k}. \tag{3.21}$$

Note that the s-transform of the sampled time signals can be obtained from its z-transform by substituting $z = e^{s\Delta t}$.

It is known that any periodic signal can be expressed in terms of a *Fourier series*, being a linear combination of cosine and sine terms. An arbitrary discrete-time input signal of the frequency ω, captured for m samples at Δt-spaced instants $\tau = 0, 1, \ldots, m-1$, can be expressed in terms of the following Fourier series:

$$u\left(\tau\right) = \sum_{i=-\infty}^{\infty} \hat{U}\left(i\right) e^{j(\omega\tau\Delta t)i} = \sum_{i=-\infty}^{\infty} \hat{U}\left(i\right) e^{j\left(\frac{2\pi\tau}{m}\right)i}, \tag{3.22}$$

$$i = -\infty, \ldots, 0, 1, \ldots, \infty,$$

where $\hat{U}\left(i\right)$ are the scaling coefficients (weighting amplitudes) for the frequencies in the decomposition.

For performing numerical analysis, the infinite series must be truncated, which is equivalent in practice to consideration of a limited frequency range. Thus the input signal becomes:

$$u\left(\tau\right) = \sum_{k=0}^{m-1} U\left(k\right) e^{j\left(\frac{2\pi\tau}{m}\right)k}, \qquad k = 0, 1, \ldots, m-1. \tag{3.23}$$

In signal processing, the weighting amplitudes $U(k)$ need to be determined, which can be obtained through the inverse transform of (3.23):

$$U(k) = \frac{1}{m} \sum_{n=0}^{m-1} u\,(n)\, e^{-j\left(\frac{2\pi k}{m}\right)n}, \qquad n = 0,\, 2,\, \ldots,\, m-1. \tag{3.24}$$

The formula (3.24) is called the *Discrete Fourier Transform* (DFT) of the sampled input sequence $u(n)$. The Inverse Discrete Fourier Transform (IDFT) is naturally the formula (3.23).

Note that $U(k)$ is periodic, such that $U(k) = U(k+m)$. As the sequence $u(n)$ is real, the equality $U(-k) = U^*(k)$ holds, where * means a complex conjugate. One can show that the values of $U(k)$ $(k = m/2,\, m/2+1, \ldots, m)$ are the complex conjugates of the values $U(k)$ $(k = m/2,\, m/2-1, \ldots, 0)$. Only three values $U(0)$, $U(m/2)$ and $U(m)$ are always real, the other ones are complex. This implies that $U(k)$ should be uniquely determined over the interval $[0, m/2]$, corresponding to $[0,\, \pi]$. Consequently, the maximum frequency captured by the DFT (due to truncation of the infinite series) depends upon the time interval Δt as follows:

$$f_{nyq} = \frac{m}{2} \frac{1}{m\Delta t} = \frac{1}{2\Delta t}. \tag{3.25}$$

This is the *Nyquist frequency*. Hence, the highest frequency that can be estimated by the sampling rate $1/\Delta t$ Hz is only half that rate.

Applying the DFT directly in numerical analysis would require $\mathrm{O(m^2)}$ operations in contrast to only O(mlog(m)) ones performed by the efficient *Fast Fourier Transform* (FFT) algorithm. There exists a number of the FFT implementations, the most popular of which is the Cooley–Tukey [14] version. Performing the z-transform (3.21) of the weighting sequence description (3.7), one transfers the input/output relation to the frequency domain:

$$y(z) = \sum_{\tau=0}^{\infty} Y_\tau z^{-\tau} u\,(z), \qquad z = e^{j\omega\Delta t}, \tag{3.26}$$

where

$$G(z_k) = \sum_{\tau=0}^{\infty} Y_\tau z_k^{-\tau}, \qquad z_k = e^{j\frac{2\pi k}{m}} \tag{3.27}$$

is the *Frequency Response Function* (FRF). The infinite series in (3.27) can be truncated if the system is asymptotically stable.

A periodic input $u(n)$ will produce a similar periodic output $y(n)$ scaled in amplitude and shifted in phase. Thus having DFTs (cf. (3.24)) $U(k)$, $Y(k)$ of

the input $u(n)$ and output $y(n)$ signal for a single-input single-output (SISO) system, the Frequency Response Function can be simply expressed as:

$$G(z_k) = \frac{Y(k)}{U(k)}. \tag{3.28}$$

The relation between the state-space model and the FRF is given by:

$$G(z_k) = C(z_k I_n - A)^{-1} B + D. \tag{3.29}$$

For a multi-input multi-output (MIMO) system with N data records available, the FRF becomes:

$$G(z) = \frac{\sum_{i=1}^{N} Y_i(k) U_i^*(k)}{\sum_{i=1}^{N} U_i(k) U_i^*(k)}. \tag{3.30}$$

An equivalent of FRF for the Laplace transform (continuous time) is the *transfer function.*

In system identification one uses the correlation functions for the input and output signals. The *cross-correlation* function R_{yu} between the output and input (delayed by the time $i\,\Delta t$), and the *auto-correlation* functions of the input R_{uu} and output R_{yy} signals are defined as (cf. (3.15)):

$$\begin{aligned} R_{yu}(i) &= \frac{1}{m}\sum_{k=0}^{m-1} y(k)\, u^T(k-i) = E\left[y_k\, u_{k-i}^T\right], \\ R_{uu}(i) &= \frac{1}{m}\sum_{k=0}^{m-1} u(k)\, u^T(k-i) = E\left[u_k\, u_{k-i}^T\right], \\ R_{yy}(i) &= \frac{1}{m}\sum_{k=0}^{m-1} y(k)\, y^T(k-i) = E\left[y_k\, y_{k-i}^T\right]. \end{aligned} \tag{3.31}$$

One can show the validity of the following relation:

$$R_{yu}(i) = \sum_{k=0}^{\infty} Y_k R_{uu}(i-k) \tag{3.32}$$

which is similar in form to (3.7) and is therefore called the *correlation weighting sequence description* for linear systems.

The DFTs S_{yu}, S_{uu}, S_{yy} of the correlation functions (3.31) are called *cross-spectral* and *auto-spectral densities*, respectively. For SISO systems, the FRF can be expressed in terms of the densities as (cf. (3.28)):

$$G(z_k) = \frac{S_{yu}(k)}{S_{uu}(k)}, \tag{3.33}$$

Spectral methods frequently use the formula, which expresses the output spectrum in terms of the input one and the related FRFs, as:

$$S_{yy} = G(z_k) S_{uu} G^*(z_k). \tag{3.34}$$

The spectral density of the signal multiplied by an appropriate factor will produce the power carried by the signal and can be plotted in the *power spectral density* (spectral power distribution). Due to the finite representation of the DFT, a phenomenon called *spectral leakage* occurs, resulting in blurring the input frequencies in the power spectral density. To minimize the leakage, various *windowed signals* are used (e.g. Hamming, Bartlett, etc.), which means introducing some weight in Eq. (3.24).

Historically, the system identification methods started from the input-output deterministic approach to turn towards the output-only stochastic approach nowadays. The first step of classical input-output system identification is determination of the Markov parameters. Depending upon the data available, the methods of determining Markov parameters are divided into time domain or frequency domain methods. The data for extracting the Markov parameters may be either input and output signals in the time domain, measured with relatively simple equipment or FRFs in the frequency domain, measured directly with sophisticated spectrum analyzers. The initial approach of determining the Markov parameters is the time domain, off-line approach consisting of three steps:

1) performance of FFT (cf. (3.24)) for input and output signals measured, for some period of time (transformation of the problem into the frequency domain),
2) calculation of FRF (cf. (3.28)) on the basis of the FFTs from step 1),
3) performance of inverse FFT on the FRF (cf. (3.27)) to get the Markov parameters (back transformation of the problem into the time domain).

An alternative for the off-line approach is an on-line one for determining observed Markov parameters in the time domain. It is called recursive least squares method, which can be arranged in a special modular structure called Least Squares Lattice Filter (see Juang [41]), especially suitable for on-line application. The method is quite sophisticated, however the possibility of on-line use is its greatest asset. A review of up-to-date recursive identification methods is given in Ljung, Soderstrom [50].

