DOI 10.2478/ama-2020-0010

acta mechanica et automatica, vol.14 no.2 (2020)

TEMPERATURE-DEPENDENT FATIGUE CHARACTERISTICS OF P91 STEEL

Władysław EGNER*, Piotr SULICH, Stanisław MROZIŃSKI***, Halina EGNER***

*Institute of Applied Mechanics, Faculty of Mechanical Engineering, Cracow University of Technology,
31-864 Kraków Al. Jana Pawła II 37, Poland
**Huta Stalowa Wola S.A., ul. gen. Tadeusza Kasprzyckiego 8, 37-450 Stalowa Wola, Poland
***UTP University of Science and Technology, Faculty of Mechanical Engineering, 85-225 Bydgoszcz ul. Kordeckiego 20, Poland

wladyslaw.egner@pk.edu.pl, p.sulich@hsw.pl, stmpkm@utp.edu.pl, halina.egner@pk.edu.pl

received 17June 2019, revised 12 May 2020, accepted 15 May 2020

Abstract: In this paper, the experimental investigations, constitutive description and numerical modelling of low-cycle fatigue behaviour of P91 steel in non-isothermal conditions are presented. First, experimental tests are performed to recognise different aspects of material behaviour. Then, an appropriate constitutive model is developed within the framework of thermodynamics of irreversible processes with internal state variables. The model describes two phases of cyclic softening, related to plastic mechanisms. An important goal of the presented research is to include thermomechanical coupling in the constitutive modelling. Next, the model parameters are identified based on the available experimental data. Some parametric studies are presented. Finally, numerical simulations are performed, which indicate the significant influence of thermomechanical coupling on the response of the constitutive model in thermomechanical fatigue conditions.

Key words: low-cycle fatigue, thermomechanical coupling, constitutive modelling

1. INTRODUCTION

The interest in prediction of cyclic behaviour of materials and structures is growing in the industry. The 9%–12% Cr steels are being extensively used as structural materials for the components of reactors and power plants (Fournier et al., 2010). Research on this new steel grade dedicated for operation at elevated temperatures began in the USA at the end of the 1970s. As a result of the extensive research, the optimal chemical composition of steel was determined, which resulted in a good combination of mechanical and thermal properties (Kannan et al., 2013) such as high creep resistance and good ductility, high resistance to thermal fatigue, good weldability and the ability to easily make connections with other ferritic steel grades, good thermal conductivity and low thermal expansion coefficient, and good corrosion resistance and fracture toughness in water–steam and gas environments (in the presence of hydrogen) (Fournier et al., 2008; Nagesha et al., 2002; Sauzay et al., 2005; Kyaw et al., 2016)). P91 steel, which is one of the 9%–12% Cr steel family members, is perfectly suited for thick-walled pipes or forgings for the construction of boilers with extremely high operating requirements and steam generators, nuclear reactors and other responsible devices operating at temperatures up to 650°C (Duda, 2015). P91 steel has been also used for many years in power plant installations during renovations to replace austenitic steel and 21/4Cr–1Mo steel pipes. The benefits of its use are the reduction in the diameter and thickness of tubular elements, and thus their mass, which is the result of the higher creep strength of P91 steel.

During the process of frequent starting up and shutting down of the power plant units, high-temperature components are subjected to alternating cyclic changes in temperature and mechanical load, i.e. to thermomechanical fatigue. High-chromium steels subjected to high-temperature loads are subject to the evolution of the microstructure. This has been well documented in Nagesha et al. (2002) and Farragher (2014). This evolution manifests itself in the form of subgrain coarsening, where the boundaries between the martensitic laths and subgrain boundaries disappear (Sauzay et al., 2005), which leads to a coarser microstructure. The subgrain coarsening phenomenon depends on the applied plastic strain (Kyaw et al., 2016). There is also a decrease in dislocation density (Nagesha et al., 2002). As a consequence of these microstructural rearrangements, the 9%–12% Cr steels commonly exhibit cyclic softening (Kyaw et al., 2016; Farragher, 2014; Shankar et al., 2006; Sulich et al., 2017). This softening is observed regardless of the test temperature and the level of deformation.

The present investigation aims at a qualitative and quantitative description of the cyclic softening of P91 steel during low cycle fatigue at elevated and changing temperatures. Due to the variability in temperature, an important goal of the presented research is to include in the constitutive modelling the thermomechanical coupling. The effect of temperature (and its changes) on the phenomenon of cyclic softening of P91 steel is investigated experimentally, and an appropriate constitutive model is developed. The model parameters are determined here from strain-controlled, low-cycle fatigue tests at several isothermal test temperatures with the use of optimisation software. The results obtained at discrete temperatures are then interpolated into the temperature functions of material characteristics. Then the qualitative and quantitative impact of thermomechanical coupling on the response of the model is illustrated.

2. MATERIAL BEHAVIOUR

Experimental tests were performed on P91 steel specimens cut out of a boiler pipe (see Fig. 1; cf. Sulich et al., 2017). The chemical composition of steel was 0.127 C, 0.442 Si, 0.489 Mn, 0.017 P, 0.005 S, 8,82 Cr, 0.971 Mo, 0.307 Ni, 0.012 Al, 0.017 Co, 0.036 Cu, 0.074 Nb, 0.004 Ti and 0.201 V. Low-cycle fatigue tests were strain controlled, with a constant total strain amplitude ε_{ac} and a constant temperature θ in each test. Five levels of total strain amplitude (0.25%, 0.30%, 0.35%, 0.50% and 0.60%) and three levels of temperature (20°C, 400°C and 600°C) were applied. Experiments were performed on the testing machine Instron 8502 equipped with a heating chamber.

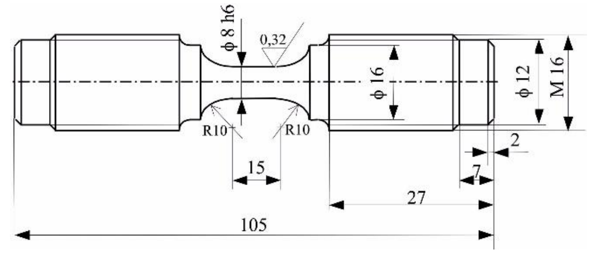


Fig. 1. Shape and dimensions of the tested specimen

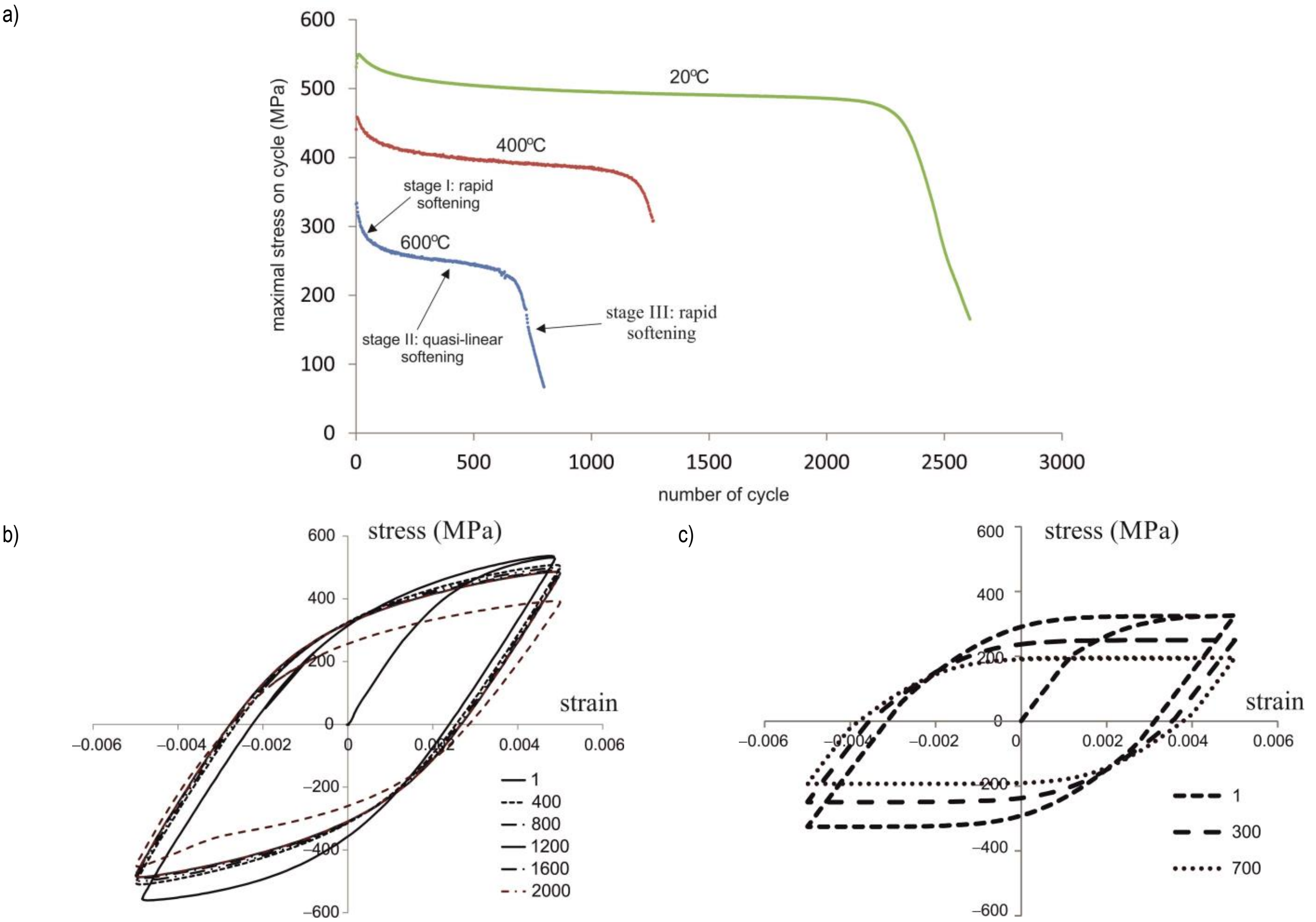


Fig. 2. (a) Maximum stress on cycle versus number of cycles for strain amplitude $\varepsilon_{ac} = 0.5\%$ and three test temperatures, 20°C, 400°C and 600°C; (b) chosen hysteresis loops for the test temperature 20°C; (c) chosen hysteresis loops for the test temperature 600°C

The results (see Fig. 2) indicate that except the first several cycles, the tested steel exhibits cyclic softening, regardless of the testing temperature and strain amplitude (Saad et al., 2011; Li et al., 2016; Mroziński and Golański, 2014; Zhao et al., 2017). This softening could be divided into three phases, which are: the rapid softening phase, followed by the slow quasi-linear softening and finally again fast softening caused by micro-damage development in the material that ultimately leads to failure of the tested sample (Mroziński, 2011). As mentioned before, the tested material generally exhibits an initial slight hardening during the very first cycles, the extent of it being dependent on the test temperature (decreasing with increasing temperature, see Fig. 2a, cf. also Nagesha et al., 2002; Kruml and Polák, 2001. It may be the effect of several mechanisms, such as the interaction between dislocations and solute atoms, mutual interaction among dislocations and formation of fine precipitates on dislocations during testing. After this very short consolidation phase, the maximum stress on the cycle continuously drops without a saturation period within the range of considered strain amplitudes. For higher plastic strain amplitudes, the tendency to saturation may be observed.

The existing literature suggests various mechanisms responsible for the cyclic softening/hardening of the high-strength steel group. In general, this phenomenon is explained by the modification of dislocation structure and density, carbide morphology, density and chemical composition (Zhang et al., 2002). In the study of Jones and Van Den Avyle (1980), it was found for 2.25Cr–1Mo steel that cyclic loading-induced dislocation shuttling accelerates the formation of Mo–C–Mo clusters, leading finally to the replacement of Mo–C pairs, which results in a continuous drop of the stress amplitude. Other mechanisms that have been proposed to explain the phenomenon of cyclic softening of this type of steel include stress reduction associated with the surface oxide film (Kim and Weertman, 1988), and change from the original lath structure to cells or equiaxed subgrains.

There are several parameters that influence the softening amplitude and kinetics (temperature, total or plastic strain amplitude, testing frequency, initial hardness, etc.). The present paper is focused on the description of the influence of variable temperature on the cyclic behaviour in non-isothermal conditions, while the dependence on other test conditions (Xie et al., 2018) and strain range memorisation effect (Zhou et al., 2018) are disregarded.

3. CONSTITUTIVE MODEL

3.1. Basic assumptions

Cyclic constitutive models have been widely developed for metallic materials and successfully used in academic research and industrial applications during the past decades. In constitutive modelling, the formalism of thermodynamics of irreversible processes with internal state variables and the local state method are often adopted (Skrzypek and Kuna-Ciskał, 2003; Egner and Ryś, 2017; Zhou et al., 2018; Xie et al., 2018). The current state of a material is determined by certain values of some independent variables called variables of state (observable or internal). The choice of internal variables depends on dissipative phenomena taking place in the material that need to be regarded in the theoretical model. The present model is based on the following assumptions: small strains, rate-independent plasticity, mixed isotropic/kinematic plastic hardening.

3.2. State variables

For elastic–plastic material exhibiting mixed hardening, the following set of state variables is defined:

$$\{V_\alpha\} = \left\{\varepsilon^{\mathrm{e}}_{ij}, \theta; \alpha^{(\mathrm{m})}_{ij}, r^{(\mathrm{n})}\right\} \tag{1}$$

where $\varepsilon^{\mathrm{e}}_{ij}$ are components of the reversible (elastic) strain tensor, θ is the absolute temperature in Kelvin degrees, $\alpha^{(\mathrm{m})}_{ij}$ correspond to kinematic plastic hardening, while $r^{(\mathrm{n})}$ are related to isotropic plastic hardening. Indexes m and n denote the number of different physical mechanisms (different dislocation populations) resulting in the occurrence and nature of hardening (Cailletaud and Saï, 1995; Egner and Egner, 2016). To reduce the number of model parameters, in the present analysis, the simplest case, $\mathrm{m} = \mathrm{n} = 1$, is considered.

3.3. Cyclic softening

Within the framework of hardening models, isotropic hardening is generally used to express the cyclic evolution of the material's mechanical strength with respect to the plastic flow (Chaboche, 2008). The dimension of the elasticity domain can be controlled with a law of the type:

$$R = Q(\theta)\left(1 - e^{-b(\theta)r}\right) \tag{2}$$

where $b(\theta)$ and $Q(\theta)$ are two coefficients that are material and temperature dependent. However, such a description leads to a typical saturation; therefore, it is not suitable for steels that soften continuously without a saturation period. To take into account the non-saturating cyclic softening observed experimentally in the case of P91 steel, the drag stress R can be divided into two parts, R_1 and R_2 (Zhang et al., 2002; Taleb and Cailletaud, 2010; Egner and Egner, 2014):

$$R_1 = Q(\theta)\left(1 - e^{-b(\theta)r}\right), R_2 = H^{\mathrm{R}}(\theta)r \tag{3}$$

The first part, R_1, corresponds to the strong softening typically taking place during the first hundred of cycles, while the second one, R_2, allows reflecting continuous softening. Due to the quasi-linear character of the second stage of softening, drag stress R_2 was here adopted in a linear form (Saad et al., 2013, Lu et al., 2015), where H^{R} reflects the slope of the second stage of cyclic softening (see Fig. 2a).

3.4. State equations

The thermodynamic forces conjugated to state variables (1) result from the assumed form of the state potential, which is here the Helmholtz free energy $\rho\psi$, decomposed into thermoelastic and thermoplastic parts (Sulich et al., 2017). In the present analysis, the following state equations define the thermodynamic forces:

$$\sigma_{ij} = \rho\frac{\partial\psi}{\partial\varepsilon^{\mathrm{e}}_{ij}} = E_{ijkl}(\theta)\left(\varepsilon_{kl} - \varepsilon^{\mathrm{p}}_{kl}\right) - \beta_{ij}(\theta)(\theta - \theta_0) \tag{4}$$

$$X_{ij} = \rho\frac{\partial\psi}{\partial\alpha_{ij}} = \frac{2}{3}C(\theta)\alpha_{ij} \tag{5}$$

$$R = \rho\frac{\partial\psi}{\partial r} = Q(\theta)\left(1 - \mathrm{e}^{-b(\theta)r}\right) + H^{\mathrm{R}}(\theta)r = R_1 + R_2 \tag{6}$$

Symbol σ_{ij} denotes the Cauchy stress tensor (thermodynamic force conjugated to elastic strain), X_{ij} denotes back stress (conjugated to plastic kinematic hardening variable α_{ij}) and drag stress R is conjugated to plastic isotropic hardening variable r. Additionally, $E_{ijkl}(\theta)$ denote the components of the elastic stiffness tensor, while $C(\theta)$, $Q(\theta)$, $b(\theta)$ and $H^R(\theta)$ stand for temperature-dependent material characteristics.

3.5. Dissipation potential and evolution equations

To establish the rate laws, the potential approach is applied here, based on the assumption of the existence of dissipation potential F, being closed, convex and scalar-valued function of the thermodynamic forces (4)–(6) and some other possible variables.

To allow for a nonlinear plastic hardening description (Frederick and Armstrong, 2007), the potential of plastic dissipation F is assumed to be not equal to plastic yield surface here (non-associated plasticity):

$$F\left(J_\alpha, Z_\beta, \theta\right) = f\left(J_\alpha, Z_\beta, \theta\right) + \frac{3\gamma(\theta)}{4C(\theta)}X_{ij}X_{ij} \tag{7}$$

$$f(J_\alpha, Z_\beta, \theta) = \sqrt{\tfrac{3}{2}(s_{ij} - X_{ij})(s_{ij} - X_{ij})} - \sigma_{\mathrm{y}}(\theta) - R \tag{8}$$

where $f(J_\alpha, Z_\beta, \theta)$ is the von Mises–type plastic yield function of thermodynamic conjugated forces J_α (deviatoric stress s_{ij}, back stress X_{ij} and drag stress R) and material characteristics Z_β (yield stress $\sigma_{\mathrm{y}}(\theta)$ and material parameters $C(\theta)$ and $\gamma(\theta)$). According to the generalised normality rule (Chaboche, 2008), the following rate equations are obtained:

$$\dot{\varepsilon}_{ij}^{\mathrm{p}} = \frac{3}{2}\dot{\lambda}^{\mathrm{p}} \frac{s_{ij} - X_{ij}}{\sqrt{\frac{3}{2}(s_{kl} - X_{kl})(s_{kl} - X_{kl})}} \tag{9}$$

$$\dot{\alpha}_{ij} = \dot{\varepsilon}_{ij}^{\mathrm{p}} - \frac{3\gamma}{2C} X_{ij}\dot{r} \tag{10}$$

$$\dot{r} = \dot{\lambda}^{\mathrm{p}} \tag{11}$$

4. THERMOMECHANICAL COUPLING

4.1. Rates of thermodynamic forces

The constitutive model involves a number of temperature-dependent model characteristics. According to Eqs (4)–(6), the rates of thermodynamic conjugate forces involve explicitly the rate of temperature and are expressed as:

$$\dot{\sigma}_{ij} = E_{ijkl}(\theta)(\dot{\varepsilon}_{kl} - \dot{\varepsilon}_{kl}^{\mathrm{p}}) - \underbrace{\left[-\frac{\partial E_{ijkl}}{\partial \theta}(\varepsilon_{kl} - \varepsilon_{kl}^{\mathrm{p}}) + \frac{\partial \beta_{ij}}{\partial \theta}(\theta - \theta_0) + \beta_{ij}\right]}_{\text{coupling term } P_{ij}^{\theta}} \dot{\theta} \tag{12}$$

$$\dot{X}_{ij} = \frac{2}{3}C\dot{\alpha}_i + \underbrace{\frac{2}{3}\frac{\mathrm{d}C}{\mathrm{d}\theta}\alpha_{ij}}_{\text{coupling term } C_{ij}^{\theta}} \dot{\theta} \tag{13}$$

$$\dot{R} = (Qb\mathrm{e}^{-br} + H^{\mathrm{R}})\dot{r} + \underbrace{\left[\frac{\mathrm{d}Q}{\mathrm{d}\theta}(1 - \mathrm{e}^{-br}) + \left(Q\frac{\mathrm{d}b}{\mathrm{d}\theta}\mathrm{e}^{-br} + \frac{\mathrm{d}H^{\mathrm{R}}}{\mathrm{d}\theta}\right) r\right]}_{\text{coupling term } Q^{\theta}} \dot{\theta} \tag{14}$$

In the presence of thermoplastic coupling, the rate of a considered thermodynamic force depends, therefore, not only on the rate of internal variable conjugated to it, but also on the rate of temperature. In the following sections, the significance of coupling terms appearing in the kinetic equations (12)–(14)

$$P_{ij}^{\theta} = -\frac{\partial E_{ijkl}}{\partial \theta}(\varepsilon_{kl} - \varepsilon_{kl}^{\mathrm{p}}) + \frac{\partial \beta_{ij}}{\partial \theta}(\theta - \theta_0) + \beta_{ij} \tag{15}$$

$$C_{ij}^{\theta} = \frac{2}{3}\frac{\mathrm{d}C}{\mathrm{d}\theta}\alpha_{ij} \tag{16}$$

$$Q^{\theta} = \frac{\mathrm{d}Q}{\mathrm{d}\theta}(1 - \mathrm{e}^{-br}) + \left(Q\frac{\mathrm{d}b}{\mathrm{d}\theta}\mathrm{e}^{-br} + \frac{\mathrm{d}H^{\mathrm{R}}}{\mathrm{d}\theta}\right) r \tag{17}$$

will be investigated for the case of thermomechanical cyclic loading.

4.2. Heat balance equation

To determine the temperature distribution within the body, the heat balance equation is used, which is derived from the first law of thermodynamics by substituting into it the internal energy density together with Fourier's law. The heat balance equation takes the following form (Egner and Egner, 2016) (c_ε^{θ} is the specific heat capacity at a constant strain, q_i stands for the heat flux and r^{θ} is the distributed heat source per unit volume):

$$\rho c_\varepsilon^{\theta}\dot{\theta} = -q_{i,i} + r^{\theta} + \underbrace{\sigma_{ij}\dot{\varepsilon}_{ij}^{\mathrm{p}} - R\dot{r} - X_{ij}\dot{\alpha}_{ij}}_{\text{mechanical dissipation}} + \underbrace{\theta\left[\frac{\partial R}{\partial \theta}\dot{r} + \frac{\partial X_{ij}}{\partial \theta}\dot{\alpha}_{ij} - P_{ij}^{\theta}\left(\dot{\varepsilon}_{ij} - \dot{\varepsilon}_{ij}^{\mathrm{p}}\right)\right]}_{\text{thermo-mechanical coupling term}} \tag{18}$$

According to Eq. (18), determination of the temperature distribution within the body is coupled not only to the total strain rate $\dot{\varepsilon}_{ij}$, but also to the inelastic rates, $\dot{\varepsilon}_{ij}^{\mathrm{p}}$, and fluxes $\dot{\alpha}_{ij}$ and $\dot{r}$.

4.3. Loading/unloading conditions

Since the thermodynamic conjugate forces are functions of state variables, the consistency relation for the development of dissipative phenomena takes the following form:

$$\dot{f} = \frac{\partial f}{\partial \sigma_{ij}}\dot{\sigma}_{ij} + \frac{\partial f}{\partial X_{ij}}\dot{X}_{ij} + \frac{\partial f}{\partial R}\dot{R} + \frac{\partial f}{\partial \theta}\dot{\theta} = 0 \tag{19}$$

where the term $\frac{\partial f}{\partial \theta}$ explicitly includes the dependence of the initial yield stress σ_{y} on temperature.

Using the chain rule and Eqs (9)–(11) and (12)–(14), the consistency condition becomes:

$$\dot{f} = \frac{\partial f}{\partial \sigma_{ij}}\dot{\sigma}_{ij} - \dot{\lambda}^{\mathrm{p}}H - \dot{\theta}S = 0 \tag{20}$$

In the above equation, H is a generalised hardening modulus,

$$H = \frac{2}{3}\frac{\partial f}{\partial \sigma_{ij}}\left(C\frac{\partial f}{\partial \sigma_{ij}} - \frac{3}{2}\gamma X_{ij}\right) - \frac{\partial f}{\partial R}(Qb\mathrm{e}^{-br} + H^{\mathrm{R}}) \tag{21}$$

while S reflects the sensitivity of the yield surface on temperature changes (Egner, 2012):

$$S\begin{cases} > 0 \Rightarrow \text{yield surf. contracts with increasing } \theta \\ = 0 \Rightarrow \text{yield surf. remains constant with increasing } \theta \\ < 0 \Rightarrow \text{yield surf. expands with increasing } \theta \end{cases} \tag{22}$$

where

$$S = \frac{\partial f}{\partial \sigma_{ij}}C_{ij}^{\theta} - \frac{\partial f}{\partial R}Q^{\theta} - \frac{\partial f}{\partial \theta} \tag{23}$$

Concerning the majority of experiments, the physical meaning has the case when the yield surface contracts with increasing temperature.

Expression (20) determines the consistency multiplier:

$$\dot{\lambda}^{\mathrm{p}}(\theta) = \frac{1}{w}\left[\frac{\partial f}{\partial \sigma_{ij}}E_{ijkl}\dot{\varepsilon}_{kl} - \underbrace{(\frac{\partial f}{\partial \sigma_{ij}}P_{ij}^{\theta} + S)}_{\text{coupling term}}\dot{\theta}\right] \tag{24}$$

where $w = \frac{\partial f}{\partial \sigma_{ij}}E_{ijkl}\frac{\partial f}{\partial \sigma_{kl}} + H > 0.$

Temperature influences the mechanical properties of conventional engineering materials. Generally speaking, degradation of mechanical properties is observed (referred to as thermal softening), accompanied by increasing values of thermal properties. The numerical examples presented below will investigate the qualitative and quantitative influence of thermomechanical coupling in the set of Eqs (12)–(18) and (24).

5. RESULTS

5.1. Numerical implementation

The constitutive model is implemented into numerical subroutines by the use of the fully implicit backward Euler scheme (which is always stable and very accurate) and the Newton–Raphson method. The iterative solution procedure is defined as

$$\Delta\mathbf{Y}^{(k+1)}=\Delta\mathbf{Y}^{(k)}-\left[\mathbf{J}^{(k)}\right]^{-1}\mathbf{R}^{\mathrm{res}(k)} \tag{25}$$

where $\Delta\mathbf{Y}=\{\Delta V_{\alpha}\}$ is the vector containing the increments of the unknowns, $[\mathbf{J}]=\partial\mathbf{R}^{\mathrm{res}}/\partial\Delta\mathbf{Y}$ is the Jacobian matrix and $\mathbf{R}^{\mathrm{res}}=\{R_{\Delta Y_i}\}$ is a residual vector containing the components $R_{\Delta Y_i}=\Delta Y_i-\Delta\widehat{Y}_i$, where ΔY_i is a variable while $\Delta\widehat{Y}_i$ denotes the function resulting from the evolution rule for i-th variable Y_i. In the case of uniaxial tension/compression, it is

$$\Delta\mathbf{Y}=\begin{bmatrix}\Delta\boldsymbol{\varepsilon}\\ \Delta\theta\\ \Delta\lambda\\ \Delta\boldsymbol{\alpha}\end{bmatrix},\ \mathbf{R}^{\mathrm{res}}=\begin{bmatrix}\mathbf{R}_{\Delta\varepsilon}\\ R_{\Delta\theta}\\ R_{\Delta\lambda}\\ \mathbf{R}_{\Delta\alpha}\end{bmatrix},\ \mathbf{J}=\begin{bmatrix}\frac{\partial\mathbf{R}_{\Delta\varepsilon}}{\partial\boldsymbol{\varepsilon}} & \frac{\partial\mathbf{R}_{\Delta\varepsilon}}{\partial\theta} & \frac{\partial\mathbf{R}_{\Delta\varepsilon}}{\partial\boldsymbol{\alpha}} & \frac{\partial\mathbf{R}_{\Delta\varepsilon}}{\partial\lambda}\\ \frac{\partial R_{\Delta\theta}}{\partial\boldsymbol{\varepsilon}} & \frac{\partial R_{\Delta\theta}}{\partial\theta} & \frac{\partial R_{\Delta\theta}}{\partial\boldsymbol{\alpha}} & \frac{\partial R_{\Delta\theta}}{\partial\lambda}\\ \frac{\partial R_{\Delta\lambda}}{\partial\boldsymbol{\varepsilon}} & \frac{\partial R_{\Delta\lambda}}{\partial\theta} & \frac{\partial R_{\Delta\lambda}}{\partial\boldsymbol{\alpha}} & \frac{\partial R_{\Delta\lambda}}{\partial\lambda}\\ \frac{\partial\mathbf{R}_{\Delta\alpha}}{\partial\boldsymbol{\varepsilon}} & \frac{\partial\mathbf{R}_{\Delta\alpha}}{\partial\theta} & \frac{\partial\mathbf{R}_{\Delta\alpha}}{\partial\boldsymbol{\alpha}} & \frac{\partial\mathbf{R}_{\Delta\alpha}}{\partial\lambda}\end{bmatrix} \tag{26}$$

The condition $\mathbf{R}^{\mathrm{res}}(\Delta\mathbf{Y})=\mathbf{0}$ defines the solution; therefore, the iteration procedure is stopped when the norm of $\mathbf{R}^{\mathrm{res}}$ is sufficiently small.

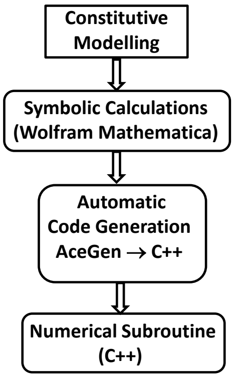


Fig. 3. Scheme of numerical implementation

To solve the problem numerically, the classical concept of elastic predictor/plastic corrector was applied. Wolfram Mathematica 11 (Wolfram Mathematica 11.2, 2017) software was used to generate the expressions for components of the Jacobian matrix. Numerical subroutines were automatically generated in C++ language by the use of AceGen software (Korelc, 2016) (see Fig. 3).

5.2. Material data

Experimental observations indicate that P91 steel is not a stable material during fatigue, i.e. the steel microstructure can be modified by the thermomechanical cycle. This is the case, however, only when the temperature reaches or exceeds the tempering temperature during some period of time (Zhang et al., 2002). In general, there are two temperature ranges (separated by tempering temperature) in which the fatigue behaviour of P91 steel is different. Above the tempering temperature, pronounced effects of ageing are observed, additionally enhanced by the fatigue loading. In this temperature range, changes in the mechanical properties of steel depend not only on the current temperature, but also on its history. Identification of model parameters should, therefore, take place jointly for all temperatures in a given load programme (Cailletaud et al., 2000). To do this, experimental thermomechanical fatigue tests should be performed and used to identify the temperature history–dependent parameters.

On the other hand, ageing at temperatures below tempering remains almost constant (which is the case considered here) and mechanical properties depend only on the current temperature, and not on its history. This means that there is no need to include all temperatures together in the identification procedure. Instead of that, it is possible to carry out several isothermal tests at different temperatures, identify material parameters in each test temperature and finally to take into account the influence of temperature on the material parameters by the use of interpolation techniques with polynomial or spline functions. In the present analysis, P91 steel was tested in the temperature range between 20°C and 600°C, while the tempering temperature of this steel is 730°C–760°C.

The identification of model parameters for each isothermal test was performed with the application of the SIMULIA-Isight package (Dassault, 2016). Two key components from the point of view of identification were used: “Data Matching”, which offers the ability to calculate different error measures of two or more data sets (e.g. stresses obtained from the experiment and stress calculated numerically) and “Optimisation” component, which allows for various methods of finding the minimum of a multivariable function (see Fig. 4). The following vector of normalised material parameters $\overline{P}_i\in\langle-1,1\rangle$ was searched (the temperature dependence of Poisson’s ratio ν and thermal expansion coefficient α^{θ} was disregarded):

$$\overline{P}_i=\frac{2P_i-(U_i+L_i)}{U_i-L_i},\ P_i\in\langle L_i,U_i\rangle,\ \{P_i\}=\{\sigma_{\mathrm{y}},E,\gamma,C,b,Q,H^{\mathrm{R}}\} \tag{27}$$

Parameters P_i are bounded between their respective lower bounds L_i and upper bounds U_i.

In the present analysis, the following error measure was used:

$$F_{\mathrm{obj}}(\overline{\mathbf{P}})=w_1\sum_{k=1}^{m}\left|\sigma_k(\overline{\mathbf{P}})-\sigma_k^{\exp}\right|+ \\ +mw_2\mathrm{Max}\left|\sigma_j(\overline{\mathbf{P}})-\sigma_j^{\exp}\right|_{j=1,\dots,m} \tag{28}$$

where $\sigma_k^{\exp}$ denote the experimental stress data and $\sigma_k(\overline{\mathbf{P}})$ are the stress data calculated numerically by the use of current values of model parameters $\overline{P}_i$. The objective function is a weighted sum (weights w_1 and w_2) of two components: the sum of absolute differences between the experimental and numerical data (stress) and the absolute maximal differences between the experimental data and the numerically simulated data. Such function allows to effectively reduce the maximum error ($w_2\gg w_1$) or to adjust very well the data in the entire tested range ($w_1\gg w_2$).

The search time depends strongly on the selection of the starting point. To shorten the “distance” to the optimal solution, the following procedure was used to select the appropriate starting point as close as possible to the optimal solution:

- Elastic parameters (initial yield stress σ_{y} and elastic modulus E) were determined manually, taking into account the initial part of the first hysteresis loop.

- Next, considering the entire first hysteresis loop, the approximate values of the kinematic hardening parameters were determined (γ and C) assuming that the isotropic hardening in the first loop can be disregarded ($b = 0$, $Q = 0$ and $H^{\mathrm{R}} = 0$).
- Then, parameters related to isotropic hardening were searched based on several selected hysteresis loops. As a result, approximate values of all material parameters were determined as the starting point for optimisation.
- Finally, the identification of all material parameters $(\sigma_{\mathrm{y}}, E, \gamma, C, b, Q, H^{\mathrm{R}})$ was carried out again, but in a substantially reduced range around the starting point.

The results of the identification procedure described above are presented in Table 1. They ensure the best match of numerical results to experimental results for the selected objective function $F_{\mathrm{obj}}(\overline{\mathbf{P}})$ and the assumed constitutive model.

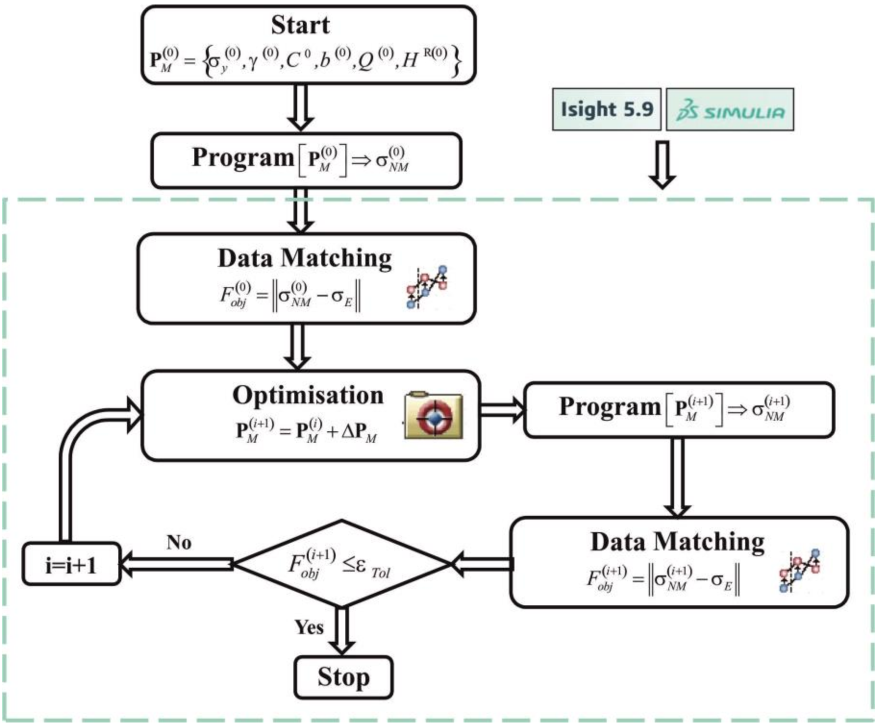


Fig. 4. Algorithm of parameter identification

Table 1. Material parameters ($\varepsilon_{\mathrm{ac}} = 0.5\%$)

	20°C	**400°C**	**600°C**
σ_{y} (MPa)	371	281	199
E (MPa)	210464	184782	173791
ν	0.3	0.3	0.3
γ	634	658.7	663.5
C (MPa)	125806	123726	99433
b	0.68	1.17	3.27
Q (MPa)	−45.5	−47.7	−66.3
H^{R} (MPa)	−1.27	−2.61	−5.46

The influence of temperature on the material parameters was then determined by interpolation techniques with spline functions according to the following scheme:

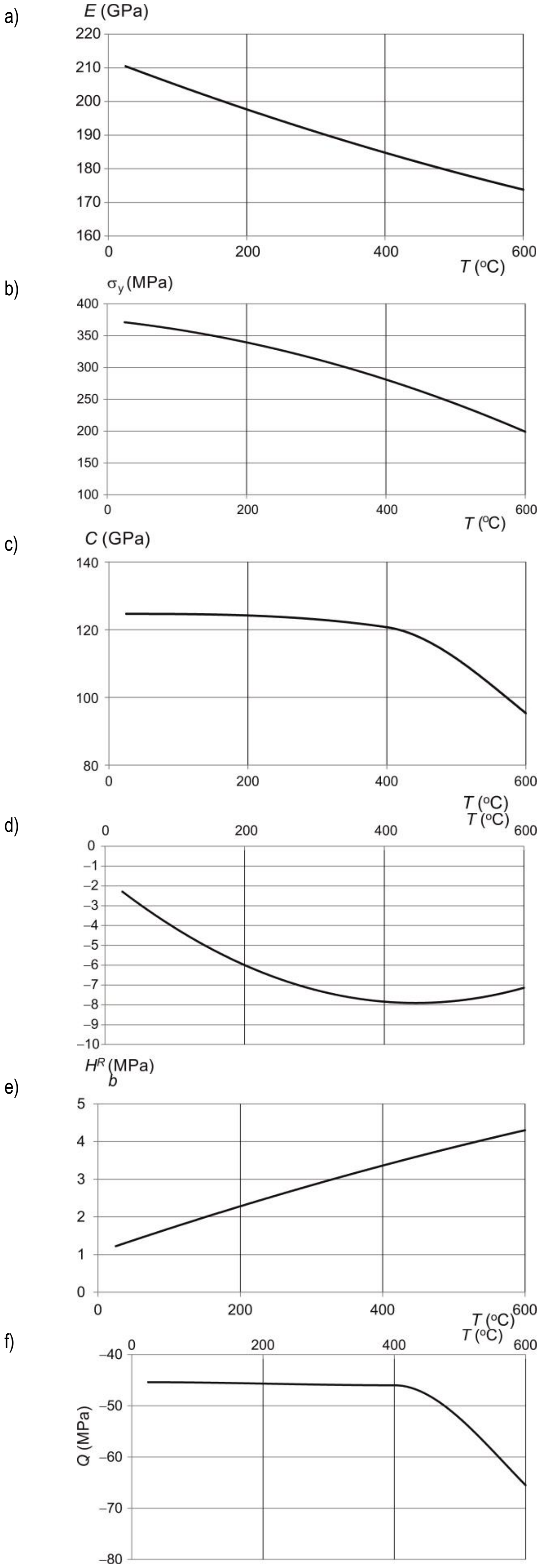
$$g(\theta) = a_i(\theta - \theta_i)^3 + b_i(\theta - \theta_i)^2 + c_i(\theta - \theta_i) + d_i \quad (29)$$

for $\theta_i \le \theta \le \theta_{i+1}$

To illustrate the significance of temperature influence, the temperature-dependent material characteristics are plotted in Fig. 5.