If the FRF data is available directly, the frequency-domain state-space identification is possible thanks to expressing the FRF in the form of the *matrix fraction description*:

$$G(z_k) = \bar{Q}^{-1}(z_k)\,\bar{R}(z_k), \tag{3.35}$$

where

$$\begin{aligned} \bar{Q}(z_k) &= I_q + \bar{Q}_1 z_k^{-1} + ... + \bar{Q}_p z_k^{-p}, \\ \bar{R}(z_k) &= \bar{R}_0 + \bar{R}_1 z_k^{-1} + ... + \bar{R}_p z_k^{-p}. \end{aligned} \tag{3.36}$$

The multipliers $\bar{Q}_p,\ \bar{R}_p$ are nothing else but the observer Markov parameters (cf. (3.11)). Premultiplying Eq. (3.35) by $\bar{Q}(z_k)$ and rearranging, the following set of equations is obtained:

$$G(z_k) = -\bar{Q}_1 G(z_k)\, z_k^{-1} - ... - \bar{Q}_p G(z_k)\, z_k^{-p} + \bar{R}_0 + \bar{R}_1 z_k^{-1} + ... + \bar{R}_p z_k^{-p} \tag{3.37}$$

The linear algebraic set of Eqs. (3.37) can be solved for the observer Markov parameters $\left[-\bar{Q}_1\ ...\ -\bar{Q}_p\ \bar{R}_0\ ...\ \bar{R}_p\right]$ (the unknown vector) in the sense of the least squares norm.

For output-only stochastic system identification, the input is not included in the model description (cf. (13)), therefore the Markov parameters relating output with input cannot be determined. Only output is measured in this approach and it entirely replaces Markov parameters in mathematical calculations. However, thanks to the formal similarity of state-output covariances (3.17) to Markov parameters (3.8), the classical theory of system realization for time-invariant deterministic models can be applied.

A *realization* of the system is the computation of the state, input and output matrices A, B, C (sometimes the direct transmission matrix D, too) from the Markov parameters for the deterministic model and from the outputs for the stochastic model. Any system has an infinite number of realizations since there exist an infinite number of state-space representations that describe the same input-output relationship. A model with the smallest state-space dimensions is called *minimum realization.*

System realization starts with formation of the generalized αq x βr Hankel matrix (a matrix in which every antidiagonal stores the same element), composed of Markov parameters in the case of deterministic models:

$$H(k-1) = \begin{bmatrix} Y_k & Y_{k+1} & ... & Y_{k+\beta-1} \\ Y_{k+1} & Y_{k+2} & ... & Y_{k+\beta} \\ ... & ... & ... & ... \\ Y_{k+\alpha-1} & Y_{k+\alpha} & ... & Y_{k+\alpha+\beta-1} \end{bmatrix}. \tag{3.38}$$

Choosing $\alpha \geq n$ and $\beta \geq n$ (the order of the system), the Hankel matrix is of rank n. Using the definition (3.8) of Markov parameters, the Hankel matrix can be expressed as:

$$H(k-1) = O_\alpha A^{k-1} Q_\beta, \tag{3.39}$$

where:

$$O_\alpha = \begin{bmatrix} C \\ CA \\ CA^2 \\ \dots \\ CA^{\alpha-1} \end{bmatrix}, \qquad Q_\beta = \begin{bmatrix} B & AB & A^2B & \dots & A^{\beta-1}B \end{bmatrix}. \tag{3.40}$$

The block matrix O_α is the observability matrix and the block matrix Q_β is the controllability matrix. The *Eigensystem Realization Algorithm* (ERA) starts with SVD factorization of the Hankel matrix for $k = 1$, yielding:

$$H(0) = UTV^T, \tag{3.41}$$

where U and V are orthonormal matrices and T is a diagonal matrix with n non-zero components arranged in decreasing order.

Considering only the parts of matrices U, T, V corresponding to the n non-zero components of T and taking into account that $H(0) = O_\alpha Q_\beta$ (cf. (3.39)), the observability and controllability matrices can be expressed in terms of the SVD matrices (cf. (3.41)) as:

$$O_\alpha = U_n T_n^{1/2}, \qquad Q_\beta = T_n^{1/2} V_n^T \tag{3.42}$$

Thus from the shifted Hankel matrix formula $H(1) = O_\alpha A Q_\beta$ (cf. (3.39)), the state matrix A can be determined as:

$$A = T_n^{-1/2} U_n^T H(1) V_n T_n^{-1/2}, \tag{3.43}$$

Note that the input matrix B is the first r columns of the controllability matrix Q_β and the output matrix C is the first q rows of the observability matrix O_α (cf. (3.40)). The triplet A, B, C determined in this way is a minimum realization of the system.

Let us decide to use only some data stored in the observability matrix O_α. If we delete the first p rows or the last p rows from matrix O_α, we get the following reduced observability matrices (cf. (3.40), (3.42)):

$$O_\alpha^{\downarrow p} = \begin{bmatrix} C \\ CA \\ \dots \\ CA^{\alpha-2} \end{bmatrix} = U_n^{\downarrow p} T_n^{1/2}, \qquad O_\alpha^{\uparrow p} = \begin{bmatrix} CA \\ \dots \\ CA^{\alpha-2} \\ CA^{\alpha-1} \end{bmatrix} = U_n^{\uparrow p} T_n^{1/2}. \tag{3.44}$$

In order to speed up the computations of A, instead of using Eq. (3.43), another formula based on definitions (3.44) can be used:

$$A = T_n^{-1/2} \left[U_n^{\downarrow p}\right]^{+} U_n^{\uparrow p} T_n^{-1/2}, \tag{3.45}$$

where $^{+}$ means the pseudo-inverse[1)]. This procedure of computing A is called the *principal component algorithm.*

Considering only the parts of matrices U, T, V corresponding to the remaining zero components of T and using Eq. (3.41), two *canonical-form realizations* can be obtained:

$$\sum_{i=1}^{\alpha} U_{0i}^T C A^{i-1} = 0, \tag{3.46}$$

$$\sum_{i=1}^{\beta} A^{i-1} B V_{0i} = 0. \tag{3.47}$$

Equation (3.46) is the basis for the $q(\alpha-1)$-dimensional *observable canonical-form realization.* In general, this realization is not controllable and therefore it is not a minimum realization. In the observable realization, the matrix A is a function of U_0, B is a function of Markov parameters $Y_{\alpha-1}$, and C is identity. On the other hand, Eq. (3.47) is the basis for the $r(\beta-1)$-dimensional *controllable canonical-form realization.* In general, this realization is not observable and therefore it is not a minimum realization either. In the controllable realization, the matrix A is a function of V_0, B is identity, and C is a function of Markov parameters $Y_{\beta-1}$. From the computational point of view one should choose the canonical-form realization of smaller dimension.

The above considerations are basically the core of the system realization theory for deterministic, time-invariant systems (idealistic case) developed by KALMAN [43], HO, KALMAN [31]. With the advent of practical engineering applications of the system identification, the Kalman theory was extended by AKAIKE [2] and AOKI [1] for stochastic systems taking into account the state and measurement noise components (cf. (3.12)). Another practical requirement i.e. the use of ambient excitation (e.g. wind, traffic), pushed the system identification towards the output-only analysis. For observer identification, which is the case in practical situations, the related ARX model (cf. (3.10)) for deterministic input-output analysis is replaced by the ARMA model (cf. (3.19)) for stochastic output-only analysis.

The on-line, stochastic system identification can be best realized through the use of the *Kalman filter.* The role of the Kalman filter is to give an optimal

[1)] The Moore–Penrose pseudo-inverse of a non-square matrix A is defined as $A^{+} = A^T A^{-1} A^T$.

prediction of the next state vector x_{k+1} by using observations of the outputs up to time k and the related matrices. The predictions are denoted by $\hat{x}_{k+1}$. Given the known dynamics and output measurements, and assuming the initial values $\hat{x}_0 = 0$, $P_0 \equiv E\left(\hat{x}_0 \hat{x}_0^T\right) = 0$, the Kalman filter estimates $\hat{x}_{k+1}$ for a general, non-steady state are expressed by recursive formulas, describing the system, *Kalman filter gain matrix* and the state covariance matrix, respectively:

$$\begin{aligned} \hat{x}_{k+1} &= (A - KC)\,\hat{x}_k + (B - KD)\,u_k + Ky_k, \\ K_k &= \left(\Gamma - AP_kC^T\right)\left(\Lambda_0 - CP_kC^T\right)^{-1}, \\ P_{k+1} &= AP_kA^T + \left(\Gamma - AP_kC^T\right)\left(\Lambda_0 - CP_kC^T\right)^{-1}\left(\Gamma - AP_kC^T\right)^T. \end{aligned} \tag{3.48}$$

For output-only analysis, the term related to the input u_k disappears. There is a close relation between the state-space observer model (cf. (3.9), (3.6)) and the Kalman filter. The arbitrary matrix G for the state-space model can be chosen in such a way that the estimated state (the observer) approaches the theoretical one in the quickest possible way (producing minimum error). It turns out that the Kalman filter gain matrix $K = -G$ is the quickest observer. To be found in observer identification is the Kalman filter state sequence:

$$\hat{X} \equiv (\hat{x}_i\ \hat{x}_{i+1}\ ...\ \hat{x}_{i+j-1})\,. \tag{3.49}$$

In stochastic identification, the Hankel matrix is composed of output measurements only:

$$H \equiv \frac{1}{\sqrt{j}} \left[\begin{array}{cccc} y_0^{\mathrm{ref}} & y_1^{\mathrm{ref}} & ... & y_{j-1}^{\mathrm{ref}} \\ y_1^{\mathrm{ref}} & y_2^{\mathrm{ref}} & ... & y_j^{\mathrm{ref}} \\ ... & ... & ... & ... \\ y_{i-1}^{\mathrm{ref}} & y_i^{\mathrm{ref}} & ... & y_{i+j-2}^{ref} \\ \hline y_i & y_{i+1} & ... & y_{i+j-1} \\ y_{i+1} & y_{i+2} & ... & y_{i+j} \\ ... & ... & ... & ... \\ y_{2i-1} & y_{2i} & ... & y_{2i+j-2} \end{array}\right] \equiv \left(\frac{Y_{0|i-1}^{\mathrm{ref}}}{Y_{i|2i-1}}\right), \tag{3.50}$$

where the *reference sensor* block is allocated first. The reference sensors are crucial for system identification as they are placed in optimal locations able to represent all modes of vibration well. In practice, not all outputs are measured at the same time. For large structures all sensors are grouped into a number of

set-ups, arranged in such a way that the common part of neighbouring set-ups are just the reference sensors.

For the sake of computational efficiency for large systems, only a subspace of all outputs can be considered in stochastic system identification. A useful projection of the row space of non-reference outputs onto the row space of the reference outputs can be defined as:

$$\wp_i^{\mathrm{ref}} \equiv \frac{Y^{\mathrm{ref}}}{Y} = YY^{\mathrm{ref^T}} \left(Y^{\mathrm{ref}} Y^{\mathrm{ref}^T}\right)^+ Y^{\mathrm{ref}}. \tag{3.51}$$

The $H(0)$ matrix in deterministic approach is a product of the observability and controllability matrices (cf. (3.39)). The main theorem of the stochastic subspace identification formulated in Van Overschee, De Moor [84] states that similarly, the projection matrix $\wp^{\mathrm{ref}}$ can be expressed as the product of the observability matrix and the Kalman filter state sequence (cf. (3.40), (3.49)):

$$\wp_i^{\mathrm{ref}} = \begin{pmatrix} C \\ CA \\ CA^2 \\ ... \\ CA^{i-1} \end{pmatrix} (\hat{x}_i \; \hat{x}_{i+1} \; ... \; \hat{x}_{i+j-1}) = O_i \hat{X}_i \,. \tag{3.52}$$

For proving this, it is assumed that the number of measurement sessions approaches infinity ($j \to \infty$). The QR factorisation of the Hankel matrix (3.50) enables determination of the projection matrix $\wp^{\mathrm{ref}}$. Thus the Kalman filter state sequence can be calculated as:

$$\hat{X}_i = O_i^+ \wp_i^{\mathrm{ref}}. \tag{3.53}$$

Using the Kalman filter sequence, the stochastic output-only model (3.13) can be expressed as follows:

$$\begin{pmatrix} \hat{X}_{i+1} \\ Y_{i|i} \end{pmatrix} = \begin{pmatrix} A \\ C \end{pmatrix} \hat{X}_i + \begin{pmatrix} \rho_w \\ \rho_v \end{pmatrix} . \tag{3.54}$$

The noise residuals ρ_w, ρ_v are uncorrelated with $\hat{X}$, thus Eq. (3.54) can be solved in the least squares sense to give the state and output matrices:

$$\begin{pmatrix} A \\ C \end{pmatrix} = \begin{pmatrix} \hat{X}_{i+1} \\ Y_{i|i} \end{pmatrix} \hat{X}_i^+ \,. \tag{3.55}$$

Alternatively, the stochastic system identification may be performed quite analogously to the deterministic approach. The deterministic Hankel matrix of outputs is replaced by a stochastic Toeplitz matrix (a matrix in which every diagonal is composed of the same element) of output covariances and the formula to determine the state matrix A looks formally the same as the deterministic solution (3.43).

3.3. Modal analysis for an identified system

The modal decomposition of the state matrix A yields:

$$A = \Psi \Lambda_d \Psi^{-1} \,. \tag{3.56}$$

For general viscous damping, Λ_d is the diagonal matrix containing complex conjugate eigenvalues:

$$\Lambda_d = \begin{bmatrix} \Lambda & 0 \\ 0 & \Lambda^* \end{bmatrix} \quad \lambda_i,\ \lambda_i^* = -\xi_i \Omega_i \pm j\sqrt{1-\xi_i^2}\,\Omega_i \,, \tag{3.57}$$

where Ω_i is the i-th undamped natural frequency (rad/s) and ξ_i is the i-th damping ratio. Ψ is the matrix of eigenmodes:

$$\Psi = \begin{bmatrix} \Theta & \Theta^* \\ \Lambda\Theta & \Lambda^*\Theta^* \end{bmatrix} . \tag{3.58}$$

where Θ contain eigenvectors of the system for the case of non-proportional damping.

It can be shown that the (A, B, C) realization of the system can be transformed to the modal realization $(\Lambda_d,\, \Psi^{-1}B,\, C\Psi)$, identifying all modal parameters. The diagonal matrix Λ_d contains information about the eigenvalues. The matrix $\Psi^{-1}B$ defines the initial modal amplitudes (the so-called modal participation matrix) and the matrix $V = C\Psi$ – the mode shapes at sensor points. Note that the output-only methods will never provide information about modal amplitudes as the input matrix B is not identified in these methods.

By transformation of eigenvalues λ_i from the discrete to the continuous time, the system poles μ_i informing about the modal damping factors σ_i and the damped natural frequencies ω_i can be obtained using the following relation (cf. (3.5), (3.6)):

$$\mu_i = \sigma_i + j\omega_i = \frac{\ln \lambda_i}{\Delta t} \,. \tag{3.59}$$

The damping ratios ξ_i are then defined as:

$$\xi_i = \frac{-\sigma_i}{\sqrt{\omega_i^2 + \sigma_i^2}}. \tag{3.60}$$

The *Modal Assurance Criterion* (MAC) is frequently used to evaluate the correctness of the performed modal analysis:

$$\mathrm{MAC}_{p,\,p+1} = \frac{\left|v_p^H v_{p+1}\right|^2}{\left(v_p^H v_p\right)\left(v_{p+1}^H v_{p+1}\right)}, \tag{3.61}$$

where the superscript ^{H} means the complex conjugate transpose. The MAC coefficient varies between 0 and 1 and indicates whether there is a good correlation between the mode shapes for the identified model order p and the mode shapes for the one-order-higher model.

3.4. Methods of system identification

3.4.1. Frequency domain. The frequency domain methods use formula (3.34), relating the output and input spectra. The methods are often called *non-parametric* system identification methods since no physical parameters are directly identified through them.

3.4.1.1. Peak Picking. The earliest engineering application utilizing correlation and spectral analysis for determination of modal parameters is the *Peak Picking* (PP) method. As the name suggests it identifies eigenfrequencies of a system, occurring as peaks in a spectrum plot. With the strongly limiting assumption of low damping and well-separated eigenmodes (no close eigenfrequencies), the spectrum matrix in the PP method is an approximation of the observed mode shape, scaled by a constant factor. However, if the column corresponds to a degree of freedom (dof), which is not excited by a certain mode (the dof is in the nodal point of the mode), such mode cannot be identified. The PP method also estimates the damping ratios, but it is unfortunately not accurate. The principal assumptions of the PP method are often not fulfilled, which leads to misinterpretations. Namely, the PP method tends to identify operational deflection shapes instead of eigenmodes. This is incorrect because for closely-spaced modes, the deflection shape is a superposition of many modes.

3.4.1.2. Complex Mode Indication Function. The *Complex Mode Indication Function* (CMIF) method consists in performing the SVD of the spectrum matrix. A complete review of the CMIF with applications can be found in Allemang and Brown [3]. The observation behind the method is the fact that the

spectrum matrix (or equivalently the FRF) at a certain frequency is practically influenced by only a few eigenmodes. The number of the influencing eigenmodes determines the rank of the spectrum matrix. At resonance, for well-separated modes, only one mode has influence on the spectrum matrix, so its rank in SVD is one. If there exist several closely-spaced modes at the same frequency, this fact is reflected by the increased rank of the spectrum matrix. Thus the CMIF method can be considered as an SVD extension of the PP method, able to identify closely-spaced modes. By performing the SVD, the CMIF method decomposes the initial system into single-degree-of-freedom systems. EWINS [20] describes the relevant parameter estimation methods for such systems.