5.3. Validation

The general procedure of constitutive modelling, numerical implementation and parameter identification was validated with the use of available experimental data. A very good agreement was obtained between numerical simulations of cyclic softening behaviour and experimental tests (see Fig. 6).



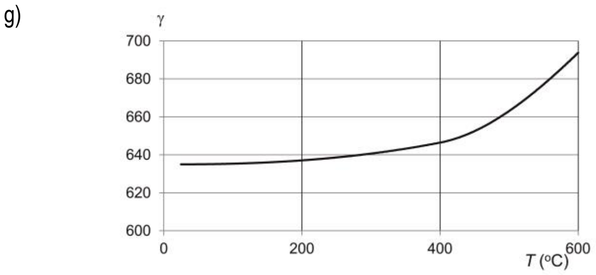


Fig. 5. Function interpolations of material characteristics (T denotes temperature expressed in Celsius degrees)

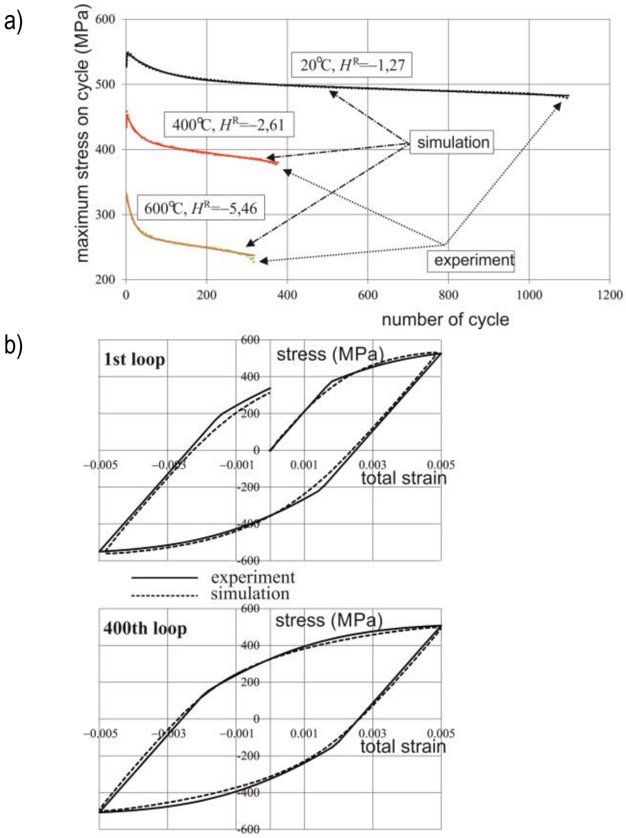


Fig. 6. Validation tests: (a) maximum stress on cycle versus the number of cycles; (b) chosen hysteresis loops at a temperature of 20°C

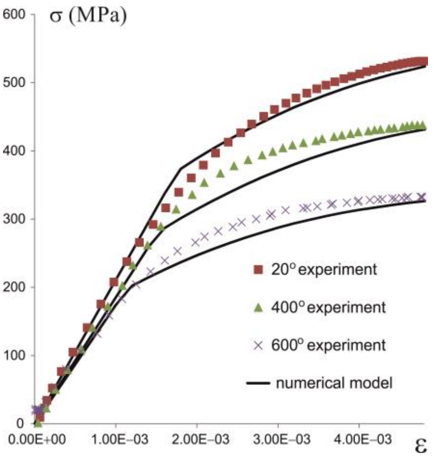


Fig. 7. Influence of linear part of isotropic softening (isothermal fatigue test at a temperature of 20°C)

The identification procedure allowed to fit very well into the softening curve, while the monotonic stress–strain curve (and individual hysteresis loops) exhibited some differences in the shape between the experiment and numerical simulations (see Fig. 7). This is most likely due to a single back stress used in the presented analysis. Increasing the number of terms in the back stress decomposition would allow the expression of a more extensive strain domain and a better description of the soft transition between elasticity and the onset of plastic flow. On the other hand, it would increase the number of model parameters, and this, in turn, would worsen some other aspects of the model behaviour. The increasing number of model parameters usually leads to multiple local minima of the error function, resulting in a non-unique set of optimal model parameters. Secondly, with a large number of parameters, the model becomes sensitive to their changes, so that a small change in a parameter may cause significant discrepancies in the simulation results.

In the subsequent sections, parametric studies were performed to show the influence of subsequent effects regarded in the constitutive description on the model response.

5.4. Isothermal tests

The influence of a linear term in Eq. (6) is shown in Fig. 8. Without the linear term in Eq. (6) ($H^{\mathrm{R}} = 0$), the shape of the maximum stress evolution cannot be reflected accurately.

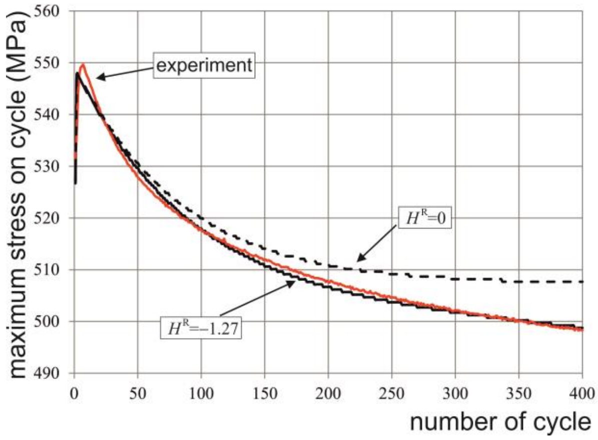


Fig. 8. Influence of linear part of isotropic softening (isothermal fatigue test at a temperature of 20°C)

It is possible to fit well into the first stage of cyclic softening, but then the second stage is highly erroneous and unrealistic (the rate of the maximum stress evolution drops to zero, cf. e.g. Saad et al., 2013). If both first and second stages are subjected to parameter optimisation described in the previous section, the error is unacceptably large and the simulated curve substantially differs from the experimental one for both, regarding here stages of material softening. The linear term R_2 included in Eq. (6) allows obtaining a very good compatibility between experimental and numerical results in the whole softening range considered.

5.5. Non-isothermal simulations

Several non-isothermal fatigue tests were subjected to numerical simulation according to the strain and temperature control scheme presented in Figs 9 and 10. Reversed strain cycles are considered in Fig. 9, with temperature changes taking place at the maximum and minimum total strain (in-phase and out-of-phase configurations are considered, cf. Fig. 9a,b). During the strain change, the temperature is kept constant. To evaluate the influence of thermomechanical coupling terms in Eqs (12)–(18), the following two cases are compared:

(case 1) Temperature rate terms in kinetic Eqs (12)–(14) are disregarded and the influence of temperature changes is accounted for only by updating the material characteristics;

(case 2) All temperature rate-dependent terms are included (full thermomechanical coupling).

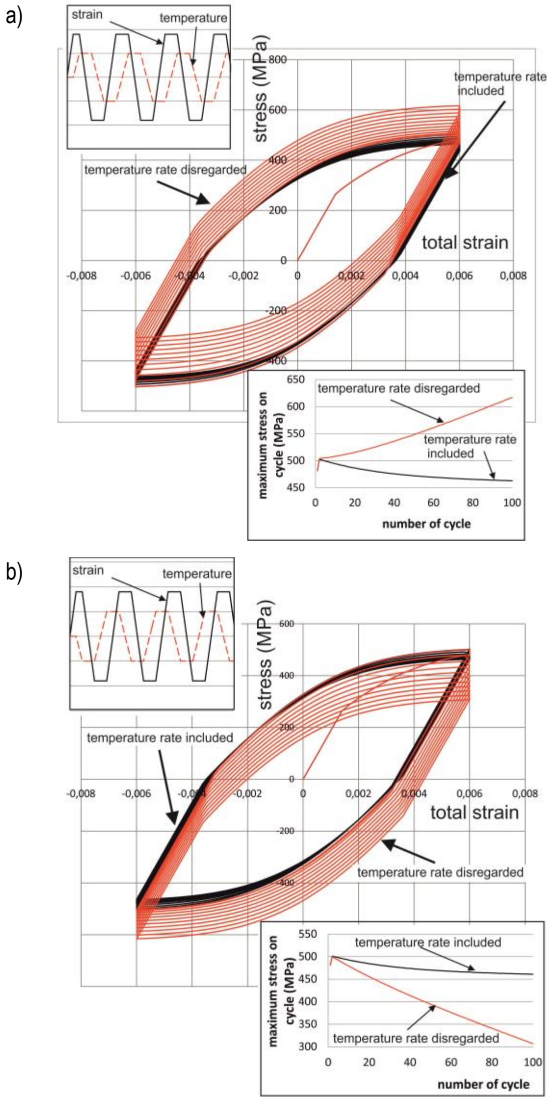


Fig. 9. Non-isothermal fatigue simulations with cyclic temperature change

The stress–strain loops for both cases are presented in Fig. 9a,b. Qualitatively, different results are obtained: without temperature rate terms (case 1), the response exhibits shift of hysteresis loops along the stress axis, while including the temperature rate terms (case 2) allows for preserving stable behaviour. Such an effect was already indicated by Chaboche (2008), and also by Egner and Egner (2016), Egner and Egner (2014) and

Besson et al. (2009). Due to this unreasonable shift, the simulated maximum stress on cycle contains a substantial error when the temperature rate-dependent terms in Eqs (12)–(14) are disregarded.

The significance of temperature rate influence depends on the specific case of the strain/temperature control scheme. For example, if cyclic strain change is accompanied by monotonic temperature increase or decrease, the difference between case (1) and case (2) is not so pronounced (see Fig. 10).

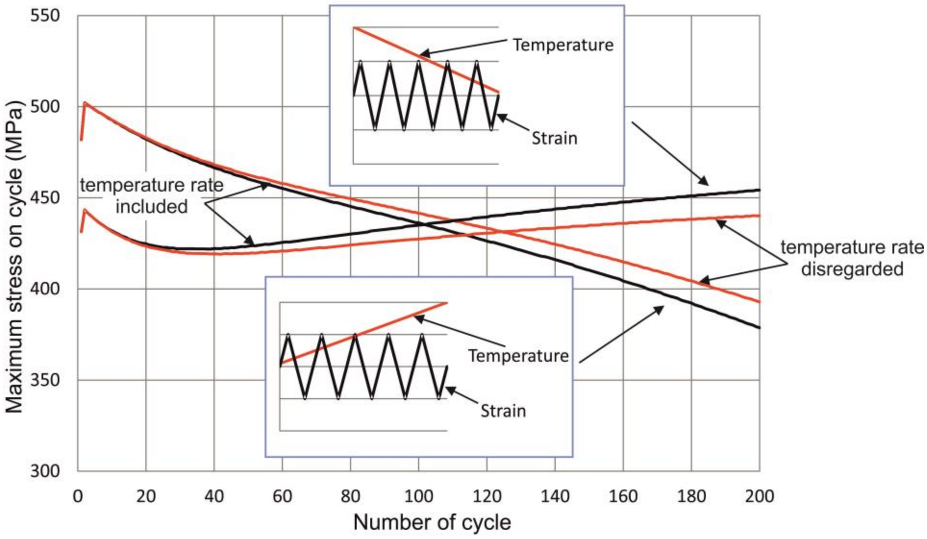


Fig. 10. Non-isothermal fatigue simulations with monotonic temperature change

6. CONCLUSIONS

This article is concerned with the description of the cyclic softening of P91 steel under non-isothermal conditions. Particular attention is paid to accounting for thermomechanical coupling in constitutive modelling and to examining the effect of this coupling on the model's response. The analysis presented in this work consists of six stages: (1) experimental testing of the material at several test temperatures, (2) constitutive modelling concerning variable temperature influence and the phenomenon of material cyclic softening, (3) numerical implementation of the mathematical model, (4) identification of model parameters at different test temperatures and then interpolating the results to obtain temperature-dependent material characteristics, (5) validating the analysis by comparing the experimental and numerical results and (6) investigating the effect of thermomechanical coupling on the model response, depending on the thermomechanical loading scheme.

To investigate the effect of full thermomechanical coupling on the model response, the classical constitutive model of Armstrong and Frederick (Frederick and Armstrong, 2007) was extended to take account of the rate of temperature explicitly. To properly reflect the first two stages of material cyclic softening, in the description of the isotropic plastic hardening, the drag stress was decomposed into two components – one of them responsible for the first stage of rapid softening, while the other for the quasi-linear softening stage. The influence of temperature on material parameters was determined using interpolation techniques with polynomial spline functions. The numerical test results revealed a significant qualitative and quantitative effect of temperature. This means that the coupling between temperature and dissipative phenomena occurring in the material can have a significant impact on the response of the constitutive model. Disregarding the temperature rate in the model equations can, therefore, lead to erroneous results, while the amount of error depends on the thermomechanical cyclic load configuration.

The constitutive description presented in this work was based on a model of the Armstrong and Frederick type. Such a model is not capable to reflect all the physical mechanisms able to produce material nonlinearities on a macro scale (such as unilateral damage, plastic strain range memorisation effect, etc.). Nevertheless, even this relatively simple model involves seven parameters that have to be identified for several temperatures of fatigue tests. This is a laborious and time-consuming task, and each additional parameter significantly increases the size of the problem. Therefore, in the presented analysis, a relatively simple model is considered, but the described procedure is general and can also be applied to more comprehensive modelling (Saanouni and Devalan, 2012).

REFERENCES

1. **Besson J., Cailletaud G., Chaboche S.** (2009), Non-linear mechanics of materials, Springer.
2. **Cailletaud G., Depoid C., Massinon D., Nicouleau-Bourles E.** (2000), Elastoviscoplasticity with ageing in aluminium alloys, [In:] *Continuum Thermomechanics: The Art and Science of Modelling Material Behaviour (Paul Germain's Anniversary Volume)*, Solid Mechanics and Its Applications, Kluwer Academic Publishers, 75–86.
3. **Cailletaud G., Saï K.** (1995), Study of plastic/viscoplastic models with various inelastic mechanisms, *International Journal of Plasticity*, 11, 991–1005.
4. **Chaboche J.L.** (2008), A review of some plasticity and viscoplasticity constitutive theories, *International Journal of Plasticity*, 24(10), 1642-1693.
5. **Dassault.** (2016), SIMULIA Abaqus Extended Products, Abaqus 6.14 - AP Isight 5.9. http://www.3ds.com/products/simulia
6. **Duda P.** (2015), Solution of an inverse axisymmetric heat conduction problem in complicated geometry. International Journal of Heat and Mass Transfer, 419–428.

7. **Egner H.** (2012), On the full coupling between thermo-plasticity and thermo-damage in thermodynamic modeling of dissipative materials, International Journal of Solids and Structures, 49(2), 279–288.
8. **Egner H., Egner W.** (2014), Modeling of a tempered martensitic hot work tool steel behavior in the presence of thermo-viscoplastic coupling, *International Journal of Plasticity*, 57, 77–91.
9. **Egner H., Ryś M.** (2017), Total energy equivalence in constitutive modeling of multidissipative materials, I*nternational Journal of Damage Mechanics*, 26(3),, 417–446.
10. **Egner W., Egner H.** (2016), Thermo-mechanical coupling in constitutive modeling of dissipative materials, *International Journal of Solids and Structures*, 91, 78–88.
11. **Farragher T.F.** (2014), Thermomechanical Analysis of P91 Power Plant Components, PhD Thesis, National University of Ireland Galway, https://aran.library.nuigalway.ie/xmlui/handle/10379/4161.
12. **Fournier B., Salvi M., Dalle F., De Carlan Y., Caës C., Sauzay M., Pineau A.** (2010), Lifetime prediction of 9-12%Cr martensitic steels subjected to creep-fatigue at high temperature, *International Journal of Fatigue*, 32(6), 971–978.
13. **Fournier B., Sauzay M., Caës C., Noblecourt M., Mottot M., Bougault A., Rabeau V., Man J., Gillia O., Lemoine P., Pineau A.** (2008), Creep-fatigue-oxidation interactions in a 9Cr-1Mo martensitic steel. Part III: Lifetime prediction, *International Journal of Fatigue*, 30(10–11), 1797–1812.
14. **Frederick C.O., Armstrong P.J.** (2007), A mathematical representation of the multiaxial Bauschinger effect, *Materials at High Temperatures*, 24(1), 1–26.
15. **Jones W.B., Van Den Avyle J.A.** (1980), Substructure and strengthening mechanisms in 2.25 Cr-1 Mo steel at elevated temperatures, *Metallurgical Transactions*, A 11A, 1275–1286.
16. **Kannan R., Sankar V., Sandhya R., Mathew M.D.** (2013), Comparative evaluation of the low cycle fatigue behaviours of P91 and P92 steels, *Procedia Engineering*, 55, 149–153.
17. **Kim S., Weertman J.R.** (1988), Investigation of microstructural changes in a ferritic steel caused by high temperature fatigue, *Metallurgical Transactions*, A 19A, 999–1007.
18. **Kruml T., Polák J.** (2001), Fatigue softening of X10CrAl24 ferritic steel, *Materials Science and Engineering*, A 319, 564–568.
19. **Kyaw S.T., Rouse J.P., Lu J., Sun W.** (2016), Determination of material parameters for a unified viscoplasticity-damage model for a P91 power plant steel, *International Journal of Mechanical Sciences*, 115–116, 168–179.
20. **Li M., Barrett R.A., Scully S., Harrison N.M., Leen S.B., O'Donoghue P.E.** (2016), Cyclic plasticity of welded P91 material for simple and complex power plant connections, *International Journal of Fatigue*, 87, 391–404.
21. **Lu J., Sun W., Becker A., Saad A.A.** (2015), Simulation of the fatigue behaviour of a power plant steel with a damage variable,. *International Journal of Mechanical Sciences*, 100, 145–157.
22. **Mroziński S.** (2011), The influence of loading program on the course of fatigue damage cumulation, **Journal of Theoretical and Applied Mechanics**, 49(1), 83–95.
23. **Mroziński S., Golański G.** (2014), Influence of temperature change on fatigue properties of P91 steel, *Materials Research Innovations*, 18 (2), 504–508.
24. **Nagesha A., Valsan M., Kannan R., Bhanu Sankara Rao K., Mannan S.L.** (2002), Influence of temperature on the low cycle fatigue behaviour of a modified 9Cr-1Mo ferritic steel, *International Journal of Fatigue*, 1285–1293.
25. **Korelc J.** (2016), AceGen 6.824 Windows.
26. **Saad A.A., Hyde T.H., Sun W., Hyde C.J., Tanner D.W.J.** (2013), Characterization of viscoplasticity behaviour of P91 and P92 power plant steels, *International Journal of Pressure Vessels and Piping*, 111–112, 246–252.
27. **Saad A.A., Sun W., Hyde T.H., Tanner D.W.J.** (2011), Cyclic softening behaviour of a P91 steel under low cycle fatigue at high temperature, *Procedia Engineering*, 1103–1108.
28. **Saanouni K., Devalan P.** (2012), *Damage Mechanics in Metal Forming: Advanced Modeling and Numerical Simulation*, Wiley.
29. **Sauzay M., Brillet H., Monnet I., Mottot M., Barcelo F., Fournier B., Pineau A.** (2005), Cyclically induced softening due to low-angle boundary annihilation in a martensitic steel, *Materials Science and Engineering A*, 400–401(1-2), 241–244.
30. **Shankar V., Valsan M., Rao K.B.S., Kannan R., Mannan S.L., Pathak S.D**. (2006), Low cycle fatigue behavior and microstructural evolution of modified 9Cr-1Mo ferritic steel, *Materials Science and Engineering* A, 437, 413–422.
31. **Skrzypek J.J., Kuna-Ciskał H.** (2003), Anisotropic Elastic-Brittle-Damage and Fracture Models Based on Irreversible Thermodynamics, *Lecture Notes in Applied and Computational Mechanics*, 9, 143–184.
32. **Sulich P., Egner W., Mroziński S., Egner H.** (2017), Modeling of cyclic thermo-elastic-plastic behaviour of P91 steel, Journal of Theoretical and Applied Mechanics, 55(2), 595–606.
33. **Taleb L., Cailletaud G.** (2010), An updated version of the multimechanism model for cyclic plasticity, *International Journal of Plasticity*, 26(6), 859–874.
34. Wolfram Mathematica 11.2. (2017), http://www.wolfram.com/mathematica/new-in-11/
35. **Xie X., Jiang W., Chen J., Zhang X., Tu S.-T.** (2019), Cyclic hardening/softening behavior of 316L stainless steel at elevated temperature including strain-rate and strain-range dependence: Experimental and damage-coupled constitutive modeling, *International Journal of Plasticity*, 114, 196-214.
36. **Zhang Z., Delagnes D., Bernhart G.** (2002), Anisothermal cyclic plasticity modelling of martensitic steels, *International Journal of Fatigue*, 24(6), 635–648.
37. **Zhao P., Xuan F. Z., Wu D.L.** (2017), Cyclic softening behaviors of modified 9–12%Cr steel under different loading modes: Role of loading levels, *International Journal of Mechanical Sciences*, 131–132, 278–285.
38. **Zhou J., Sun Z., Kanouté P., Retraint, D.** (2018), Experimental analysis and constitutive modelling of cyclic behaviour of 316L steels including hardening/softening and strain range memory effect in LCF regime, *International Journal of Plasticity*, 107, 54–78.