3.4.1.3. Maximum Likelihood. The most advanced spectrum-driven method is the *Maximum Likelihood* (ML) method, described in PINTELON, SCHOUKENS [68]. Unlike PP or CMIF, the ML estimates the parameters of a model by minimizing an error norm. It is thus a parametric method, solving a non-linear optimization problem in an iterative way. Despite the complexity of the approach due to the nonlinear formulation and the related high computational burden, the ML methods proved to be a robust tool, identifying modal parameters from large and noisy data. GUILLAUME *et al.* [29] shows that the ML can operate both with FRFs (input/output) or spectra (output-only), which makes the method versatile.

3.4.2. Time domain. The time domain methods follow strictly the system realization theory, i.e. identify the system by determining the triplet (A, B, C). Thanks to the modal analysis (see Sec. 3.3), all modal parameters can be identified by the time-domain methods. Therefore these methods are often referred to as *parametric* identification methods.

3.4.2.1. Determnistic approaches. Two most popular traditional (input-output) approaches for system identification in the time domain were developed simultaneously in the early 80's. One of them called the Eigensystem Realization Algorithm (ERA) was elaborated in the NASA Langley Research Center by JUANG [41] in the field of aerospace engineering. It uses the SVD factorization of the Markov parameter Hankel matrix (cf. (3.43), (3.45)) for finding the triplet (A, B, C) to identify the system. The deterministic ERA algorithm is based on both the input and output data. The canonical form of ERA (cf. (3.46), (3.47)) is very similar to another approach developed in the field of structural engineering called the *Polyreference Time Domain* (PTD), proposed by VOLD, RUSSELL [85]. The PTD was probably the most widespread time-domain input-output method of system identification applied in engineering before the advent of output-only methods. It encompasses two particular earlier versions called

Ibrahim Time Domain (see IBRAHIM, MIKULCIK [40]) and *Least Squares Complex Exponential* methods. The relations between the methods can be studied in ALLEMANG [4].

3.4.2.2. Stochastic approaches. Stochastic methods in the time domain began their fast development with proving the validity of formula (3.17), which enabled the use of classical realization theory in output-only considerations. Many stochastic methods can be derived from deterministic ones by substituting the impulse responses with output covariances. For instance, the *Instrumental Variable* (IV) method formally corresponds to the PTD method after performing such substitution. It identifies only the AR parameters, which results in a linear problem solved in the least squares sense, however the ARMA model structure is preserved. Advanced covariance-driven methods use only a subset of sensors (reference sensors) for system identification and are therefore called stochastic subspace methods. *Stochastic Subspace Identification-cov* (SSI-cov) uses the SVD factorization and corresponds to the deterministic ERA after substituting the impulse responses with output covariances. Depending upon the selection of weighting matrices for the Hankel matrix, the SSI-cov has two versions, namely *Balanced Realization* (BR – no weighting) and *Canonical Variate Analysis* (CVA – the weighting suited for the less excited modes to be better identified). Another subspace method, the *Stochastic Subspace Identification-dat* (SSI-dat) uses direct outputs (i.e the Hankel matrix (3.50)) instead of their covariances. This is an advantage over SSI-cov, because the data is not squared up (cf. (3.31)), which leads to higher numerical accuracy. SSI-dat determines the projection matrix (3.51), (3.52) and then employs the Kalman filter state sequence (3.49), (3.53) in a robust and stable least squares algorithm, identifying the system matrices (cf. (3.55)).

4. Low-frequency SHM methods

The low-frequency SHM methods operate in the non-ultrasonic range of frequencies i.e. below 20 kHz. They strongly rely on modal analysis. Therefore, they are often called vibration-based SHM. Most of the low-frequency SHM methods require a finite element model and many of them use the state-space form (cf. (12)) of equations of motion. Collecting experimental or in-situ measurements is necessary for obtaining the real response of a structure. Then, the FE model is fit to the measured data by means of the system identification procedures (see Sec. 3) in order to provide a reliable reference model. Damage identification is usually performed by subsequent updating of the model. The low-frequency methods often use quite sophisticated mathematical tools, but the effort of implementing them in numerical algorithms pays off when it comes

to results interpretation, which is rather straightforward. Since the methods look at a structure globally and examine a relatively broad inspection zone (possibly the whole structure, e.g. a bridge), they tend to identify damage of considerable extension, e.g. stiffness degradation or corrosion. Presently, low-frequency SHM methods are first of all applied in civil and mechanical engineering.

Several damage-sensitive parameters of the model have been tried to diagnose structural health. The damage identification procedure consists of updating the model parameters to best match the experimental response of a damaged structure. Friswell, Mottershead [23] describe difficulties in solving the resulting inverse problem inherent to model updating. Natural frequencies were used first to assess damage, but they are sensitive to structural defects of rather significant intensity. Williams, Messina [87] demonstrate the accuracy improvement of damage predictions through adding anti-resonance frequencies. Besides natural frequencies, FRFs are also used to identify damage. Yang, Lee [90] incorporate mode shapes to damage assessment algorithms as being more informative than frequencies. Multi-component objective functions, combining both frequencies and modal shapes, are constructed to improve the effectiveness of damage identification. As an alternative to mode shapes, Zimmermann [93] proposes to analyse the so-called *Ritz vectors*. Ho, Ewins [32] and Maeck, De Roeck [52] identified modal curvature as a parameter very sensitive to damage. Ceravolo *et al.* [11] and Modena *et al.* [57] use damping-related coefficients as parameters in damage detection. Zhang *et al.* [92] and Worden *et al.* [88] formulate SHM algorithms on the basis of modal strain energy, while Fritzen, Bohle [25] use modal kinetic energy. Friswell, Mottershead [24] and Fritzen, Bohle [26] consider model reduction as an important issue in model updating, which may make the analysis feasible or not for large engineering structures. Teughels [79] performed successful model reduction by the application of linear and parabolic damage functions. Basseville [5] deals with the problem of optimal placement of sensors, which plays an important role in effective identification.

Spina [75] uses standard deterministic realization theory to perform damage identification via the ARX family models. By discovering the formal similarity between the Markov parameters and output covariances (cf. (3.17)), a link was found between the stochastic methods and the classical realization theory. Since then, the influence of noise could be effectively analysed in SHM methods and moved them out of laboratories to in-field applications. De Stefano *et al.* [18] apply an ARMA model to perform modal analysis of a bridge girder with unknown random excitation. Bodeux, Golinval [8] use an ARMA model to test the performance of buildings during earthquakes. Two stochastic system identification methods (SSI-cov and SSI-dat) with an application to a wind-excited steel mast are described in Peeters, De Roeck [65]. A comparison of IV,

BR and CVA approaches (see 3.4.2.2) applied to three different industrial case studies (car – acoustics, aircraft, bridge) can be found in Hermans, Van Der Auweraer [30].

The Virtual Distortion Method (VDM), belonging to fast reanalysis methods (see Holnicki–Szulc, Gierlinski [33]), was applied to detecting damage via the solution of an inverse problem. Kołakowski *et al.* [47] demonstrate the effectiveness of VDM in the time domain whereas Świercz *et al.* [78] present its advantages in the frequency domain.

An interesting alternative to inverse methods in damage assessment, which are the vast majority of all approaches, is the Direct Stiffness Calculation proposed by Maeck [53]. The approach identifies structural stiffness directly as the quotient of the bending moment to beam curvature. With numerous sensors (accelerometers, optical fibres), both components of the quotient can be measured and theoretically no structural model is required.

5. NDT/E methods

Theoretical and practical aspects of the NDT/E methods can be these days quickly studied on the Internet, where the discussion panels addressing practical problems are of additional value (see e.g. http://www.ndt.net/ [38]). A selection of contemporary NDT/E methods is presented in Lewinska–Romicka [48].

5.1. Ultrasonic testing

Ultrasonic testing (UT) is the interrogation of materials using stress waves of the frequency higher than 20 kHz. This NDT/E method is capable of detecting both volume as well as planar (surface) defects. The essence of UT is to inject stress waves into the material to be examined and then monitor the transmitted or reflected signal. UT consists in scanning the surface of material with a probe. Piezoelectric probes are widely used to both induce and receive the stress waves. It is crucial that the stress waves propagate efficiently between the probe and the material, which requires a good acoustic coupling between them. This can be ensured by a coupling medium such as water, gels or greases. Electromagnetic acoustic transducers (EMAT) are also used to induce ultrasonic waves. No coupling medium is required for EMATs.

The type of defect detected by ultrasonic scanning depends upon the type of induced waves and their frequency. The theoretical detectable defect size is of the order of the wavelength. Hence, the high-frequency waves are more sensitive to defects. However, low frequency waves can penetrate the material to greater depths. So the choice of frequency is a compromise between sensitivity and penetration.

There are two major scanning techniques used in UT:

- Pulse echo technique – utilizes the phenomenon of reflection of ultrasonic waves. At a boundary of two different materials, a portion of the waves is reflected and the rest is transmitted. The portion reflected depends on the angle of incidence and the acoustic impedance (product of material density and wave velocity) of the two materials. The undeniable advantage of the pulse echo technique is the required access to only one surface of the tested object.
- Pitch and catch technique – requires access to two surfaces of the examined object. It is used for materials with high damping (pulse echo method not efficient) or for detecting near-surface defects. Two separate probes are placed on two opposite surfaces of the object. One of the probe is transmitting the signal whereas the other is receiving it. A material fault manifests itself in attenuation of energy of the transmitted wave. Amplitude analysis of the transmitted signal is the basis for quantification of defects.

The best detectability is observed for material discontinuities perpendicular to the injected ultrasonic beam.

The advantages of UT are:

- possibility of detecting both volume and planar defects,
- safe operation conditions for the staff,
- relatively low cost of testing.

The drawbacks of UT are:

- no detection of discontinuities positioned along the ultrasonic beam,
- difficulty in examining rough surfaces or irregular shapes,
- necessity of using the coupling medium for piezoelectric probes.

5.2. Radiographic testing

Radiographic testing (RT) is an NDT/E method detecting volume and planar defects. It is based on the measurement of the attenuation of electromagnetic radiation after having passed through a tested sample. The intensity of the incident radiation is reduced when passing through material, depending on its thickness, density, and the wavelength of radiation used. The penetrating power of the radiation increases with the growth of its energy (reduction of wavelength).

RT is performed using:

- the X radiation produced by Roentgen lamps (wavelength range $10^{-13} \div 10^{-9}$),
- the γ radiation produced by radioactive isotopes e.g. Se, Yb, Co (wavelength range $3 \times 10^{-14} \div 10^{-10}$).

The X or γ radiation triggers the photochemical phenomenon i.e. the decay of photo-emulsion. RT takes advantage of the fact that the difference in attenuation as the radiation passes through a defect, is sufficient to reveal the defect by exposition of a photographic film.

The advantages of RT are:

- possibility of testing objects of various, complicated shapes,
- good detectability of volume defects and planar defects along the radiation beam,
- no necessity of contact between the object and equipment.

The drawbacks of RT are:

- harmful operation conditions for the staff,
- poor detection of discontinuities positioned perpendicular to the radiation beam,
- relatively high cost of testing and bulky equipment.

5.3. Acoustic emission

Acoustic emission (AE) is the phenomenon of generating an elastic wave, in the range of ultrasound usually between 20 kHz and 1 MHz, by the rapid release of energy from the source within a material, which is deforming or fracturing under external loading. The energy comes from crack growth, dislocations or phase transformation. The elastic wave propagates through the solid to the surface, where it can be recorded by one or more piezoelectric sensors. In this way, information about the existence and location of possible sources is obtained.

AE differs from other NDT/E methods, which actively probe the structure. AE is a passive method, which listens to emissions from active defects and is very sensitive to defect activity when a structure is loaded beyond its service load in a test. AE analysis is a very useful method for the investigation of local damages in materials. One of the advantages compared to other NDT/E methods is the possibility to observe the damage the processes during the entire load history.

Operational environments are generally very noisy, and the AE signals are usually very weak. Hence, signal discrimination and noise reduction are very difficult, but extremely important for successful AE applications.

The advantages of AE are:

- no necessity of exciting the object or structure (service load is sufficient),
- ability of differentiating between stable and growing defects.

The drawbacks of AE are:

- difficulties in quantifying defects (commercial AE systems can only estimate qualitatively the extension of damage in the material and tell approximately how long the components will last),
- weak signals, therefore high sensitivity to environmental noise.

5.4. Magnetic particle inspection

Magnetic particle inspection (MT) consists in magnetizing an object made of a ferromagnetic material and examining the attenuation of magnetic field (flux), which reveals the places of defects. MT is suitable for planar defect detection only. By definition, MT is limited to ferromagnetic materials e.g. iron, ferromagnetic steel, cast iron, cast steel, iron-based alloys (Fe-Co, Fe-Ni).

The first stage of the MT method is magnetizing the tested material with a magnetic defectoscope, which is a device for inducing magnetic field in the material directly (flux defectoscopes) or indirectly by using electric current (current defectoscopes). Then, one of the following MT techniques may be used:

- magnetic powder inspection – determining the attenuation of magnetic flux using magnetic powders. As a result, a defectogram is obtained, in which the magnetic powder reflects the existing defects,
- magnetic transducer examination – measuring the attenuation of magnetic flux with magnetic transducers (induction transducers are the most widely used).

After the testing has been completed, the tested object has to be demagnetized.

The advantages of MT are:

- possibility of detecting planar (even narrow or shallow) discontinuities on/underneath the surface,
- low cost of equipment.

The drawbacks of MT are:

- harmful operation conditions for the staff (toxic powder components),
- limited applicability to ferromagnetic materials.

5.5. Eddy currents

Eddy current testing (ET) is a method for revealing surface defects in electrical conductors. The method is based on application of the Faraday's law of electromagnetic induction. The material is placed in varying magnetic field (produced by induction transducers), which induces varying electrical field in conductive materials. The amplitude and phase of the signal from the transducers

contains information about the geometry of material defects. The basic equipment is an ET defectoscope, providing power supply for induction transducers and analyzing their output signals.

The advantages of ET are:

- possibility of detecting planar discontinuities on/underneath the surface,
- applicability to coated (e.g. painted) surfaces.

The drawbacks of ET are:

- necessity of having extensive knowledge and experience by the operator,
- difficulties in interpretation of the signal affected by electric conductivity, magnetic penetrability and operating frequency.

5.6. Penetrant testing

Penetrant testing (PT) is a method of similar applicability to MT. It is used in materials in which MT or ET cannot be used. The essence of PT is application of a penetrant to a tested surface and subsequent application of a developer. The penetrant is usually an oil-base fluid and the developer is a dry powder or a powder suspended in water. The penetrant usually includes toxic additives e.g. fluorescent substance, therefore it is necessary to wash the examined object after the test.

The advantages of PT are:

- possibility of detecting planar discontinuities on/underneath the surface,
- possibility of detecting discontinuities in locations of rapid changes of cross-section (hardly possible by any other method),
- low cost of equipment.

The drawbacks of PT are:

- harmful operation conditions for the staff (toxic components),
- no applicability to coated (e.g. painted) surfaces,
- tedious (many stages) process of testing.

5.7. Optical holography

Optical holography (OH) is an imaging method, which records the amplitude and phase of light reflected from an object as an interferometric pattern on film. It thus allows reconstruction of the full 3D image of the object. In OH, the test sample is compared by interferometric measurement in two different stress states. Loading can be mechanical (e.g. vibration), thermal, etc. The resulting interference pattern contours the deformation undergone by the specimen between the two recordings. Surface as well as sub-surface defects reveal distortions in the otherwise uniform pattern. In addition, the characteristics of the compo-

nent, such as vibration modes, mechanical properties, residual stress, etc., can be identified.

The light used to illuminate the surface of the specimen must be coherent, which means that it must also be monochromatic, and the only practical source is a laser. With fast development of charge coupled devices (CCD) and digital image processing, OH offers tremendous flexibility and real-time visualization. Automated defect detection and analysis is possible thanks to computerized analysis of patterns.

OH is a dynamically developing method of many advantages, however one serious drawback for the moment – its huge cost.

6. High-frequency SHM methods

In contrast to the low-frequency SHM methods, the high-frequency ones operate in the ultrasonic range of frequencies i.e. above 20 kHz. The physical background for the methods is the phenomenon of elastic wave propagation in solid media. The high-frequency SHM methods heavily utilize experimental instrumentation for all stages of analysis i.e. generation and detection of elastic waves as well as results interpretation. Especially in the last aspect, they differ from the low-frequency methods, because the interpretation of results (often performed on-line) may not be easy and in principle requires much experience. Most of the high-frequency methods can be placed in the category of pattern recognition at the stage of damage identification. The methods look at a local part of a structure and therefore are focused on precise identification of a relatively small defect (e.g. a crack) in a narrow inspection zone (e.g. the connection of a wing to the fuselage in an aircraft). The most numerous applications of the high-frequency SHM methods can be observed in the aerospace and also mechanical engineering.