Acknowledgements: This work was supported by the National Science Centre of Poland through the Grant No. 2017/25/B/ST8/02256.

DOI 10.2478/ama-2020-0011 acta mechanica et automatica, vol.14 no.2 (2020)

MODELLING OF IMPULSE LOAD INFLUENCE ON THE STRESS STATE OF FOAM MATERIALS WITH POSITIVE AND NEGATIVE POISSON'S RATIO

Heorhiy SULYM[*], **Olena MIKULICH**[**], **Vasyl' SHVABYUK**[**]

[*]Faculty of Mechanical Engineering, Department of Mechanics and Applied Computer Science,
Bialystok University of Technology, ul. Wiejska 45C, 15-351 Bialystok, Poland
[**]Faculty of Architecture, Construction and Design, Department of Applied Mathematics and Mechanics,
Lutsk National Technical University, 75 Lvivska st., Lutsk, 43018, Ukraine

shyprao@gmail.com, h.sulym@pb.edu.pl, v.shvabyuk@gmail.com

received 3 December 2019, revised 21 May 2020, accepted 26 May 2020

Abstract: The influence of impulse load applied for different duration on the distribution of normalised dynamic radial stresses in positive and negative Poisson's ratio medium was investigated in this study. For solving the non-stationary problem in the case of plane deformation for structurally inhomogeneous materials, the model of Cosserat continuum was applied. This model enables accounting for the influence of shear-rotation deformation of micro-particles of the medium. In the framework of Cosserat elasticity, on applying the Fourier transforms for time variable and developing the boundary integral equation method, solving of the non-stationary problem reduces to the system of singular integral equations, where the components that determine the influence of shear-rotation deformations are selected. The numerical calculations were performed for the foam medium with positive and negative Poisson's ratio for different values of time duration of impulse. Developed approach can be used to predict the mechanical behaviour of foam auxetic materials obtained at different values of a volumetric compression ratio under the action of time variable load based on analysis of the distribution of radial stresses in foam medium.

Key words: Stress state, Foam, Poisson's ratio, Cosserat elasticity

1. INTRODUCTION

In recent years, the interest of scientists has grown to the creation of new materials, which have additional functional capabilities. These additional functional capabilities are wider than the properties which are defined by the composition of the material. Such materials are called smart materials. Expansion of their functional capabilities is based on the peculiarities of their internal structure. These materials also include the materials with negative Poisson's ratio ν (Scarpa et al., 2016; Brighenti, 2014). Such materials are called auxetics (Evans, 1991). Under the action of stretching force, they become thicker perpendicular to the applied force.

Reviews of auxetic materials and the possibility of their application were made by Carneiro et al. (2013), Novak et al. (2016), Ren et al. (2018) and others.

Developing the technology for creation of materials with a negative Poisson's ratio is important for using them in sports applications (Duncan et al., 2018) and other aspects (Naik et al., 2019).

Thus, in the work of Lakes (1991), the method of creating synthetic materials with negative Poisson's ratio was described. This method was based on triaxial compression and heat treatment of as-received foam for converting it to negative Poisson's ratio foam. An alternative method converting open-cell polyurethane foam into auxetic foam was based on using chemo-mechanical process (Underhill, 2017).

The researches of Lakes (2016), Rueger et al. (2016), Li et al. (2016) and others were dedicated to the investigation and comparison of elastic characteristics of foam materials with positive and negative Poisson's ratio. These experiments were performed in the framework of Cosserat elasticity. Applying the model of the moment theory of elasticity, we can account for the influence of material microstructure through consideration of shear-rotation deformation of micro-particles of a medium.

The results of the research performed (Lakes, 2016; Rueger et al., 2016; Li et al., 2016, etc.) confirm that the refined models of continuous medium mechanics should be used for investigation of the deformation processes in structurally inhomogeneous materials since the obtained results cannot be correctly described using the equations of the classical theory of elasticity. However, accounting for the influence of shear-rotation deformations leads to the complication of motion equation of the micropolar medium and obtaining the solutions with corresponding problems (Sulym et al., 2018).

A lot of papers investigated the dynamic behaviour of auxetics with different internal structures using computational analysis. In the work of Zang et al. (2014), the in-plane dynamic crushing behaviour of honeycomb auxetics with various cell wall angles was studied using dynamic finite element simulation. Strek et al. (2019) performed computational analysis of the dynamic behaviour of a three-layered sandwich beam with a metal foam core.

Besides, comparing the stress distribution in the initial foam material and auxetic foam obtained based on numerical modelling of the influence of time variable load on the stress state of these materials has a practical interest too. These numerical calculations must be performed in the framework of Cosserat elasticity.

Sylum et al. (2018) developed the method of investigating the stress state of structurally inhomogeneous materials in Cosserat elasticity. This method was based on the application of Fourier transforms of time variable and development of boundary integral equation method in the framework moment theory of elasticity.

Therefore, the aim of the work is comparison of the dynamic stress state of foam media with positive and negative Poisson's ratios under the action of impulse load which is applied to the boundary of tunnel cavity.

2. STATEMENT AND SOLUTION OF THE PROBLEM

2.1. Statement of the problem

Consider a foam medium with the tunnel cavity of constant cross section (Sulym et al., 2018). To the boundary of the cavity, non-stationary load is applied in the normal direction. Denote the boundary of a cross section of the cavity by *L* (Fig. 1). The centre of gravity is placed at the origin of a Cartesian coordinate system Ox_1x_2.

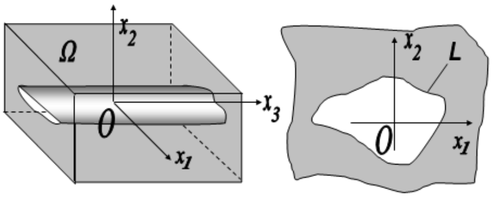


Fig. 1. 3D full geometry and 2D plane geometry of the problem

The problem consists in the determination and comparison of the dynamic stresses at negative and positive Poisson's ratio media with the tunnel cavity under the action of impulse load applied for different duration in a direction normal to the boundary of cavity.

The boundary conditions of the problem are written as (Sulym et al., 2018):

$$\sigma_n|_L = \sigma_0 P(t), \tau_{sn}|_L = 0, \ m_n|_L = 0, \tag{1}$$

where σ_0 is a constant and $P(t)$ is a known function which describes the change of load intensity over time.

2.2. Solution of the problem

If applying load in a direction normal to the boundary of cavity one can solving of 3D problem convert to solving 2D problem in case of plane deformation.

For the investigation of the stress state of the micropolar medium under the action of time variable load, the motion equations of Cosserat elasticity are used (Nowacki, 1974):

$$\sigma_{ji,j} + X_i = \rho \ddot{u}_i, \tag{2}$$

$$\in_{kij} \sigma_{ij} + m_{jk,j} + Y_k = J \ddot{\phi}_i, \tag{3}$$

where σ_{ji} is the force stress, m_{ji} is the couple stress, ρ is the material density, $\mathbf{X} = \{X_i\}$ is the mass forces vector, $\mathbf{Y} = \{Y_i\}$ is the couple forces vector, J is the inertia of unit volume rotation, $\in_{klm}$ is the permutation symbol, $\boldsymbol{u} = \{u_i\}$ is the displacement vector and $\boldsymbol{\phi} = \{\phi_k\}$ is the rotation vector. Functions u_i and ϕ_k are continuous functions.

Here and further, the Einstein summation convention is used. A comma at subscript denotes differentiation with respect to a coordinate indexed after the comma, i.e. $u_{j,i} = \partial u_j / \partial x_i$. Under the condition of plane, strain indices vary from 1 to 2, and k = 3.

The dependencies for determining force and couple stresses are written as (Nowacki, 1974):

$$\sigma_{ji} = (\mu + \alpha)\gamma_{ji} + (\mu - \alpha)\gamma_{ij} + \lambda \gamma_{kk} \delta_{ij},$$

$$m_{ji} = (\gamma + \varepsilon)\kappa_{ji} + (\gamma - \varepsilon)\kappa_{ij} + \beta \kappa_{kk} \delta_{ij}, \tag{4}$$

where α, β, γ and κ are the elastic constants required to describe an isotropic constrained Cosserat elastic solid, λ and μ are Lame parameters, $\gamma_{ij} = u_{i,j} - \in_{kji} \phi_k$ is the asymmetric deformation tensor and $\kappa_{ij} = \phi_{i,j}$ is the torsion bending tensor.

For solving the non-stationary problem of a dynamic stress state investigation in auxetic foam materials, Fourier transforms is used:

$$\hat{f}(\boldsymbol{x}, \omega) = \int_{-\infty}^{\infty} f(\boldsymbol{x}, t) e^{-i\omega t} dt, \tag{5}$$

where $i = \sqrt{-1}$ and ω is the frequency.

Applying Fourier transforms (5) to the motion equations (2) and (3), we obtain the equations which are equivalent to the equations of time-harmonic motion with cyclic frequency ω. Applying the weighted residual approach and collocation method (Vincent et al., 1994) to the obtained equations, the representations of displacement and microrotation transforms for the case of absence of concentrated forces are written as (Sulym et al., 2018):

$$\hat{u}_i = \int_L p_j \cdot U_{ij}^* dL + \int_L m_k \cdot \Phi_{kj}^* dL, \tag{6}$$

$$\hat{\phi}_k = \int_L p_j \cdot U_{kj}^{**} dL + \int_L m_k \cdot \Phi_{kk}^{**} dL, \tag{7}$$

where $U_{ij}^*, U_{kj}^{**}, \ \Phi_{kj}^*$ and Φ_{kk}^{**} are the fundamental functions for displacements and microrotations and p_j, m_k are unknown functions (Sulym et al., 2018). Sulym et al. (2018) showed that the obtained representations for the fundamental functions for displacements and microrotations $U_{ij}^*, U_{kj}^{**}, \Phi_{kj}^*$ and Φ_{kk}^{**} contain the components which correspond of the classical theory of elasticity and the components which are obtained in the framework of Cosserat elasticity.

In the field of Fourier transforms, we substitute the displacement and microrotation representations (6) and (7) with the formulas for determining force and couple stresses (4), thus satisfying the Fourier transforms of boundary condition (1).

Therefore, for applying the approach which was developed by Mikulich et al. (2017), we write integral functions in a complex form. Selecting the irregular components and using the formulas of Sokhotski–Plemelj for the limiting boundary values of the Cauchy integrals, we obtain the integral equations for determining the unknown on the boundary functions $p_1, p_2 \ m_3$ in the form:

$$\frac{Re(q)}{2} + PV \int_L (f_1(\zeta, \overline{\zeta}, z, \bar{z}) \cdot q \cdot dt + f_2(\zeta, \overline{\zeta}, z, \bar{z}) \cdot \bar{q} \cdot d\bar{t} + f_3(\zeta, \overline{\zeta}, z, \bar{z}) \cdot m_3 \cdot dL) = \sigma_0 \hat{P}(\omega), \tag{8}$$

$$\frac{Im(q)}{2} \vartheta_1 + PV \int_L (g_1(\zeta, \overline{\zeta}, z, \bar{z}) \cdot q \cdot dt + g_2(\zeta, \overline{\zeta}, z, \bar{z}) \cdot \bar{q} \cdot d\bar{t} + g_3(\zeta, \overline{\zeta}, z, \bar{z}) \cdot m_3 \cdot dL) = 0, \tag{9}$$

$$\frac{m_3}{2} + PV\int_L (G_1(\zeta,\overline{\zeta},z,\bar{z}) \cdot q \cdot dt + G_2(\zeta,\overline{\zeta},z,\bar{z}) \cdot \bar{q} \cdot d\bar{t} + G_3(\zeta,\overline{\zeta},z,\bar{z}) \cdot m_3 \cdot dL) = 0, \qquad (10)$$

where $f_l(\bullet), g_l(\bullet)$ and $G_l(\bullet)$ are the known complex functions (Sulym et al., 2018), $\vartheta_1 = \left(1 - \left(\frac{\alpha}{\mu}\right)^2\right)$ is a constant, and $\zeta = x_1 + i \cdot x_1$, $z = x_1^0 + i \cdot x_2^0$, $p \cdot dL = i \cdot q \cdot dt$, $p = p_1 + i \cdot p_2$. Here, the integrals are understood in the sense of Cauchy principal value.

The system of integral equations (8)–(10) is solved numerically by applying the method which is based on the mechanical quadrature method and collocation method (Mikulich et al., 2017).

For determining the Fourier transform of radial stresses, the integral representation is used (Sulym et al., 2018):

$$\widehat{\sigma}_r = PV\int_L (h_4(\boldsymbol{x},\boldsymbol{x}^0) \cdot p_1 + h_5(\boldsymbol{x},\boldsymbol{x}^0) \cdot p_2 + +h_6(\boldsymbol{x},\boldsymbol{x}^0) \cdot m_3)dL, \qquad (11)$$

where $h_k(\bullet)$ are known functions.

On the basis of the solutions of a system of integral equations, the values of dynamic stresses are obtained in a numerical form. For calculation of originals of the dynamic hoop and radial stresses, the modified discrete Fourier transform is used (Sulym et al., 2018):

$$\sigma_r(t_k) = \frac{2}{T}\mathrm{Re}\sum_{n=0}^{K-1}\left(\widehat{\sigma}_r(\omega_n)\exp\left(2\pi i\frac{n \cdot k}{K}\right) - - \textstyle\sum_{n=0}^{K-1}\widehat{\sigma}_r(\omega_n)\right), \qquad (12)$$

where K is the number of elements of selection, $\omega_n = 2\pi n/t$ is the frequency and $t_k = k \cdot t/K$, $k = 0 \ldots K-1$ is time.

3. NUMERICAL CALCULATION OF STRESSES

Let us study the influence of time duration of impulse load on the distribution of radial stresses in positive and negative Poisson's ratio polymer foam media with a tunnel cavity of circular cross section.

For numerical calculations, change of the load intensity over time is chosen in the form of weak shock impulse (Mikulich et al., 2017):

$$P(t) = p_*\bar{t}^{n_*}e^{-\alpha_*\bar{t}}, \bar{t} > 0, n_* \geq 0, \qquad (13)$$

where p_*, n_* and α_* are constants, $\bar{t} = \frac{t \cdot c_l}{a}$ is the time parameter, $c_l = \sqrt{\frac{\lambda+2\mu}{\rho}}$ is the speed of the expansion wave and a is the character size (for the case of tunnel cavity of circular cross section, $a = R$, where R is the radius of cavity).

The numerical calculations are performed for the foam material with negative and positive Poisson's ratio foam. The values of elastic characteristics of positive and negative foam materials are used with the work of Rueger et al. (2016). Here auxetic foam material was obtained under triaxial compression and heat treatment of positive Poisson's ratio foam.

The internal structure of positive Poisson's ratio foam is modelled as the complex of honeycomb with average cell size h (Fig. 2a), whereas the internal structure of the negative Poisson's ratio foam, obtained via triaxial compression, is modelled as the complex of re-entrant cells (Fig. 2b) (Novak et al., 2016; Ren et al., 2018).

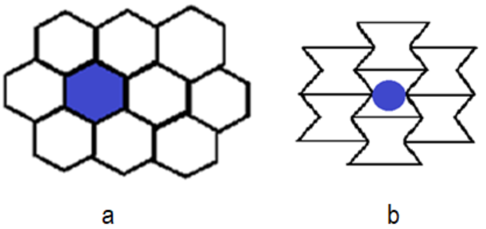


Fig. 2. 2D plane model of the foam material with positive (a) and negative (b) Poisson's ratio

In Rueger et al.'s (2016) study for positive Poisson's ratio foam, values of the physical characteristics were obtained: the density of foam was $\rho = 30\ \frac{\mathrm{kg}}{\mathrm{m}^3}$, the shear modulus was $\mathrm{G} = 45\ \mathrm{kPa}$, length characteristics in Cosserat elasticity were $\ell_\mathrm{t} = 2.1\ \mathrm{mm}$ and $\ell_\mathrm{b} = 9\ \mathrm{mm}$ and Poisson's ratio was $\nu = 0.3$. The inertia of unit volume rotation of honeycomb elements in case of $\mathrm{h} = 1.2\ \mathrm{mm}$ was $\mathrm{J} = 72.4 \cdot \times\ 10^{-6}\ \frac{\mathrm{kg}}{\mathrm{m}}$.

For a foam material with negative Poisson's ratio with a volumetric compression ratio 3.2, the value of an acute angle of a re-entrant element was 58.43°. This value is obtained without accounting for bending deformations of the edges of the cells. The inertia of unit volume rotation of re-entrant elements was $\mathrm{J} = 29.34 \cdot \times\ 10^{-6}\ \frac{\mathrm{kg}}{\mathrm{m}}$. For auxetic foam, the physical characteristics of the obtained material were as follows: the density of foam specimens was $\rho = 96\ \frac{\mathrm{kg}}{\mathrm{m}^3}$, the shear modulus was $\mathrm{G} = 16\ \mathrm{kPa}$, characteristics of length in Cosserat elasticity were $\ell_\mathrm{t} = 2.3\ \mathrm{mm}$ and $\ell_\mathrm{b} = 3.9\ \mathrm{mm}$ and negative Poisson's ratio was $\nu = -0.63$.

Moreover, on the basis of length characteristics in Cosserat elasticity, the values of the elastic constants required to describe an isotropic constrained Cosserat elastic solid are $\alpha = 28.36\ \mathrm{kPa}$, $\gamma = 0.0845\ \mathrm{N}$ and $\varepsilon = 0.8885\ \mathrm{N}$ for negative Poisson's ratio foam and $\alpha = 4.54\ \mathrm{kPa}$, $\gamma = 0.199\ \mathrm{N}$ and $\varepsilon = 14.382\ \mathrm{N}$ for positive Poisson's ratio foam.