A comprehensive review of the high-frequency SHM methods is given in Staszewski [76]. Detailed mathematical description of the phenomenon of elastic wave propagation is provided in Rose [69]. Muravin [60] and Holroyd [35] characterize acoustic emission (AE), which utilizes structure-borne stress waves generated by internal material defects under external load applied. It is a passive method – no excitation, except for the service load, is required. All other methods require high-frequency excitation. Ultrasonic testing (UT) relies on the transmission and reflection of bulk waves and utilizes various phenomena (e.g., wave attenuation, scattering, reflections, mode conversions, energy partitioning) for damage detection. Presently, the most frequently used damage detection method based on guided (between two boundaries of a plate) ultrasonic waves is the Lamb wave inspection, in which the symmetric and antisymmetric modes of 0-th order (below 0.5 MHz) are most frequently analyzed. Koehler

et al. [46] report that the shear horizontal waves with non-dispersive 0-th order mode are also effective in damage detection. MALLET *et al.* [56] identify structural defects by examining Lamb wave attenuation and mode conversions. BIEMANS *et al.* [7] and BETZ *et al.* [6] propose *acousto-ultrasonics* as another high-frequency SHM method, combining elements of AE, UT and Lamb wave inspection. Acousto-ultrasonics uses an impulse excitation, resulting in propagation of a number of mixed wave modes, difficult to analyse.

Piezotransducers are the most versatile devices used in SHM. The use of piezoelectric transducers for generation and detection of elastic waves in plates was first studied in TRACY, CHANG [81] and continued by many researchers. GIURGIUTIU *et al.* [28] use piezoelectric-wafer active sensors (PWAS) for Lamb wave inspection. PARK and INMAN [63] propose an interesting SHM method utilizing the direct and converse piezoelectric effect simultaneously, to obtain an impedance signature and identify damage by monitoring its changes.

Finite Element Method has been extensively used for modelling wave propagation. However, there are problems of accuracy and efficiency in the high-frequency regime. To overcome those, OSTACHOWICZ and KRAWCZUK [62] developed the Spectral Element Method providing high accuracy for a relatively coarse mesh thanks to defining the element stiffness matrix in the frequency domain. Except for the Finite Element Method, STAVROULAKIS [77] successfully uses the Boundary Element Method for high-frequency crack identification as well. LOPEZ–DIEZ *et al.* [51] employ the Statistical Energy Analysis as a tool for incipient damage detection in a spacecraft panel.

7. Artificial intelligence in SHM

In damage identification algorithms, two major approaches can be met – either the formulation is analytical and relevant formula-based methods are used e.g. sensitivity-based optimization, or the formulation avoids analytical formulas and then artificial intelligence (also named soft-computing) methods are employed.

In fact, most structural health monitoring methods include Artificial Intelligence (AI). The SHM methods make use of wavelet transformations (see MALLAT [55], HOU *et al.* [36], HOU, HERA [37]) at the stage of signal processing. Artificial neural networks (see CHANG *et al.* [12], YANA *et al.* [89]), genetic algorithms (see CHOU, GHABOUSSI [13]), and statistical analysis (see FARRAR *et al.* [21], MONACO *et al.* [58], SOHN *et al.* [73]) are employed at the stage of damage identification. The AI tools are efficient, although they disregard the physical interpretation of the analysed phenomena. Case-based reasoning (CBR) proposed by MUJICA et al. [59] is a worth-noting method, combining wavelet transformation with Kohonen-like neural networks (self-organizing maps). However, the use

of knowledge-based approaches (such as CBR) has not been extensively exploited for damage detection yet.

8. SHM applications

One of the earliest applications of an SHM method was proposed by Cawley for a pipeline [10]. Presently, SHM methods have found applications mainly in civil, aerospace and mechanical engineering.

The Civil Engineering applications are first of all bridges (see Fritzen, Bohle [25], [26], Wang *et al.* [86], Peeters *et al.* [66], Maeck, De Roeck [54], Teughels, De Roeck [80]) masts (see Peeters, De Roeck [65]), buildings (see Skjaerbaek *et al.* [71], Spina [75]) and historical buildings (see Del Grosso *et al.* [15]).

The aerospace applications encompass aircrafts (see Hunt, Hebden [39]), spacecrafts (see Lopez–Diez [51]), satellites (see Kabashima *et al.* [42]) and composites (see Zak *et al.* [91], Yana *et al.* [89]).

The mechanical applications concentrate on rotating machinery (see Uhl *et al.* [83], Kasarda *et al.* [45]).

Other applications include e.g. water networks (see Holnicki–Szulc *et al.* [34]).

9. Final remarks

The Structural Health Monitoring has started as a research field applied exclusively in the branches of engineering requiring high-precision standards, e.g. in spacecraft or rotating machinery design. However the problem of aging infrastructure has accelerated the development of the Structural Health Monitoring in civil engineering as well, especially for bridges. An important step forward for the low-frequency methods was the formulation of the stochastic identification approach, enabling the consideration of measurement noise. Another essential achievement was the use of output-only data, which enabled to perform the *Operational Modal Analysis* (OMA), basing solely on ambient excitation. The newest trend seems to combine the advantages of the OMA and methods dealing with the defined excitation, resulting in the *Operational Modal Analysis with eXogenous inputs* (OMAX).

Piezoelectric transducers are commonly used in present-day SHM as they have the rare feature of being the actuator and sensor in one device. Optical fibres are gaining more and more attention, especially in bridge applications. Wireless sensor communication is important in permanent monitoring. Once the wireless sensors are mounted, the collected data can be transmitted long distance and analysed in a central unit far away from the monitored structure. In view of

the abundance of recorded data on large structures, data condensation techniques have become a critical issue in present-day SHM (see SPENCER *et al.* [74]).

Thus far, the majority of research done in SHM is limited to linear problems. Nonlinear analysis has been recently proposed by FARRAR *et al.* [22], DEMARIE *et al.* [16], UHL *et al.* [83]. Another important and relatively unexplored problem is the distinction of damage identification from the influence of environment on the structure (see PEETERS *et al.* [67], SOHN *et al.* [73]). DERAEMAEKER and PREUMONT [17] propose to use spatial filters to handle the problem.

It seems that the SHM is becoming more and more important as a research area of practical value to the engineering community, whose development is in fact driven by maintenance requirements of industrial structures. The SHM community is getting better organized. Many European research centres, dealing with the area, were co-operating in the Structural Assessment Monitoring and Control (SAMCO – (http://www.samco.org/) network, EU-financed in the years 2002–2006, and still continued as a voluntary initiative. Another SHM-related organization is the European Association for the Control of Structures (http://dipmec.unipv.it/eacs/).

ACKNOWLEDGMENTS

The paper has been financially supported by the national research project (EU structural funds) – DIADYN, *"Integrated dynamic system of risk assessment, diagnostics and control of structures and technological processes"* PBZ–KBN–105/T10/2003 (2005–2008). The author would like to express his gratitude to Prof. Wiesław Ostachowicz (IMP PAN, Gdańsk) for making available several proceedings from the DAMAS conferences and to Prof. Jan Holnicki–Szulc (IPPT PAN, Warsaw) for comments.

REFERENCES

1. AOKI M., *State space modelling of time series*, Springer, Berlin 1987.
2. AKAIKE H., *Stochastic theory of minimal realization*, IEEE Transactions on Automatic Control, **19**, 667–674, 1974.
3. ALLEMANG R.J., BROWN D.L., *A complete review of the complex mode indicator function (CMIF) with Applications*, Proc. of the International Conference on Noise and Vibration Engineering ISMA2006, Leuven, 3209–3246, 2006.
4. ALLEMANG R.J., *Vibrations: Experimental modal analysis*, Course Notes, 7-th edition, Structural Dynamics Research Laboratory, University of Cincinnati, OH, USA 1999. http://www.sdrl.uc.edu/course_info.html
5. BASSEVILLE M., *On sensor positioning for structural health monitoring,* Proc. of the 2nd European Workshop on Structural Health Monitoring, 7–9 July, Munich, Germany, 583–590, 2004.

6. Betz D., Thursby G., Culshaw B., Staszewski W.J., *Acousto-ultrasonic sensing using fiber Bragg gratings*, Smart Materials and Structures, **12**, 122–128, 2003.

7. Biemans C., Staszewski W.J., Boller C., Tomlinson G.R., *Crack detection in metallic structures using broadband excitation of acousto-ultrasonics*, Journal of Intelligent Material Systems and Structures, **12**, 8, 589–597, 2001.

8. Bodeux J.B., Golinval J.C., *ARMAV model technique for system identification and damage detection*, European COST F3 Conference on System Identification and Structural Health Monitoring, Madrid, Spain, 303–312, 2000.

9. Briley B.E., *An introduction to fiber optics system design*, Elsevier Science Publishers, 1990.

10. Cawley P., *Long range inspection of structures using low frequency ultrasound*, Structural Damage Assessment Using Advanced Signal Processing Procedures, Proc. of DAMAS '97, University of Sheffield, UK, 1–17, 1997.

11. Ceravolo R., De Stefano A., Molinari F., *Developments and comparisons on the definition of an instantaneous damping estimator for structures under natural excitation*, Proc. of the International Conference on Damage Assessment of Structures DAMAS'01, 25–28 June, Cardiff, UK, 231–240, 2001.

12. Chang C.C., Chang T.Y.P., Xu Y.G., Wang M.L., *Structural damage detection using an iterative neural network*, Journal of Intelligent Material Systems and Structures, **11**, 1, 32–42, 2000.

13. Chou J-H. and Ghaboussi J., *Genetic algorithm in structural damage detection*, Computers and Structures, **79**, 14, 1335–1353, 2001.

14. Cooley J.W., Tukey S.W., *An algorithm for the machine calculation of complex fourier series*, Math. Computers, **19**, 297–301, 1965.

15. Del Grosso A., Torre A., Rosa M., Lattuada B. G., *Application of SHM techniques in the restoration of historical buildings: The Royal Villa Monza*, Proc. of the 2nd European Workshop on Structural Health Monitoring, 7–9 July, Munich, Germany, 205–212, 2004.

16. Demarie G.V., Ceravolo R., De Stefano A., *Instantaneous identification of rigid bodies on nonlinear support based on volterra series representation*, Proc. of the 2nd International Conference on Structural Health Monitoring of Intelligent Infrastructure, 16–18 Nov, Shenzhen, China, 913–919, 2005.

17. Deraemaeker A., Preumont A., *Vibration-based damage detection using large array sensors and spatial filters*, Mechanical Systems and Signal Processing, **20**, 1615–1630, 2006.

18. De Stefano A., Sabia D., Sabia L., *Structural identification using armav models from noisy dynamic response under unknown random excitation*, Structural Damage Assessment Using Advanced Signal Processing Procedures, Proc. of DAMAS'97, University of Sheffield, UK, 419–428, 1997.

19. Eisenmann R.C. Sr., Eisenmann R.C. Jr., *Machinery malfunction diagnosis and correction*, Prentice Hall, Upper Saddle River, NJ 1998.

20. Ewins D.J., *Modal testing: theory, practice and application*, Taylor and Francis Group, 2-nd edition, 2001.

21. Farrar C.R., Duffey T.A., Doebling S.W., Nix D.A., *A statistical pattern recognition paradigm for vibration based structural health monitoring*, Structural Health Monitoring 2000, Stanford University, Palo Alto, California, 764–773, 1999.

22. Farrar C.R., Sohn H., Robertson A.N., *Application of nonlinear system identification to structural health monitoring*, Proc. of the 2-nd European Workshop on Structural Health Monitoring, 7–9 July, Munich, Germany, 59–67, 2004.

23. Friswell M.I., Mottershead J.E., *Inverse methods in structural health monitoring*, Proc. of the International Conference on Damage Assessment of Structures DAMAS'01, 25–28 June, Cardiff, UK, 201–210, 2001.

24. Friswell M.I., Mottershead J.E., *Finite Element Model Updating in Structural Dynamics*, Kluwer Academic Publishers, 1995.

25. Fritzen C.-P., Bohle K., *Damage identification using a modal kinetic energy criterion and "output-only" modal data–application to the Z24-brigde*, Proc. of the 2nd European Workshop on Structural Health Monitoring, 7–9 July, Munich, Germany, 185–194, 2004.

26. Fritzen C.-P., Bohle K., *Identification of damage in large scale structures by means of measured FRFs-procedure and application to the I40-highway bridge*, Proc. of the International Conference on Damage Assessment of Structures DAMAS'99, Dublin, Ireland, 310–319, 1999.

27. Gautschi G., *Piezoelectic sensorics*, Springer, 2002.

28. Giurgiutiu V., Blackshire J.L., Thomas D.T., Welter J.T., Yu L., *Recent advances in the use of piezoelectric wafer active sensors for structural health monitoring*, Proc. of the 2-nd European Workshop on Structural Health Monitoring, 7–9 July, Munich, Germany, 1083–1090, 2004.

29. Guillaume P., Hermans L., Van Der Auweraer H., *Maximum likelihood identification of modal parameters from operational data*, Proc. of IMAC 17 Conference on Structural Dynamics, Kissimmee, Florida, USA, 1887–1893, 1999.

30. Hermans L., Van Der Auweraer H., *Modal testing and analysis of structures under operational conditions: industrial applications*, *Mechanical Systems and Signal Processing*, **13**, 2, 193–216, 1999.

31. Ho B. L., Kalman R.E., *Effective construction of linear state-variable models from input/output data*, Regelungstechnik, **14**, 545–548, 1966.

32. Ho Y.K., Ewins D.J., *Numerical evaluation of the damage index*, Structural Health Monitoring 2000, Stanford University, Palo Alto, California, 995–1011, 1999.

33. Holnicki–Szulc J., Gierlinski J.T., *Structural analysis, design and control by the virtual distortion metod*, John Wiley and Sons, Chichester, UK 1995.

34. Holnicki–Szulc J., Kolakowski P., Nasher N., *Leakage detection in water networks*, Journal of Intelligent Material Systems and Structures, **16**, 3, 207–219, 2005.

35. Holroyd T., *The acoustic emission and ultrasonic monitoring handbook*, Coxmoor Publishing Company, Oxford 2001.

36. Hou Z., Noori M., Amand R.S., *Wavelet-based approach for structural damage detection*, Journal of Engineering Mechanics, **126**, 7, 677–683, 2000.

37. Hou Z. and Hera A., *A system identification technique using pseudo-wavelets*, Journal of Intelligent Material Systems and Structures, **12**, 10, 681–687, 2001.

38. http://www.ndt.net/

39. Hunt S.R., Hebden I.G., *Validation of the eurofighter tphoon structural health and usage monitoring system*, European COST F3 Conference on System Identification and Structural Health Monitoring, Madrid, Spain, 743–753, 2000.

40. Ibrahim S.R., Mikulcik E.C., *A Method for the direct identification of vibration parameters from the free response*, Shock and Vibration Bulletin, No. 47, Pt. 4, September, 183–198, 1977.

41. Juang J–N., *Applied system identification*, Prentice Hall PTR, Englewood Cliffs, NJ, USA 1994.

42. Kabashima S., Ozaki T., Takeda N., *Damage detection of satellite structures by optical fiber with small diameter*, Smart Structures and Materials 2000: Smart Structures and Integrated Systems, Proceedings of SPIE, Vol. 3, 985, Newport Beach, California, 343–351, 2000.

43. Kalman R.E., *Mathematical description of linear dynamical systems*, SIAM Journal on Control, **1**, 2, 152–192, 1963.

44. Kao K.C., Hockham G.A., *Dielectric fibre surface waveguides for optical frequencies*, Proc. of the IEEE, **113**, 1151–1158, 1966.

45. Kasarda M., Quinn D.D., Bash T., Mani G., Inman D., Kirk R.G., Sawicki J.T., *Magnetic bearings for non-destructive health monitoring of rotating machinery supported in conventional bearings*, Proc. of the International Conference on Damage Assessment of Structures DAMAS'05, 4–6 July, Gdańsk, Poland, 383–390, 2005.

46. Koehler B., Schubert F., Frankenstein B., *Numerical and experimental investigation of lamb wave excitation, propagation and detection for SHM*, proc. of the 2-nd European Workshop on Structural Health Monitoring, 7–9 July, Munich, Germany, 993–1000, 2004.

47. Kołakowski P., Zielinski T.G., Holnicki–Szulc J., *Damage identification by the dynamic virtual distortion method*, Journal of Intelligent Material Systems and Structures, **15**, 6, 479–493, 2004.

48. Lewinska–Romicka A., *Non-destructive testing. Fundamentals of defectoscopy* [in Polish], Wydawnictwa Naukowo–Techniczne, Warszawa 2001.

49. Ljung L , *System identification. Theory for the user*, 2-nd edition, Prentice Hall PTR, Upper Saddle River, NJ, USA 1999.

50. Ljung L., Soderstrom T., *Theory and practice of recursive identification*, The MIT Press, Cambridge, Massachusetts, USA 1983.

51. Lopez–Diez J., Torrealba M., Guemes A, Cuerno C., *Application of statistical energy analysis for damage detection in spacecraft structures*, Proc. of the International Conference on Damage Assessment of Structures DAMAS'05, 4–6 July, Gdańsk, Poland, 525–532, 2005.