For numerical calculations, values of constants of the impulse load in equation (14) are chosen as: $\alpha^* = 1.25, p^* = 2.89, n^* = 2$ for the case of $t_* = 8$; $\alpha^* = 0.9375, p^* = 1.64, n^* = 2$ for the case of $t_* = 12$ and $\alpha^* = 0.625, p^* = 0.7225, n^* = 2$ for the case of $t_* = 16$, where t_* is a dimensionless paremeter, which characterizing of the impulse duration. .

Numeric calculations of dynamic normalised radial stresses at the internal points of the foam media are shown in Figs 4–6 for the case $t_* = 8, 12$ and 16 accordingly. Here, t_* is the dimensionless parameter characterised by impulse duration: the period of growth and attenuation of load intensity.

Numeric results are calculated for the four values of the distance δ between the boundary of the cavity and the internal points of medium (Fig. 3) (Mikulich et al., 2017).

In Figs 4–6, curves 1–4 correspond to the case of $\delta = 1.5R$, $\delta = 2.5R$, $\delta = 5R$ and $\delta = 7.5R$. Numerical calculations are performed for the case of the radius of cavity cross section $R = \ell_C$, where $\ell_C = \sqrt{\frac{\gamma+\varepsilon}{2\alpha}}$ is scale parameter in Cosserat elasticity (Sulym et al., 2018).

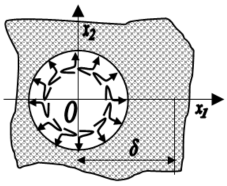


Fig. 3. 2D plane strain geometry of the problem

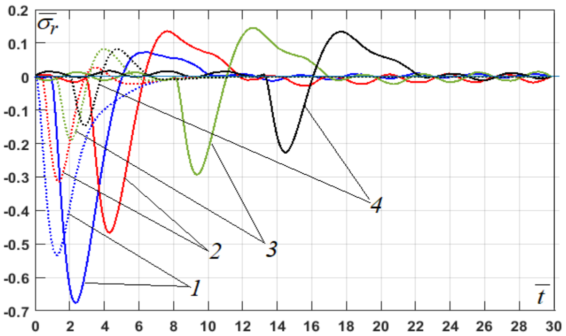


Fig. 4. Distribution of normalised dynamic radial stresses in foam media with tunnel cavity for the case $t_* = 8$

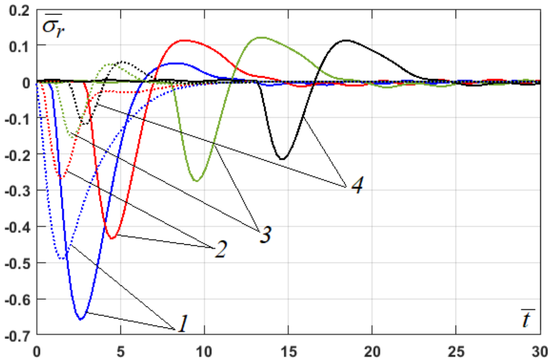


Fig. 5. Distribution of normalised dynamic radial stresses in foam media with tunnel cavity for the case $t_* = 12$

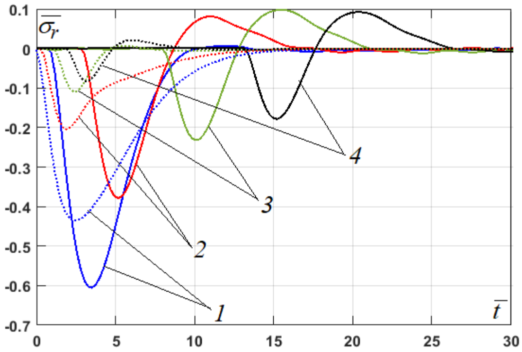


Fig. 6. Distribution of normalised dynamic radial stresses in foam media with tunnel cavity for the case $t_* = 16$

In Figs 4–6, numeric calculation results of normalised radial stresses are shown by solid lines for the case of negative Poisson's ratio foam and by dotted lines for the case of positive Poisson's ratio foam.

Analysis of results of numerical calculations which is shown in Figs 4–6 confirms that the effect of impulse loading on foam materials causes higher stresses in auxetics than in foam materials with a positive Poisson's ratio. Thus, for the foam materials which are investigated, the maximum values of radial stresses in the interior points of the foam media which are located near the boundary of the cavity are 44%–54% of the intensity applied to the boundary load for the positive Poisson's ratio foam and 60%–68% for negative Poisson's ratio.

However, the speed of wave propagation in the foam medium with negative Poisson's ratio is much lower than in medium with positive Poisson's ratio. It also explains better operational properties of the foam materials with negative Poisson's ratio due to transfer of part of the deformations into sections that are farther away from the defect and to reduce localised effects in the material during its using.

Fig. 7 shows the dependence of the ratio of the maximum values of radial stresses in foam media with a negative and positive Poisson's ratio for changing the distance δ from the tunnel cavity boundary for different impulse durations.

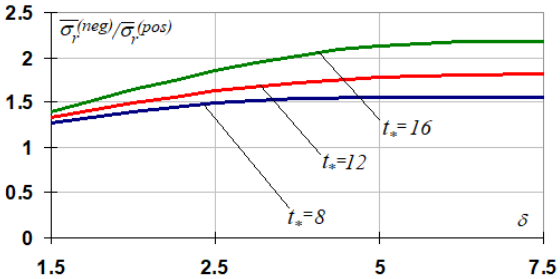


Fig. 7. The ratio of maximum values of radial stresses in foam media

Besides, based on the numerical calculations, it has been shown that for foam auxetic material, the maximum values of normalised dynamic radial stresses are higher by 1.2 to 2.2 times than the corresponding ones calculated for a positive Poisson's ratio foam. It can be explained that auxetic foam materials have a higher density than classical foams. Besides, the load that causes compression deformation in pre-compressed materials passes higher dynamic stresses than in materials with a positive Poisson's ratio.

Also, for foam materials with positive and negative Poisson's ratio, the maximum value of the normalised dynamic radial stresses in the internal points of the stress medium slightly depends on the duration of the impulse. In this case, the smaller the duration of the impulse, the higher are the values of the corresponding stresses. The oscillating character of radial stress distribution in foam materials with negative Poisson's ratio, which occurs after impulse expansion, is explained by lesser stiffness of re-entrant element compared with honeycomb.

4. CONCLUSION

Based on modify method of boundary integral equations (Sulym et al., 2018) in the framework of Cosserat continuum, comparison of the distribution of normalised dynamic radial stresses in materials with a positive and negative Poisson's ratio was performed. The results of the numerical calculations confirm the necessity of using refined theories of the continuum mechanics (moment theory of elasticity – Cosserat elasticity) to study the stress state of foam materials with positive and negative Poisson's ratio.

Based on the numerical calculations, it is established that the distributions of the relative radial stresses in foam media with positive and negative Poisson's ratio have similar characteristics, but the stresses arising from compressive stresses in materials with positive Poisson coefficient are smaller.

An adaptive approach for the investigation of the effect of time variable load on the distribution of radial stresses in foam media can be used to predict the mechanical behaviour of foam auxetic materials obtained at different values of a volumetric compression ratio.

REFERENCES

1. **Brighenti R.** (2014), Smart behaviour of layered plates through the use of auxetic materials, *Thin-Walled Structures,* 84, 432-442.
2. **Carneiro V., Meireles J., Puga H.** (2013), Auxetic materials — A review, *Materials Science-Poland,* 31(4), 561-571.
3. **Duncan O., Shepherd T., Moroney Ch., Foster L., Venkatraman Pr, Winwood K., Allen T., Alderson A.** (2018), Review of Auxetic Materials for Sports Applications: Expanding Options in Comfort and Protection, *Applied Sciences,* 8, 941, 1-33.
4. **Evans K.** (1991), Auxetic Polymers: A New Range of Materials*, Endeavour*, 15(4), 170–174.
5. **Grima J., Attard D., Gatt R., Cassar R.** (2009), A Novel Process for the Manufacture of Auxetic Foams and for Their re-Conversion to Conventional Form, *Advanced Engineering Materials,* 11(7), 533-535.
6. **Lakes R. S.** (1991), Experimental Micro Mechanics Methods for Conventional and Negative Poisson's Ratio Cellular Solids as Cosserat Continua, *Journal of Engineering Materials and Technology*, 113, 148-155.
7. **Lakes R. S.** (2016), Physical Meaning of Elastic Constants in Cosserat, Void, and Microstretch Elasticity, *Journal of Mechanics of Materials and Structues*, 11(3), 217-229.
8. **Li D., Dong L., Lakes R.** (2016), A Unit Cell Structure with Tunable Poisson's Ratio from Positive to Negative, *Materials Letters*, 164, 456-459.
9. **Mikulich O., Shvabyuk V., Sulym H.** (2017), Dynamic Stress Concentration at the Boundary of an Incision at the Plate under the Action of Weak Shock Waves, *Acta Mechanica et Automatica,* Vol. 11, No. 3, 217-221.
10. **Naik S., Dandagwhal R., Wani C., Giri S.** (2019), A review on various aspects of auxetic materials. *AIP Conference Proceedings*, 2105 (1), 10.1063/1.5100689.
11. **Novak N., Vesenjak M., Ren Z.** (2016), Auxetic Cellular Materials - a Review. *Journal of Mechanical Engineering,* 62(9), 485-493.
12. **Nowacki W.** (1974), *The Linear Theory of Micropolar Elasticity,* Springer, New York.
13. **Ren X. , Das R., Tran P., Ngo T., Xie Y.** (2018), Auxetic Metamaterials and Structures: A Review, *Smart Mater. Struct*., 27, 1-38.
14. **Rueger Z., Lakes R.S.** (2016), Cosserat elasticity of negative Poisson's ratio foam: Experiment, *Smart Materials and Structures,* Vol. 25, 1-8.
15. **Scarpa F., Alderson A., Ruzzene M., K.** (2016), Auxetics in smart systems and structures, *Smart Materials and Structures*, 25(5), 1-8.
16. **Strek T., Michalski J., Jopek H.** (2019) Computational analysis of the mechanical impedance of the sandwich beam with auxetic metal foam core, *Physica Status Solidi B*, Vol. 256 (1), 1800423, 10.1002/pssb.201800423.
17. **Sulym H., Mikulich O., Shvabyuk V.** (2018), Investigation of the dynamic stress state of foam media in Cosserat elasticity, *Mechanics and Mechanical Engineering,* Vol. 22, No.3, 739-750.
18. **Underhill R.S.** (2017), Manufacture and characterization of auxetic foams, DRDC-RDDC-2017-R099.
19. **Zhang X., Ding H., An Li.** (2014), Numerical Investigation on Dynamic Crushing Behavior of Auxetic Honeycombs with Various Cell-Wall Angles, *Advances in Mechanical Engineering, 10.1155/2014/679678.*

DOI 10.2478/ama-2020-0012

PROCESS ANALYSIS OF THE HOT PLATE WELDING OF DRIVE BELTS

Krzysztof WAŁĘSA,*[*] Ireneusz MALUJDA,[*] Krzysztof TALAŚKA,[*] Dominik WILCZYŃSKI[*]

[*]Institute of Machine Design, Poznań University of Technology, ul. Piotrowo 3, 61-138 Poznań, Poland

krzysztof.walesa@put.poznan.pl, ireneusz.malujda@put.poznan.pl,
krzysztof.talaska@put.poznan.pl, dominik.wilczynski@put.poznan.pl

received 11 February 2020, revised 25 May 2020, accepted 28 May 2020

Abstract: Most industrial machines use belt transmission for power transfer. These mechanisms often use the round belts of several millimetres in diameter that are made of thermoplastic elastomers, especially polyurethane. Their production process calls for bonding the material, which is often performed by hot plate butt welding. In order to achieve proper design of an automatic welding machine, the authors analysed the hot plate welding process of round belts. This process consists of five phases. It is necessary to recognize all the physical phenomena that occur during welding, especially those connected with thermomechanical properties of material. This knowledge is necessary to determine the temperature distribution during each step of the process. The paper presents a standard welding cycle together with an explanation of the physical phenomena in each phase. An analysis of these fundamentals will be used to derivate the function of temperature distribution during all process phases. In addition, some assumptions for calculation of temperature distribution and some fundamental physic correlations were presented.

Keywords: Hot plate welding, round drive belts, heating, thermal conducting, plasticizing

1. INTRODUCTION

Industrial grade belts, both used in drives and conveyors, are commonly utilized in machine building for transferring torque between the connected workstations or for transporting products on manufacturing lines. Both full cross-section and perforated conveyor belts are utilized in transportation, especially the latter are present in vacuum transportation of light objects (Wojtkowiak et al., 2018; Wojtkowiak and Talaśka, 2019; Wilczyński, 2019). The most commonly used types of drive belts are flat, toothed and shaped belts. Toothed belts are used both in simple transmissions as well as non-classical solutions, for example, with variable ratios (Domek and Dudziak, 2011; Domek et al., 2016; Krawiec et al., 2018; Krawiec et al., 2019). Shaped belts, in particular V-shaped and round belts, are commonly employed in drives, that is, special robotic arm joint mechanism, with two twisted small diameter belts (Inoue et al., 2016), they can be made of rubber (Kukla et al., 2015) or polyester or polyurethane based elastomers (Behabelt, 2015). Their common application in industrial machines calls for an efficient manufacturing process, which usually takes place in two stages (Sikora, 1993). First of all, a long belt is manufactured and cut down to required size. Finally, the ends are joined permanently to form a continuous loop (Wałęsa, 2018).

Due to the peculiar characteristics of thermoplastic elastomers used in the manufacturing of such belts, it is possible to join them by hot welding (Groover, 2015). A specific approach to this process is butt-welding utilizing the hot plate method; it is popular due to its simplicity and efficiency (Grewell and Benatar, 2007; Yousepour et al., 2004). This method is commonly employed in the automotive industry and civil engineering, for example, in the process of joining: tanks for utility fluids, lamp enclosures, engine instrumentation (Pietrzak et al., 2019; Grewell and Benatar, 2007) and pipes (Rzasinski, 2017; Troughton, 1997; Cocard et al., 2009). Furthermore, studies were carried out on the hot welding of inflexible polymer materials, for example, acrylonitrile butadiene styrene copolymer (ABS) (Mokhtarzadeh and Benatar, 2012), polycarbonate (PC) (Krishnan and Benatar, 2004), as well as polypropylene (PP) (Nieh and Lee, 1992). However, it needs to be pointed out that all of them apply primarily to plastics.

The authors started design work on a device for automated butt welding of drive belts utilizing the hot plate technique, which is to improve the efficiency of the manufacturing process of continuous belts. For the purpose of verification of the design assumptions, the process has to be analysed together with further study of the influence of heating parameters on weld quality (Wałęsa et al., 2019a and 2019b). One of the basic research actions regarding this topic is to analyse the heat transfer during the hot plate welding process. To this end, it is required to identify the heat transfer phenomena taking place during hot plate welding. These considerations will consequently form the basis for performing temperature distribution calculations, which describe the temperature values at every point of the material, according to time. The results obtained from calculations and research, in combination with known material characteristics, will be used for calculating the plasticized distance. Identifying this value and its dependence on process variables (e.g., hot plate temperature, time and applied force) will be used to control the hot plate welding process. This allows to anticipate the extent of belt shortening throughout the manufacturing process and to determine the best parameters to obtain a satisfactory weld quality. Moreover, analysis of plasticized distance and temperature distribution will be used in the research regarding the influence of welding process to macrostructure of material. Thanks to this, the division for three heat affected zones, commonly expected for semi-crystalline material (Casalino and Ghorbel, 2008), will be verified.

In addition, it should be noticed that a lot of thermal exchanging processes can be described in a mathematical way, especially using the FEM implements (Dyja et al., 2017; Gawrońska, 2019; Kubiak, 2019; Saternus et al., 2018 and 2019; Winczek et al., 2016). Considering the partial results of examinations of the plasticizing process and mechanical parameters derivation (Wałęsa et al., 2020a and 2020b), in further works, the hot plate welding process will be described by a mathematical model, which will take into consideration conclusions from the presented process analysis.

2. THE HOT PLATE WELDING CYCLE

Concerning the necessary technological operations, the hot plate welding process of the drive belts can be divided into five phases (Wałęsa et al., 2019a and 2019b; Klimpel, 2000; Potente et al., 2002; Jasiulek, 2006), with different physical phenomena occurring during the heat transfer in each phase. Considering this fact, the approximate temperature distribution can be expected (Potente et al., 2002). One of the most important activities is the plasticization of the belt end that enables the chemical reaction and physical interaction between the macromolecules in the joining process (Amanat et al., 2010; Amancio-Filho and dos Santos, 2009; Madej and Ozimina, 2010; Puszka, 2006; Żuchowska, 2000). The course of plasticization process conditions the rest of welding process, especially the ability of the ends of the belts to make durable joint during cooling step, where durable connections between macromolecules are made (Ciszewski and Radomski, 1989).

2.1. The matching phase

The first phase of welding is the matching of belt ends (Fig. 1). It entails plasticizing and melting the flat surfaces of the belt during their contact with the heated plate. The plate has the temperature T_p, which facilitates the welding process by gaining the welding temperature T_w in the bonded material. The welding temperature choice is very important issue and its value depends on polymer type. Exceeding the degradation temperature causes destructing the material in a few stages (Wanqing et al., 2017). On the other hand, in case of some polymers, joint strength increases with welding temperature (Evers et al., 2017).

In matching phase, the belt ends (1) are held by shaped grips (2), and are moved towards the hot plate surface (3) with velocity v_m. After the surfaces of the belt and the hot plate make contact, the belt is pressed to the plate with force F_m. In this phase, flash is formed due to the belt plasticizing in contact with the hot plate. The partial melting results in the surface of the belt adapting to the surface shape of the hot plate. This is the result of the material's reaction to the temperature of the hot plate.

The belt heating causes the temperature to increase along its axis. Two main characteristic quantities can be therefore distinguished:

- the distance p, where belt temperature exceeds the welding temperature T_w (the temperature when a material is plasticized and melted),
- the distance h, where belt temperature exceeds T_0.

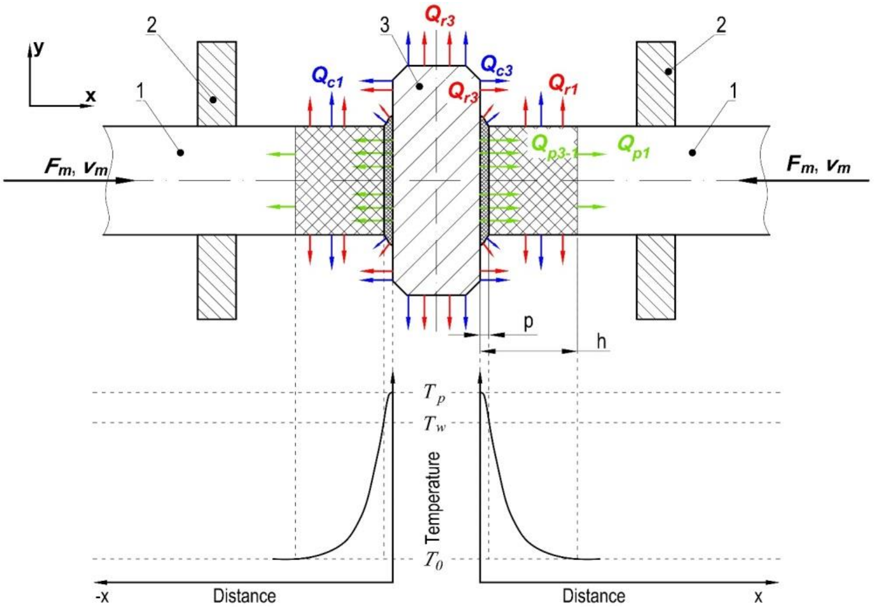


Fig. 1. The matching phase of the hot plate welding process with expected temperature distribution: 1 – belt, 2 – shaped grip, 3 – hot plate, F_m – matching force, v_m – matching velocity, T_p – the hot plate temperature, T_w – welding temperature, T_0 – ambient temperature, p – plasticize distance, h – heating distance; Q_{p3-1}, Q_{p1}, Q_{r1}, Q_{r3}, Q_{c1}, Q_{c3} – heat

During this phase, the following heat transfer phenomena are observed: conduction, convection and radiation. Considering the particular heat fluxes, the following can be identified:

- the heat Q_{p3-1}, which stands for heat transfer from the hot plate to the belt by contact conduction,
- the heat Q_{p1}, which stands for heat transfer in the volume of the belt material through conduction,
- the heat Q_{c3}, which stands for heat transfer from the hot plate through convection,
- the heat Q_{r3}, which stands for heat transfer from the hot plate through radiation,
- the heat Q_{c1}, which stands for heat transfer from the outer surface of the belt through convection,
- the heat Q_{r1}, which stands for heat transfer from the outer surface of the belt through radiation.

In this case, the dominant methods of heat transfer are: conduction in the belt material and contact conduction between the hot plate and the belt surface. The radiation is not as significant as the phenomena mentioned above, because the process temperature is relatively low. In further analysis, it is possible to disregard the issue of convection because the airflow around the belt or hot plate will be not forced; therefore, forced convection is not present. It is also possible to disregard natural convection during the analysis because the ends of the belt will be clamped in the shaped grips and dies, thus limiting the heat energy transfer to the cooler air.

The matching phase is very important. The melting of the belt surface as a result of hot plate contact allows matching the belt to the flat surface of the plate. Consequently, belt surfaces are placed parallel to each other and perpendicular to the central axis. Thus, the flat ends of the belt are properly aligned for bonding. The most important effect of material plasticizing is the removal of any inequalities, roughness and faults, which may have been incurred during the cutting process. This allows to avoid joint defect issues caused by inadequate surface preparation.

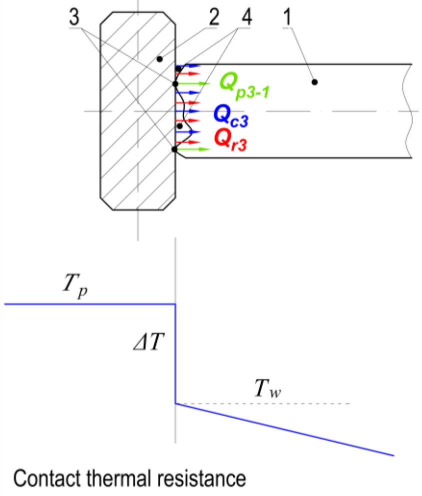


Fig. 2. Heating of the non-matched belt surface by the hot plate: 1 – belt, 2 – hot plate, 3 – contact points, 4 – cavities (filled by hot gases); Q_{p3-1}, Q_{r3}, Q_{c3} – heat

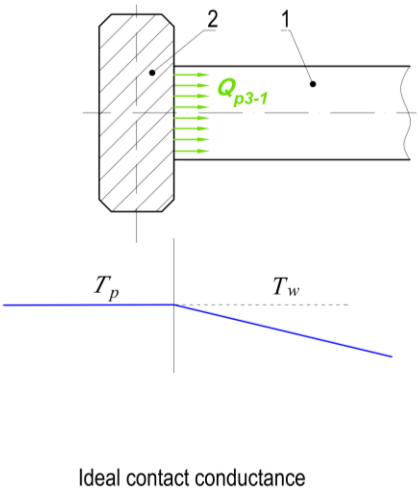


Fig. 3. Heating of the matched belt surface by the hot plate: 1 – belt, 2 – hot plate; Q_{p3-1} – heat

Another benefit of the matching phase is providing uniform conditions for heat transfer along the whole flat surface of the belt (Amancio-Filho and dos Santos, 2009). This allows to replace the point thermal conduction (Q_{p3-1}), convection (Q_{c3}) and radiation (Q_{r3}) occurring through the cavities in the material (Fig. 2) with whole surface conduction (Qp_{3-1}) resulting in more stable heat transfer conditions (Fig. 3).

2.2. The heating phase

During the heating phase (Fig. 4), the ends of the belt (1) are pressed against the hot plate (3) by the shaped grips (2) with force F_h. In this phase, the heat transfer phenomena are similar to the previous one. The area of the convection (Q_{c1}) and the radiation (Q_{r1}) effect is expanded. This is the result of the expansion of the heated area. Due to this fact, a new heat transfer mechanism occurs: conduction between the outer surface of the belt and shaped grips (Q_{p1-2}).

The main assumptions of the heating phase are as follows:
- increase of the plasticized (melted) distance p to the value ensuring the best conditions to perform the joining (chemical reactions and mechanical interactions between macromolecules), the increase in the h distance is the side effect of the heating,
- the F_h force value is significantly lower than the matching force F_m (F_h is about 10%–20% of F_m) because of the sharp decrease in polymer viscosity together with the increase in temperature (Klimpel, 2000). If too much clamping force is applied during the heating phase, it causes excessive flash formation. The first negative consequence is unnecessary loss of material. Secondly, the plasticized polymer is displaced from the joining area, which is disadvantageous because it can impede the joining process and form cavities in the central area of the joint. On the other hand, it is required to utilize a small value of F_h force, as otherwise, issues with contact breaking might occur due to the outflow of material from the heating area. In this case, the belt ends should be pressed toward the hot plate, but this is impossible without applying F_h force. It should be noticed that this action can be performed automatically owing to the thermal expansion of the belt material.

2.3. The switchover phase

The switchover phase consists in removing the hot plate from the area between the ends of the belt (Fig. 5).

The belt (1) is moved slightly by the shaped grips (2) to achieve a small clearance from the hot plate; however, this action can be disregarded in the heat transfer analysis. In this phase, the heating process ends, because of the source of energy is being withdrawn. Consequently, the switchover phase should be as short as possible. The convection (Q_{c1}) and the radiation (Q_{r1}) from the belt surfaces (especially in the flat surfaces of the weld) cause unnecessary loss of energy. Considering this fact, it is necessary to control the belt temperature, which has a tendency to decrease mainly in the flat, plasticized area. This is the result of the convection phenomenon, which becomes more significant with higher surface temperature. It is critical to maintain the temperature of the welding area greater than T_w. For this reason, the switchover time should be as short as possible.

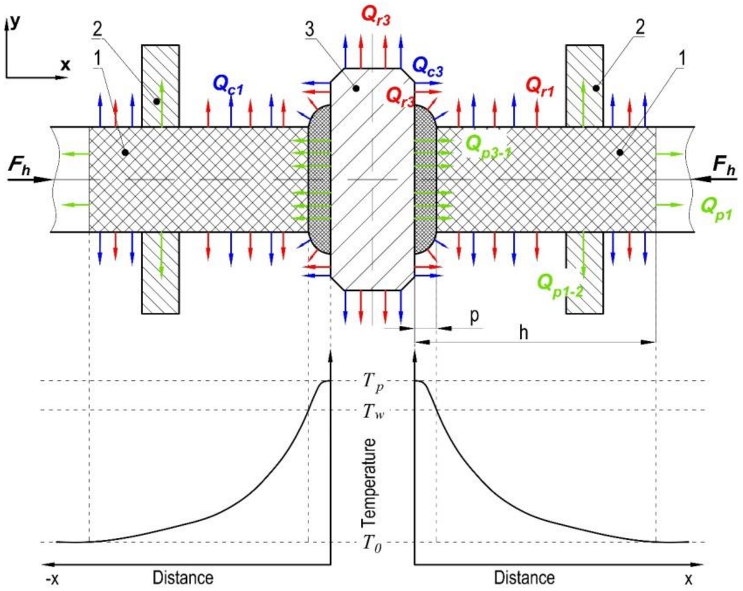


Fig. 4. The heating phase of the hot plate welding process with expected temperature distribution: 1 – belt, 2 – shaped grip, 3 – hot plate, F_h – force during heating, T_p – the hot plate temperature, T_w –welding temperature, T_0 – ambient temperature, p – plasticized distance, h – heated region length; Q_{p3-1}, Q_{p1-2}, Q_{p1}, Q_{r1}, Q_{r3}, Q_{c1}, Q_{c3} – heat

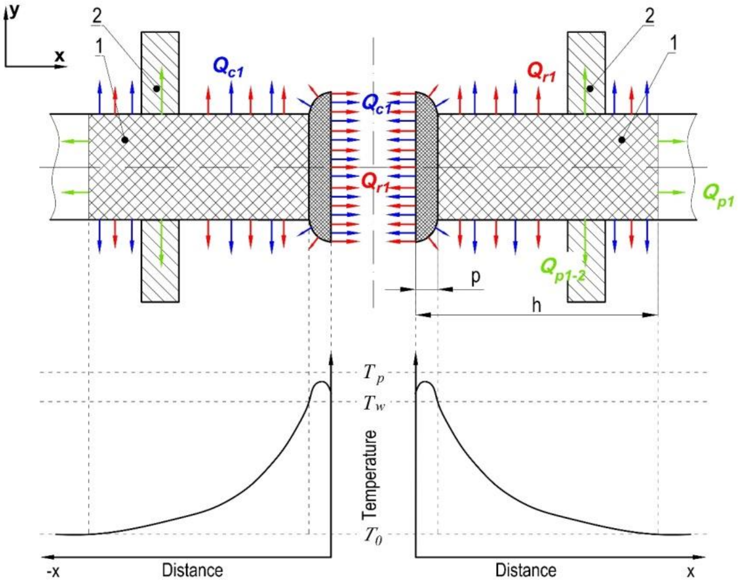


Fig. 5. The switchover phase of the hot plate welding process with expected temperature distribution: 1 – belt, 2 – shaped grip, T_p – the hot plate temperature, T_w –welding temperature, T_0 – ambient temperature, p – plasticize distance, h – heated region length; Q_{p1-2}, Q_{p1}, Q_{r1}, Q_{c1} – heat

2.4. The joining phase

In the joining phase, the ends of the belt (1) are clamped by shaped grips (2) and moved towards each other, which results in joining their surfaces (Fig. 6).

The joining phase is critical, because it is where the material bonding begins. It comprises two main characteristic phenomena:

- chemical reactions between molecules of the hydrocarbons,
- mechanical interactions between macromolecules which result in their splicing (Grewell and Benatar, 2007).

During the joining phase, one needs to consider a compromise between the short duration of the phase and low joining velocity v_j. The short phase time is associated with the relatively high velocity v_j, which is beneficial because of reduced energy loss by convection (Q_{c1}) and radiation (Q_{r1}). On the other hand, the joining velocity v_j and the joining force F_j should be as low as possible, to avoid the problem of excessive distortion of the plasticized region.

2.5. The cooling phase

In the cooling phase (Fig. 7), the ends of the belt (1) are pressed against each other. Free cooling occurs at the joint and the belt as a result of the interaction with the air surrounding the belt. In the ideal scenario, the duration of this activity should be sufficiently long, allowing the belt temperature to decrease to the level of ambient temperature T_0. Afterwards, the belt can be removed from the grips and the joint can be subject to further processing to remove the flash.

The main assumptions for the cooling phase are:

DOI 10.2478/ama-2020-0012

- belt temperature is equalized by contact conduction,
- continuation of the chemical reactions and the mechanical interactions occurring between macromolecules,
- main heat transfer occurs through conduction (Q_{p1}) in the material,
- heat transfer with the environment only occurs via natural convection (Q_{c1}) and radiation (Q_{r1}),
- the duration of this step depends on belt temperature – its value should decrease to the level of ambient temperature T_0 at the end of the phase.

The cumulative hot plate welding process time, ranging from over a dozen seconds to a few dozen minutes (Cocard et al., 2009), depends on the dimensions of the welded parts. Usually, the longest is the cooling phase (Klimpel, 2000).

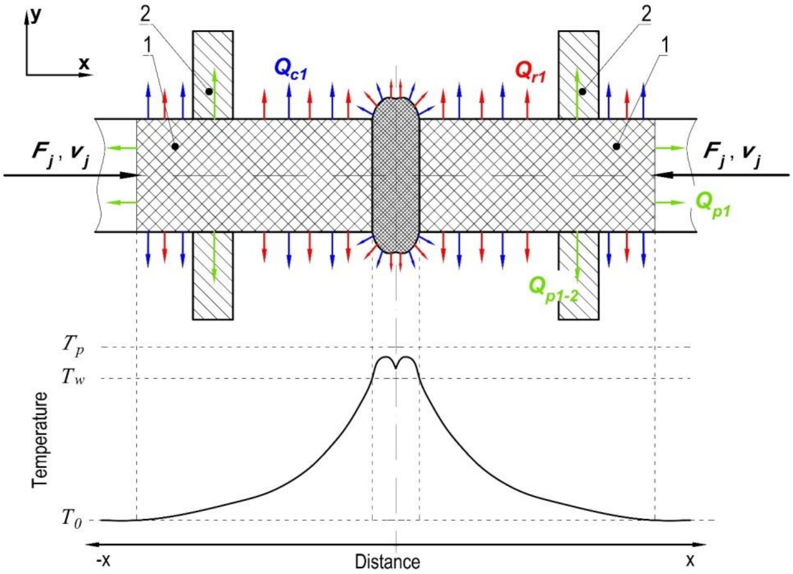


Fig. 6. The joining phase of the hot plate welding process with expected temperature distribution: 1 – belt, 2 – shaped grip, F_j – joining force, v_j – joining velocity, T_p – the hot plate temperature, T_w –welding temperature, T_0 – ambient temperature, p – plasticize distance, h – heated region length; Q_{p1-2}, Q_{p1}, Q_{r1}, Q_{c1} – heat

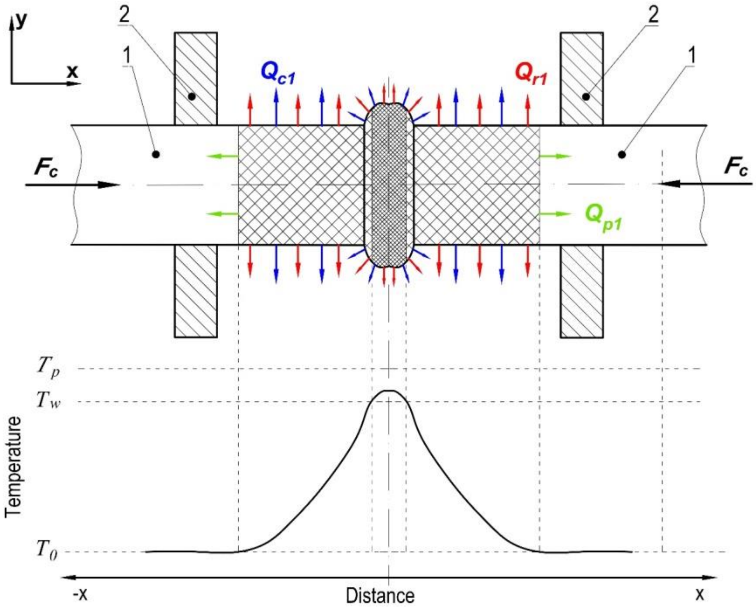


Fig. 7. The cooling phase of the hot plate welding process with expected temperature distribution: 1 – belt, 2 – shaped grip, F_c – force during cooling, T_p – the hot plate temperature, T_w –welding temperature, T_0 – ambient temperature, Q_{p1}, Q_{r1}, Q_{c1} – heat

3. MODELLING OF THE HEAT TRANSFER PHENOMENA

The complete analysis of the hot plate welding process calls for establishing temperature distribution. This function describes the temperature values at all the points of the heated region, according to time (Carslaw and Jaeger, 1959). To obtain this derivation, it is necessary to solve the heat transfer equations.

DOI 10.2478/ama-2020-0012

3.1. Simplifications regarding the heat transfer phenomena

To simplify analysing the heat transfer process in preliminary description of the hot plate welding process, some factors can be omitted. Furthermore, several assumptions can be made for the purpose of performing the mathematical calculation:

- the transient heat transfer during the entire process, that is, the temperature in every point of the belt changes throughout the process duration,
- belt as a long body, with high length to diameter ratio,
- uniform temperature distribution in the belt cross section, because the belt material is homogeneous in the macro scale; therefore, it can be treated as isotropic,
- the phenomenon of heat generation is not present,
- considering the heat transfer along the main axis of the belt as unidimensional, heated from a flat surface together as well as the isotropic model of the material,
- heating of both ends of the belt is symmetrical,
- disregarding natural convection from the outer surface of the belt and the hot plate, because the airflow is not forced, and therefore, natural convection is not a significant factor,
- omitting the radiation from the outer surface of the belt and the hot plate. This assumption is based on the relatively low temperature and a short time,
- disregarding heat conduction from the belt to shaped grips, which are a relatively long distance away from the heating area, and therefore, their contribution is not significant,
- constant hot plate temperature on the whole surface, this allows to simplify the issue because the temperature change due to the energy transfer can be omitted,
- disregarding the effect of thermal expansion,
- density, heat transfer coefficient and specific heat of the belt are not dependent on temperature. Considering the preliminary calculations, we can assume that these parameters do not vary with temperature increase,
- the hot plate is a rigid and non-deformable body.

These assumptions can be made, due to relatively low process temperature, which does not exceed 300°C (Wałęsa et al., 2019b and 2020a). These simplifications allow to perform the calculation on temperature distribution taking into consideration the commonly known heat transfer cases, that is, the semi-infinite space, the round rod placed between the walls or the round ribs (Carslaw and Jaeger, 1959).

3.2. Possibilities of mathematical modelling

The heat transfer in almost every phase of hot plate welding is primarily through conduction. Considering this fact, it is required to solve the heat conduction equation (1). In general case, it takes the following form (Staniszewski, 1979):

$$\frac{\partial}{\partial x}\left(\lambda\frac{\partial T}{\partial x}\right)+\frac{\partial}{\partial y}\left(\lambda\frac{\partial T}{\partial y}\right)+\frac{\partial}{\partial z}\left(\lambda\frac{\partial T}{\partial z}\right)+q_v=c_p\cdot\rho\cdot\frac{\partial T}{\partial t}. \qquad (1)$$

Using the simplifications mentioned above, it can be assumed that: thermal conductivity (λ), latent heat (c_p) and density (ρ) are constant within the whole temperature range. Moreover, internal heat generation (q_v) can be omitted. Consequently, the simplified Fourier's equation (2), shall have the following form:

$$a\cdot\nabla^2 T=\frac{\partial T}{\partial t}, \qquad (2)$$

where: a stands for thermal diffusivity, described in equation:

$$a=\frac{\lambda}{c_p\cdot\rho}. \qquad (3)$$

Assuming the unidimensional character allows to further simplify equation (2) to achieve the following form:

$$a\cdot\frac{\partial^2 T}{\partial x^2}=\frac{\partial T}{\partial t}. \qquad (4)$$

The solution of this differential equation is not obvious, and its form depends on the starting and boundary condition. For the matching and heating phase, it can utilize the first type of boundary conditions, where the value of the hot plate and the ambient temperature is set. The temperature distribution in switchover phase can be simplified by transferring the coordinates of heating temperature distribution. For the joining and cooling phase, it is required to use the third type boundary condition in the following form (Carslaw and Jeager, 1959):

$$\frac{\partial T(x,t)}{\partial x}=-\frac{\alpha}{\lambda}\cdot T(x,t), \qquad (5)$$

where: α is the heat transfer coefficient.

Obviously, other assumptions may be made as well. For example, it is possible to account for melting. In this case, thermodynamic constants (α, λ, c_p) and density (ρ) will be different in solid and liquid phase (Taler and Duda, 2003). This fact calls for further analysis.