52. Maeck J., De Roeck G., *Dynamic bending and torsion stiffness derivation from modal curvatures and torsion rates*, Journal of Sound and Vibration, **225**, 1, 153–170, 1999.

53. Maeck J., *Damage assessment of Civil Engineering structures by vibration monitoring*, PhD Thesis, Katholieke Universiteit Leuven, Belgium 2003.
http://www.kuleuven.be/bwm

54. Maeck J., De Roeck G., *Damage assessment using vibration analysis on the Z24 bridge*, Mechanical Systems and Signal Processing, **71**, 1, 133–142, 2003.

55. Mallat S., *A wavelet tour of signal processing*, San Diego, Academic Press 1999.

56. Mallet L., Lee B.C., Staszewski W.J., Scarpa F., *Structural health monitoring using scanning laser vibrometry: II. Lamb waves for damage detection*, Smart Materials and Structures, **13**, 2, 261–269, 2004.

57. Modena C., Sonda D., Zonta D., *Damage localization in reinforced concrete structures by using damping measurements*, Proc. of the International Conference on Damage Assessment of Structures DAMAS'99, Dublin, Ireland, 132–141, 1999.

58. Monaco E., Franco F., Lecce L., *Experimental and numerical activities on damage detection using magnetostrictive actuators and statistical analysis*, Journal of Intelligent Material Systems and Structures, **11**, 7, 567–578, 2000.

59. Mujica L.E., Vehi J., Rodellar J., Kolakowski P., *A hybrid approach of knowledge-based reasoning for structural assessment*, Smart Materials and Structures, **14**, 1554–1562, 2005.

60. Muravin G., *Inspection, diagnostics and monitoring of construction materials by the acoustic emission methods*, Minerva Press, London 2000.

61. Natke H., Cempel C., *Model-aided diagnosis of mechanical systems*, Springer, Berlin 1997.

62. Ostachowicz W., Krawczuk M., *Damage detection of structures using spectral element method*, [in:] Advances in Smart Technologies in Structural Engineering, J. Holnicki-Szulc, C. A. Mota Soares [Eds.], Springer, 69–88, 2003.

63. Park G. Inman D.J., *Impedance-based structural health monitoring*, [in:] Inman D., Farrar C., Lopes V. Jr., Steffen V. Jr. [Eds.] Damage Prognosis for Aerospace, Civil and Mechanical Systems, John Wiley and Sons 2005.

64. Peeters B., *System identification and damage detection in Civil Engineering*, PhD Thesis, Katholieke Universiteit Leuven, Belgium 2000. http://www.kuleuven.be/bwm

65. Peeters B., De Roeck G., *Reference-based stochastic subspace identification for output-only modal analysis*, Mechanical Systems and Signal Processing, **13**, 6, 855–878, 1999.

66. Peeters B., De Roeck G., Caetano E., Cunha A., *Dynamic study of the Vasco Da Gama bridge*, Proc. of the ISMA 2002, Conference on Noise and Vibration Engineering, September, Leuven, Belgium 2002.

67. Peeters B., Maeck J., De Roeck G., *Vibration-based damage detection in civil engineering: Excitation sources and temperature effects*, Smart Materials and Structures, **10**, 518–527, 2001.

68. Pintelon R., Schoukens J., *System identification: a frequency domain approach*, IEEE Press, New York 2001.

69. Rose J.L., *Ultrasonic waves in solid media*, Cambridge University Press, Cambridge 1999.

70. Rytter A., *Vibration based inspection of civil engineering structures*, PhD Thesis, Department of Building Technology and Structural Engineering, Aalborg University, Denmark 1993.

71. Skjaerbaek P.S., Kirkegaard P.H., Nielsen S.R.K., *Shaking table tests of reinforced concrete frames*, Structural Damage Assessment Using Advanced Signal, Processing Procedures, Proc. of DAMAS'97, University of Sheffield, UK, 441–450, 1997.

72. Sohn H., Farrar C.R., Hemez F.M., Shunk D.D., Stinemates D.W., Nadler B.R., *A Review of structural health monitoring literature: 1996–2001*, Los Alamos National Laboratory Report, LA-13976-MS, 2003.

73. Sohn H., Worden K., Farrar C.R., *Statistical damage classification under changing environmental and operational conditions*, Journal of Intelligent Material Systems and Structures, **13**, 9, 561–574, 2002.

74. Spencer B.F. Jr., Gao Y., Yang G., *Distributed computing strategy for damage monitoring employing smart sensors*, Proc. of the 2-nd International Conference on Structural Health Monitoring of Intelligent Infrastructure, 16–18 Nov, Shenzhen, China, 35–47, 2005.

75. Spina D., *Modal parameters identification of buildings using ARX models and seismic experimental data*, Proc. of the International Conference on Damage Assessment of Structures DAMAS'99, Dublin, Ireland, 328–334, 1999.

76. Staszewski W.J., *Structural health monitoring using guided ultrasonic waves*, [in:] Advances in Smart Technologies in Structural Engineering, J. Holnicki-Szulc, C.A. Mota Soares [Eds.], Springer, 117–162, 2003.

77. Stavroulakis G.E., *Inverse and crack identification problems in engineering mechanics*, Kluwer Academic Publishers, Dordrecht 2000.

78. Świercz A., Kolakowski P., Holnicki–Szulc J., *Damage identification utilizing harmonic excitation in static-like frequency-based inverse analysis*, Proc. of the 2nd International Conference on Structural Health Monitoring of Intelligent Infrastructure, 16-18 Nov, Shenzhen, China, 971–976, 2005.

79. Teughels A., *Inverse modelling of civil engineering structures on operational modal data*, PhD Thesis, Katholieke Universiteit Leuven, Belgium 2003.
http://www.kuleuven.be/bwm

80. Teughels A., De Roeck G., *Structural damage identification of the highway bridge Z24 by FE model updating*, Journal of Sound and Vibration, **278**, 3, 589–610, 2004.

81. Tracy M., Chang F.-K., *Identifying impacts in composite plates with piezoelectric strain sensors, Part I: Theory, Part II: Experiment*, Journal of Intelligent Material Systems and Structures, **9**, 920–937, 1998.

82. Udd E. [Ed.], *Fibre optic smart structures*, Wiley, New York 1995.

83. Uhl T., Barszcz T., Bednarz J., *Model based SHM – Rotating machinery application*, Proc. of the International Conference on Damage Assessment of Structures DAMAS'05, 4–6 July, Gdańsk, Poland, 459–466, 2005.

84. Van Overschee P., De Moor B., *Subspace identification for linear systems:Theory-implementation-applications*, Kluwer Academic Publishers, Dordrecht, the Netherlands 1996.

85. Vold H., Russell R., *Advanced Analysis Methods Improve Modal Test Results*, Sound and Vibration, March, 36–40, 1983.

86. Wang M.L., Xu F.L., Lloyd G.M., *A Systematic Numerical Analysis of the Damage Index Method Used for Bridge Diagnostics*, Smart Structures and Materials 2000: Smart

Systems for Bridges, Structures, and Highways, Proceedings of SPIE, **3**, 988, Newport Beach, California, 154–164, 2000.

87. Williams E.J., Messina A., *Applications of the multiple damage location assurance criterion*, Proc. of the International Conference on Damage Assessment of Structures DAMAS'99, Dublin, Ireland, 256–264, 1999.

88. Worden K., Manson G., Allman D., *An experimental appraisal of the strain energy damage location method*, Proc. of the International Conference on Damage Assessment of Structures DAMAS'01, 25–28 June, Cardiff, UK, 35–46, 2001.

89. Yana Y.J., Yam L.H., Jiang J.S., *Vibration-based damage detection for composite structures using wavelet transform and neural network*, Composite Structures, **60**, 403–412, 2003.

90. Yang S.M., Lee G.S., *Effects of modeling error on structure damage diagnosis by two-stage optimization*, Structural Health Monitoring 2000, Stanford University, Palo Alto, California, 871–880, 1999.

91. Zak A., Krawczuk M., Ostachowicz W., *Vibration of a laminated composite plate with closing delamination*, Proc. of the International Conference on Damage Assessment of Structures DAMAS'99, Dublin, Ireland, 17–26, 1999.

92. Zhang L., Quiong W., Link M., *A structural damage identification approach based on element modal strain energy*, Proc. of ISMA 23, Noise and *Vibration Engineering*, Leuven, Belgium 1998.

93. Zimmermann D.C., *Looking into the Crystal Ball: The continued need for multiple viewpoints in damage detection*, Proc. of the International Conference on Damage Assessment of Structures DAMAS'99, Dublin, Ireland, 76–90, 1999.

Received June 08, 2006; revised version December 11, 2006.