4. SUMMARY

We distinguish the following heat transfer phenomena that are of key importance for the hot plate welding process: conduction, convection and radiation. It is necessary to analyse their specific parameters for each phase of the hot plate welding. Considering the phenomena in particular, together with employing a number of simplifications allows to derive the temperature distribution, which describes the temperature in every point along the belt axis. This function is to be solved analytically and the results will be examined by FEM implements. Afterwards, it is anticipated that experimental research should allow to verify the model.

The temperature distribution along the axis of the belt, together with the known material characteristics, will be used to establish the plasticized area of the belt as well as the shortening of the belt during the hot plate welding process.

After preliminary analysis, with listed simplifications of heat transfer modelling, authors will make an effort to the estimate radiation and convection coefficients. After that, the temperature distribution results will be compared with calculations regarding simplifications, to obtain precisely described temperature distribution, and real error of calculations with simplifications.

REFERENCES

1. **Amanat N., James N. L., McKenzie D.R**. (2010), Welding methods for joining thermoplastic polymers for the hermetic enclosure of medical devices, *Medical engineering & Physics,* 32, 690–699.
2. **Amancio-Filho S.T., dos Santos J.F.** (2009), Joining of Polymers and Polymer-Metal Hybrid Structures: Recent Developments and Trends, *Polymer Engineering and Science,* 49(8), 1461–1476.
3. **Behabelt** (2015), *Product Catalogue 2015/2016,* Behabelt, Glottertal.
4. **Carslaw H.S., Jeager J.C.**, 1959, *Conduction of Heat in Solids*, Oxford University Press, London.

5. **Casalino G., Ghorbel E.** (2008), Numerical model of CO_2 laser welding of thermoplastic polymers, *Journal of Materials Processing Technology*, 207, 63–71.
6. **Ciszewski A., Radomski T.** (1989), *Constructuion materials in machine design (in Polish),* PWN, Warszawa.
7. **Cocard M., Grozav I., Iacob M., Caneparu A.** (2009), Establishing the Optimum Welding Procedure for PE 100 Polyethylene Pipelines Using the Response Surface Design, *Materiale Plastice*, 46(4), 452–457.
8. **Domek G., Dudziak M.** (2011), Energy Dissipation in Timing Belts Made From Composite Materials, *Advanced Material Research,* 189-193, 4414–4418.
9. **Domek G., Kołodziej A., Dudziak M., Woźniak T.** (2016), Identification of the quality of timing belt pulleys, *Procedia Engineering*, 177, 275–280.
10. **Dyja R., Gawrońska E., Grosser A.** (2017), Numerical problems related to solving the Navier-Stokes equations in connection with the heat transfer with the use of FEM, *Procedia Engineering*, 177, 78–85.
11. **Evers F., Schöppner V., Lakemeyer P.** (2017), The influence on welding processes on the weld strength of flame-retardant materials, *Weld World*, 61, 161–170.
12. **Gawrońska E.** (2019), A sequential approach to numerical simulations of solidification with domain and time decomposition, *Applied Sciences,* 9(10).
13. **Grewell D., Benatar A.** (2007), Welding of plastics: fundamentals and new developments, *International Polymer Processing*, 22(1), 43–60.
14. **Groover M. P.** (2015), *Fundamentals of modern manufacturing*, Willey, 503–510.
15. **Inoue T., Miyata R., Hirai S.** (2016), Antagonistically Twisted Round Belt Actuator System for Robotic Joints, *Journal of Robotics and Mechatronics*, 28(6), 842–853.
16. **Jasiulek P.** (2006), *Joining of plastics by welding, glueing and laminating (in Pollish),* Wydawnictwo 'KaBe', Krosno.
17. **Klimpel A.** (2000), *Welding of termoplastics materials (in Polish),* Wydawnictwo Politechniki Śląskiej, Gliwice.
18. **Krawiec P., Grzelka M., Kroczak J., Domek G., Kołodziej A.** (2019), A proposal of measurement methodology and assessment of manufacturing methods of nontypical cog belt pulleys, *Measurement,* 132, 182–190.
19. **Krawiec P., Waluś K., Warguła Ł., Adamiec J.** (2018), Wear evaluation of elements of V-belt transmission with the application of the optical microscope, *MATEC Web of Conferences,* 157**,** 01009.
20. **Krishnan C., Benatar A.** (2004), Analysis of Residual Stress in Hot Plate Welded Polycarbonate, *ANTEC 2004 Proceedings: Plastics,* 1149–1153.
21. **Kubiak M.** (2019), Prediction of microstructure composition in steel plate heated using high power Yb:YAG laser radiation, *MATEC Web of Conferences,* 254**,** 02023.
22. **Kukla M., Tarkowski P., Górecki J., Malujda I., Talaśka K.** (2015), The Effect of Magnetic Field on Magnetorheological Composites. Artificial Neural Network Based Modelling and Experiments, *Applied Mechanics And Materials,* 816, 327–336.
23. **Madej M., Ozimina D**. (2010), *Plastics and composite materials (in Polish),* Wydawnictwo Politechniki Świętokrzyskiej, Kielce.
24. **Mokhtarzadeh A., Benatar A.** (2012), Experiments with conventional and high temperature hot plate welding of thermoplastics using temperature and pressure control, *ANTEC 2012 Proceedings: Plastics,* 1684–1690.
25. **Nieh J., Lee J.** (1992), Hot Plate Welding of Polypropylene Part I: Crystallization Kinetics, *Polymer Engineering and Science,* 38, 1121–1132.
26. **Potente H., Schneiders J., Bornemann M.** (2002), Theoretical model for the one-dimensional temperature and stress calculation of simple hot plate welded geometries, *Macromolecular Materials and Engineering,* 287**,** 843–853.
27. **Puszka A.** (2006), *Polyurethanes – sources, properties and modifications (in Polish)*, Zakład Chemii Polimerów, Wydział Chemii Uniwersytetu Marii Curie Skłodowskiej w Lublinie, Lublin.
28. **Rzasinski R.** (2017), The algorithm of verification of welding process for plastic pipes, *IOP Conference Series: Materials Science and Engineering,* 227, 012113.
29. **Saternus Z., Piekarska W., Kubiak M., Domański T., Goszczyńska-Króliszewska D.** (2018), Numerical estimation of temperature field in a laser welded butt joint made of dissimilar materials, *MATEC Web of Conferences,* 157**,** 02043.
30. **Saternus Z., Piekarska W., Kubiak M., Domański T., Goszczyńska-Króliszewska D.** (2019), Numerical modeling of cutting process of steel sheets using a laser beam, *MATEC Web of Conferences,* 254**,** 08004.
31. **Sikora R.** (1993), *Polymers processing (in Polish),* Wydawnictwo ŻAK, Warszawa.
32. **Staniszewski B.**, 1979, *Wymiana Ciepła Podstawy Teoretyczne*, PWN, Warszawa.
33. **Taler J., Duda P.**, 2003, *Rozwiązywanie prostych i odwrotnych zagadnień przewodzenia ciepła*, Wydawnictwo Naukowo – Techniczne, Warszawa.
34. **Troughton M**. (1997), *Handbook of Plastics Joining: A practical guide,* Plastics Design Library, New York.
35. **Wałęsa K., Malujda I., Górecki J.** (2020), Experimental research of the mechanical properties of the round drive belts made of thermoplastic elastomer, *IOP Conference Series: Materials Science and Engineering,* 776, 012107.
36. **Wałęsa K., Malujda I., Talaśka K.** (2018), Butt welding of round drive belts, *Acta Mechanica et Automatica,* 12(2), 115–126.
37. **Wałęsa K., Malujda I., Wilczyński D.** (2020), Experimental research of the thermoplastic belt plasticizing process in the hot plate welding, *IOP Conference Series: Materials Science and Engineering,* 776, 012011.
38. **Wałęsa K., Malujda M., Górecki J., Wilczyński D.** (2019), The temperature distribution during heating in hot plate welding process, *MATEC Web of Conferences,* 254**,** 02033.
39. **Wałęsa K., Mysiukiewicz O., Pietrzak M., Górecki J., Wilczyński D.** (2019), Preliminary research of the thermomechanical properties of the round drive belts, *MATEC Web of Conferences*, 254, 06007.
40. **Wanqing L., Changqing F., Xing Z., Youliang C., Rong Y., Donghong L.** (2017), Morphology and thermal properties of polyurethane elastomer based on representative structural chain extenders, *Thermochimica Acta*, 653, 116–125.
41. **Wilczyński D., Malujda M., Górecki J., Domek G.** (2019), Experimental research on the proces of cutting transport belts, *MATEC Web of Conferences*, 254, 05014.
42. **Winczek J., Modrzycka A., Gawrónska E.** (2016), Analytical description of the temperature field induced by laser heat source with any trajectory, *Procedia Engineering*, 149, 553–558.
43. **Wojtkowiak D., Talaśka K.** (2019), Determination of the effective geometrical features of the piercing punch for polymer composite belts, *The International Journal of Advanced Manufacturing Technology*, 104(1-4), 315–332.
44. **Wojtkowiak D., Talaśka K., Malujda I., Domek G.** (2018), Estimation of the perforation force for polimer composite conveyor belts taking into consideration the shape of the piercing punch, *The International Journal of Advanced Manufacturing Technology,* 98(9–12), 2539–2561.
45. **Yousepour A., Hojjari M., Immarigeon J-P.** (2004), Fusion Bonding/Welding of Thermoplastic Composites, *Journal of Thermoplastic Composite Materials,* 17, 303–341.
46. **Żuchowska D.** (2000), *Construction polymers (in Polish),* WNT, Warszawa.

DOI 10.2478/ama-2020-0013

ELECTRO PNEUMATIC CONTROL SYSTEM FOR INVERTED PENDULUM

Krzysztof POPŁAWSKI,[*] Leszek AMBROZIAK,[*] Mirosław KONDRATIUK[*]

[*]Bialystok University of Technology, Faculty of Mechanical Engineering, Department of Robotics and Mechatronics,
ul. Wiejska 45C, 15-351 Białystok, Poland

r.c.drifter.d1@gmail.com, l.ambroziak@pb.edu.pl, m.kondratiuk@pb.edu.pl

received 1 March 2020, revised 30 May 2020, accepted 4 June 2020

Abstract: The paper concerns the inverted pendulum control system with using pneumatic cylinder. A mathematical model of the pendulum used to derive the LQG controller was presented. Prepared laboratory stand was presented and described in detail. The main purpose of the work was experimental researches. A number of control process tests were conducted with variable model parameters such as additional mass, injected disturbances and so on. The results were shown on the time plots of the control object states.

Keywords: Inverted pendulum, pneumatic actuator, inverted pendulum control, encoder, stabilization

1. INTRODUCTION

Inverted pendulums are popular systems used in researches about control theory. Nowadays, there are many configurations of inverted pendulums. The most common pendulums are linear, rotational (Ahangar-Asr et al., 2011), (Ozbek et al., 2010), single (Astrom et al., 2000), double (Sang et al., 2011) or triple (Zhangn et al., 2017) and pendulums on a cart (Hui et al., 2016). Many analyses on inverted pendulums are conducted because of similarity to such objects as robots and rockets (Boubaker, 2012), (Hovingh et al., 2007). These analysis mostly concerned with modern control systems application like genetic adaptive (Moore et al., 2001), with visual feedback (Wang et al., 2008) or another advanced control methods (Zhijun et al., 2013) including sliding mode control and swing up control. There are also investigations concern nonlinear analysis (Boubaker, 2013), (Prasad et al., 2013). Moreover, the actuation systems for inverted pendulums were intensively studied. We can list here the electric, pneumatic or hydraulic drives applied to pendulum stabilization (White et al., 1999).

Controlling an inverted pendulum on a cart requires knowledge on implementing algorithms by means of digital devices and pneumatically actuated systems (Uszynski, 2018). Control systems based on PC stations and separate signal processing cards are very often used. Inverted pendulum systems can be controlled with different types of drives. Pneumatic actuators are not very popular, mainly because of air compressibility and significant mechanical friction (Beater, 2007). On the other hand, high durability and simple maintenance makes pneumatic drives a good replacement for electric drives.

This paper presents design and construction of inverted pendulum system controlled with pneumatic actuator and Arduino Due microcontroller. Second chapter presents the development of cart position and pendulum angle measurement system. In this part, an optical encoder was designed. Computer program that measures pendulum angle was implemented by means of Arduino microcontroller and Simulink environment. In order to define the cart position, a linear potentiometer with analogue output was used. Third chapter presents a design of inverted pendulum control system and output signal amplifier. Next chapter describes the derivation of inverted pendulum mathematical model, which was used to predefine linear-quadratic regulator and Kalman state observer. Design of pneumatically actuated inverted pendulum was developed by means of CAD software. In chapter five, experimental setup and research results in the form of plots (pendulum angle, cart position and control signal) were presented.

2. MEASUREMENT OF CART POSITION AND INVERTED PENDULUM ANGLE

Design and practical implementation of the inverted pendulum system is connected with two main tasks. First is the measurement of the pendulum deflection angle and the second is reading cart position.

2.1. Pendulum angle measuring system design

Optical encoders can be used to measure linear or rotational movement. Generally, we can divide them into two types: incremental or absolute encoders. The main difference between them is the possibility to remember the position after loss of power. Incremental encoder only sends electric pulses – that is why, during every single turning on, the device has to calibrate its position. This drawback can lead into loss of pulses and false position measurement. Absolute encoders use Gray code to define the position and need more optical sensors. This makes them more complicated and more expensive.

To measure the rotation of inverted pendulum, it would be easier to buy an absolute encoder but to reduce cost of made it, an incremental encoder from ink printer was used. The inconvenience will be the need to use the pendulum position calibration, but some software modifications can allow to do it even on running program in microcontroller. The used optical encoder AEDS-

962x (Fig. 1) from Avago Technologies does not require to amplify pulses (Avago, 2006).

It is equipped with two output channels, and thanks to this feature, it is possible to distinguish the direction of pendulum rotation. Microcontroller Arduino Due has a pin that can be used as a 3.3 V power supply for encoder. Each pulse on both channels is detected by means of external interrupts on microcontroller digital inputs. This function guarantees that no pulses will be lost and also that computing power is not harmed.

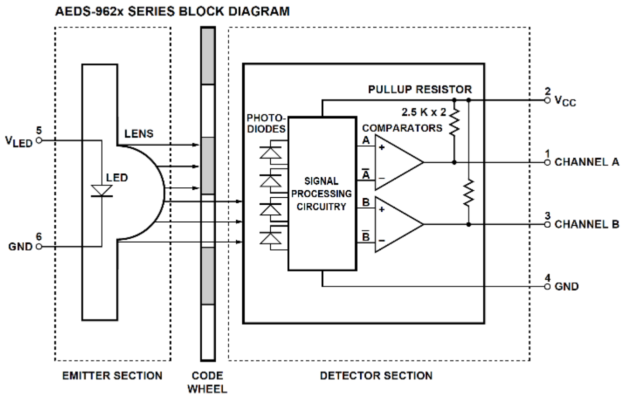


Fig. 1. Optical encoder block diagram (Avago, 2006)

The programming part of the project was realized in MATLAB/Simulink environment (MathWorks, 2017). Creating the program is quite intuitive and Arduino products are supported, but to measure the pendulum angle, a special function in the form of a block had to be developed. This process starts from empty 'S-Function Builder' block that can contain specified parts of C code. Choosing open source library for optical encoders reduced amount of code and complexity of configuration 'S-Function Builder' block. Using external interrupts on both optical encoder channels and high resolution disk resulted in reading pendulum angle by every 0.08°.

2.2. Reading cart position with linear potentiometer

Another necessary variable that had to be measured was linear cart position. Pneumatic system is based on Festo products. That is a way in which a linear encoder MLO-POT-TLF was used. This encoder has one 0-10 V range analogue output. Arduino Due has limited maximal voltage to 3.3 V; therefore, simple voltage divider based on two resistors was made. Each resistance was calculated from simple equation (1) written below:

$$U_{\text{out}} = \frac{R_2}{R_1+R_2} \cdot U_{\text{in}} \tag{1}$$

Reading cart position in Simulink is simpler since there is ready *'Analog Input'* block that has two configurable parameters: pin number and sampling time. Pendulum and cart positions can be analysed in real time using *'Time Scope'* block.

3. INVERTED PENDULUM CONTROL SYSTEM DESIGN

Electro pneumatic diagram (Fig. 2) was developed in FluidSIM software. The system consists of air preparation unit (0.1), rodless cylinder (1.0) and proportional directional control valve (1.2).

Depending on the control voltage applied to the proportional valve, its position can change. Spool position determines the speed and direction of the cylinder slider movement. Prechosen Festo DGPL cylinder has a very solid slider that allows to make simple and easily mounted construction with pendulum.

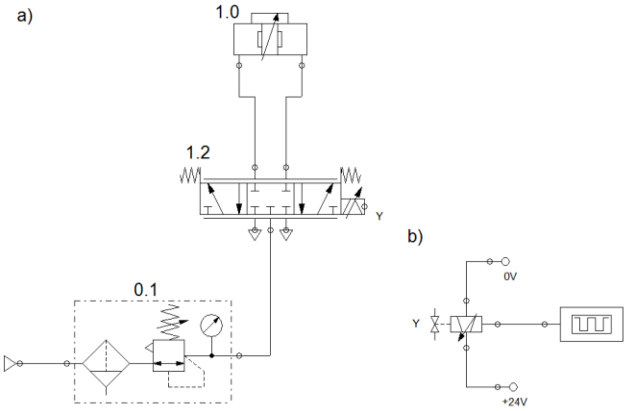


Fig. 2. Electro pneumatic system: a) pneumatic part, b) electrical part

The proportional valve MPYE-5 can be controlled with 0–10 V range analogue signal. Arduino Due has two analogue output pins but they are not standardized with industrial devices. These outputs have a range from 0.55 V to 2.75 V and 12-bit DAC converter. In order to achieve 0–10 V control signal, a simple differential amplifier was realized. Using an operational amplifier allowed to remove the 0.55 V offset and amplify the output signal in order to reach the desired value. Fig. 3 shows the diagram of the circuit.

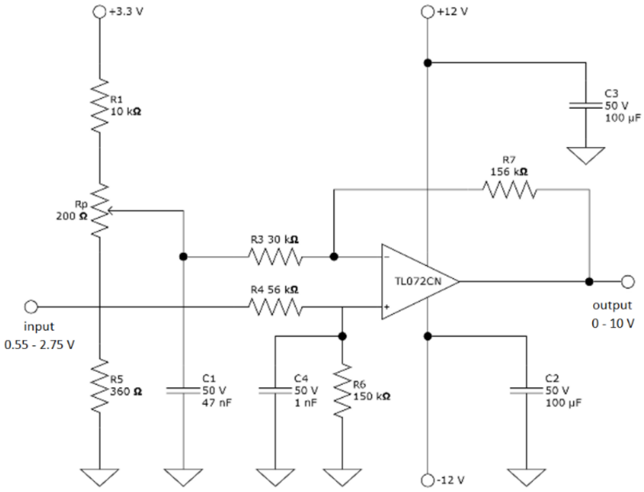


Fig. 3. Differential amplifier circuit diagram

The amplifier is based on integrated circuit TL072CN, which has to be supplied with symmetrical voltage. This small drawback also has a good effect because only this kind of amplifiers can reach exact 0 V in this application. Resistor R4 makes sure that current on Arduino output is limited and capacitor C4 reduces eventual possibility of noises in the signal.

The prepared electrical circuit was specifically developed in the form of a PCB (printed circuit board) using thermal transfer method. Fig. 4 presents the final effect of the circuit that consisted of the amplifying part and additional four voltage dividers adjusted for sensors with 0–10 V output signals.

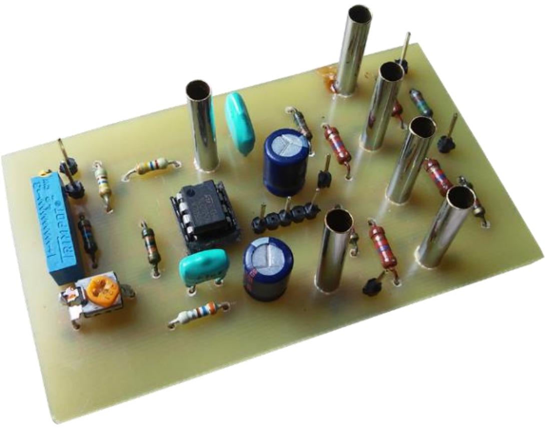


Fig. 4. Developed control signal amplifier and voltage dividers

Controlling analogue output in Simulink is similar to reading signals. This can be done with one already provided block that has one configurable parameter, which is DAC pin number. In order to be able to control, the proportional valve signal had to be biased to 5 V (valve closed position) and limited to 0–10 V range.

4. INVERTED PENDULUM MODEL AND CONTROLLER DESIGN

Fig. 5 presents the physical model of the inverted pendulum where: M – cart mass [kg], m – pendulum mass [kg], b – damping coefficient in linear motion [N·s/m], b_p – damping coefficient in rotational motion [N·s/rad], I – pendulum mass moment of inertia [kg·m^2], x – cart coordinate [m], θ – pendulum deflection angle [rad], l – pendulum half-length [m], g – acceleration of gravity [m/s^2], F – cylinder applied force [N].

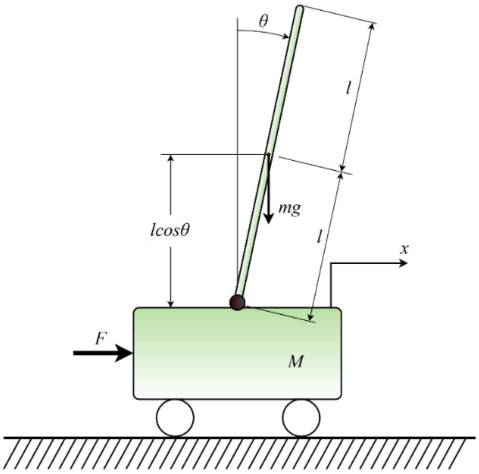


Fig. 5. Physical model of inverted pendulum

It is an unstable system in vertical pendulum position that can rotate freely in one plane. Applying force F to the cart allows to control the pendulum position.

After the calculation of forces in the inverted pendulum, the system of nonlinear equations (2) and (3) was obtained:

$$(M+m)\ddot{x} + b\dot{x} + ml\ddot{\theta}cos\theta - ml\dot{\theta}^2 sin\theta = F \tag{2}$$

$$(I+ml^2)\ddot{\theta} + ml\ddot{x}cos\theta = mglsin\theta - b_p\dot{\theta} \tag{3}$$

In order to linearize these equations, a model operating point was assumed in the top pendulum position. Assuming small angle deflections, we have the approximations: $sin\theta \approx \theta, cos\theta \approx 1$, $\dot{\theta}^2 \approx 0$. Now linearized equations can be described with equation (4) and (5):

$$(M+m)\ddot{x} + b\dot{x} + ml\ddot{\theta} = F \tag{4}$$

$$(I+ml^2)\ddot{\theta} + ml\ddot{x} + b_p\dot{\theta} = mgl\theta \tag{5}$$

Based on the above equations, we can have the following state-space model described with following equations:

$$\dot{\mathbf{x}} = \mathbf{Ax} + \mathbf{Bu}, \tag{6}$$

$$\mathbf{y} = \mathbf{Cx} + \mathbf{Du}, \tag{7}$$

where:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} \theta \\ \dot{\theta} \\ x \\ \dot{x} \end{bmatrix},$$

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{(M+m)mgl}{q_1} & \frac{-(M+m)b_p}{q_1} & 0 & \frac{mlb}{q_1} \\ 0 & 0 & 0 & 1 \\ \frac{mlb}{q_1} & \frac{mlb_p}{q_1} & 0 & \frac{-(I+ml^2)b}{q_1} \end{bmatrix},$$

$$\mathbf{B} = \begin{bmatrix} 0 \\ \frac{-ml}{q_1} \\ 0 \\ \frac{(I+ml^2)}{q_1} \end{bmatrix},$$

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

$$\mathbf{D} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

and $q_1 = (M+m)I + Mml^2$.

For stabilizing inverted pendulum system dynamics, the LQR (Linear-Quadratic-Regulator) controller was developed and implemented to ensure local stability of the control plant. LQR approach is often used and mentioned to stabilize a structurally unstable system like inverted pendulum. LQR controller uses measurements **y** of control plant to generate a control signal **u** that controls **y**. The LQR regulator minimizes the cost function:

$$J = \int_0^\infty [\mathbf{x}(t)\mathbf{Qx}(t) \ + \mathbf{u}(t)\mathbf{Ru}(t)]dt \tag{8}$$

The state control law we can be written as follows:

$$\mathbf{u} = -\mathbf{kx} \tag{9}$$

The closed loop system can be determined with equation (9) as:

$$\dot{\mathbf{x}} = [\mathbf{A} - \mathbf{Bk}]\mathbf{x} \tag{10}$$

In addition to the state-feedback gain **k**, LQR returns the solution **S** of the associated Riccati equation described as:

$$\mathbf{A^TS}+\mathbf{SA}-\mathbf{SBR^{-1}B^TS}+\mathbf{Q}=0 \tag{11}$$

where **k** is derived from **S** using following formula:

$$\mathbf{k}=\mathbf{R^{-1}B^TS} \tag{12}$$

In order to determine **Q** and **R** values, Bryson's rule was used.

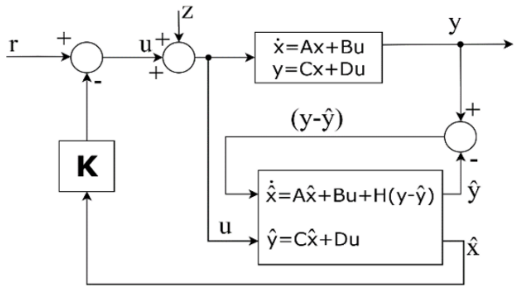


Fig. 6. Implemented control system

During the experimental setup, we cannot measure all the state vector elements. That is why, we have to estimate them. In this purpose, a state observer was designed. During this process, it is important to place poles on the left from the observed system poles. Combining LQR regulator and Kalman filter resulted in obtaining the final control algorithm for inverted pendulum system called LQG (Bryson et al., 1969). Fig. 6 presents control system.

Simulation results are presented in Fig. 7. Initially, the pendulum is deflected 3° from the vertical position and at the seventh second, an external disturbance force (100 N) is applied in order to check the control system robustness.

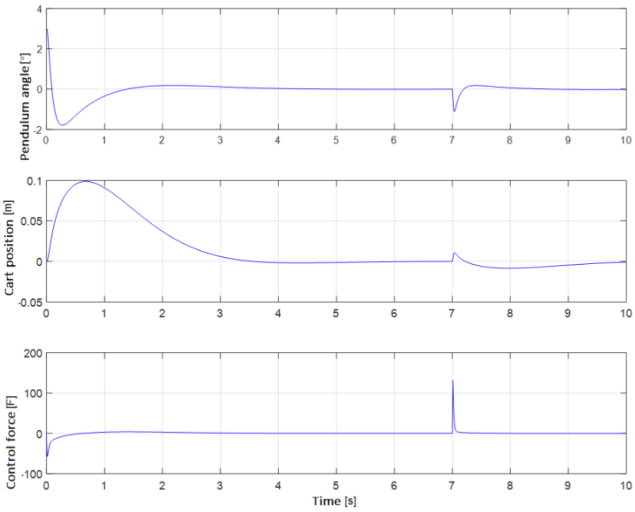


Fig. 7. Simulation results in form of a plots

Achieved results do not exceed the predicted values, which means – in this way, the control design can be considered successful.

5. EXPERIMENTAL SETUP AND RESULTS

Experimental setup is presented in Fig. 8 consists of: 1 - air compressor, 2 - air preparation unit, 3 - directional proportional control valve, 4 - rodless cylinder, 5 - linear potentiometer, 6 - optical encoder, 7 - inverted pendulum module, 8 - external power supply, 9 - signal amplifier, 10 - microcontroller, 11 - PC station.

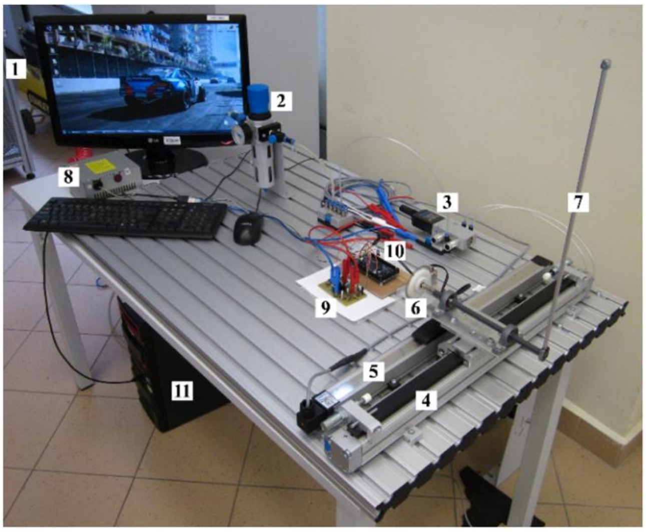


Fig. 8. Experimental setup of pneumatically actuated inverted pendulum system

All the elements are also included in the diagram in Fig. 9.

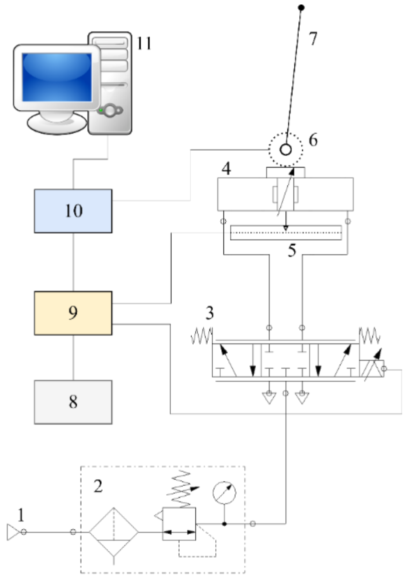


Fig. 9. Experimental system diagram

Serial transmission between PC and microcontroller allows to draw plots, save data and change online parameters during the experiment. Arduino Due can also be programmed as an independent device to stabilize the inverted pendulum, but then there is no interface to interact in such a mode.

During the experiment, the developed inverted pendulum system was tested with external disturbances in the form of a push and step change of set cart position. Each experiment was conducted at full air compressor tank at 0.8 MPa. Then the air pressure was reduced to 0.5 MPa.

In Fig. 11–14, the system responses for external disturbance are presented. After the disturbance, the cart is moving in order to stabilize the pendulum. Oscillations on the plots are the result of significant static friction during linear cylinder movement. Inverted pendulum system needs about 3 seconds to reach a stable position, maximal deflection angle was around 3° and the cart had to travel approximately 80 mm. Control signal amplitude is fading as pendulum is reaching a stable position.

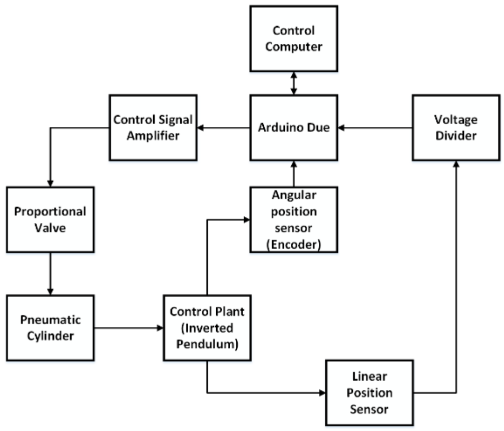


Fig. 10. Signals flow on the test rig

Tab. 1. System parameters

	Parameter value		Parameter value
Long pendulum without additional load	$m = 32\ g$	**Long pendulum with additional load**	$m = 37\ g$
	$l = 215\ mm$		$l = 215\ mm$
	$I = 0,002\ kgm^2$		$I = 0,0029\ kgm^2$
	$b_p = 0,013\ \frac{Nms}{rad}$		$b_p = 0,009\ \frac{Nms}{rad}$
Short pendulum without additional load	$m = 26\ g$	**Short pendulum with additional load**	$m = 31\ g$
	$l = 175\ mm$		$l = 175\ mm$
	$I = 0,0011\ kgm^2$		$I = 0,0016\ kgm^2$
	$b_p = 0,018\ \frac{Nms}{rad}$		$b_p = 0,015\ \frac{Nms}{rad}$
$M = 2,7\ kg$		$g = 9,81\ \frac{m}{s^2}$	$b = 65\ \frac{Ns}{m}$

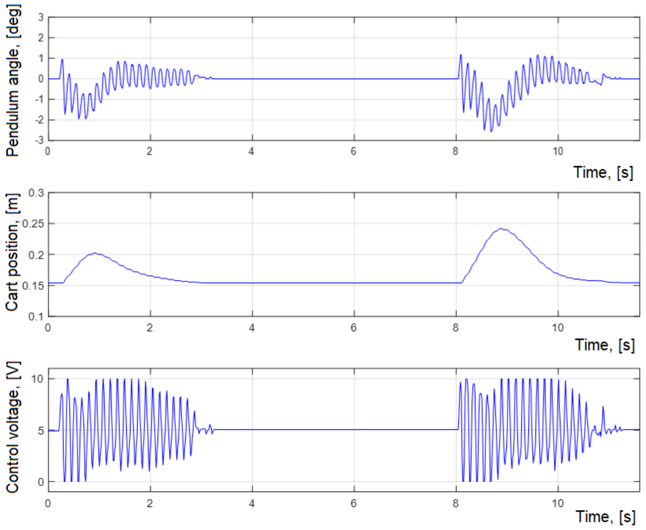


Fig. 11. Experimental results for long pendulum without additional load (external disturbance system response)

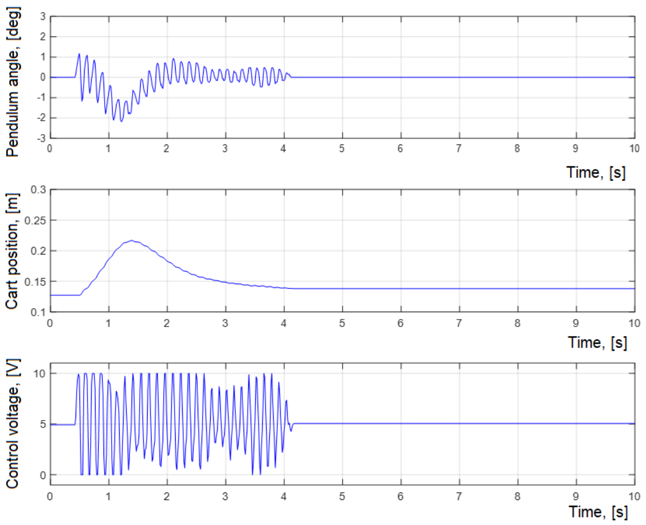


Fig. 12. Experimental results for long pendulum with additional load (external disturbance system response)

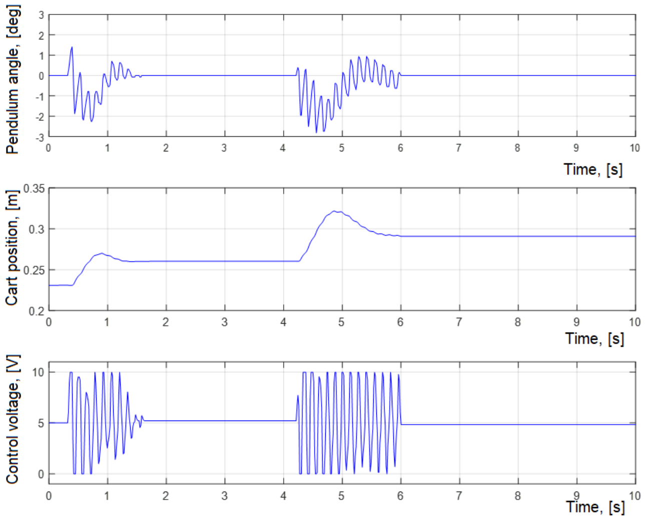


Fig. 13. Experimental results for short pendulum without additional load (external disturbance system response)

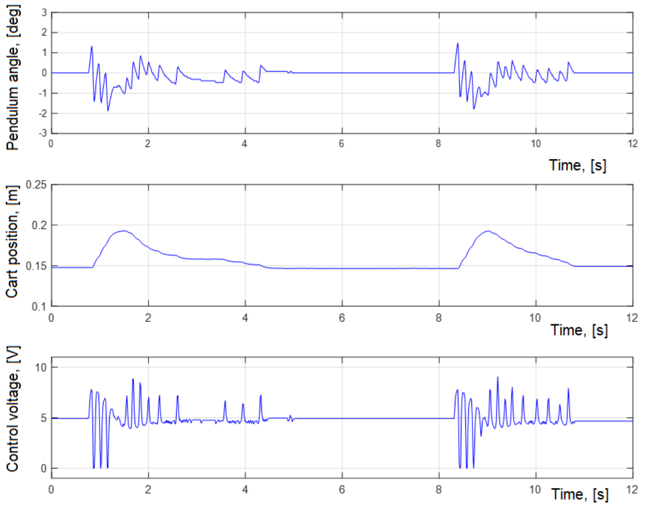


Fig. 14. Experimental results for short pendulum with additional load (external disturbance system response)

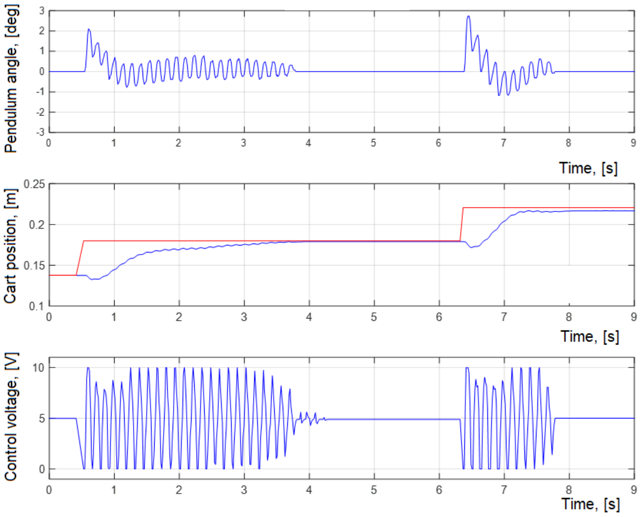


Fig. 15. Experimental results for long pendulum without additional load (step system response)

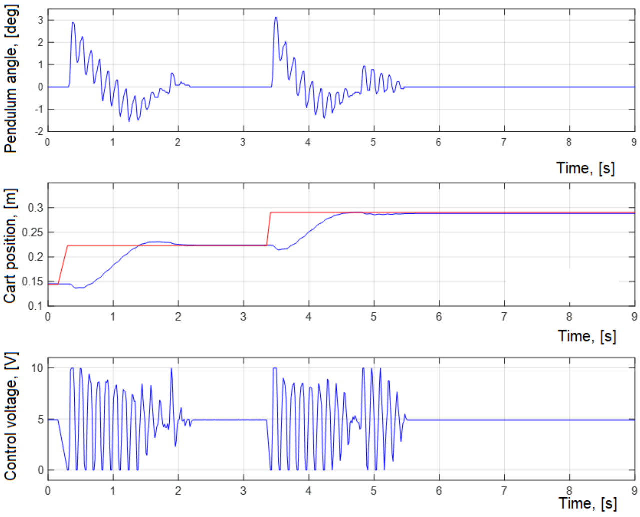


Fig. 16. Experimental results for long pendulum with additional load (step system response)

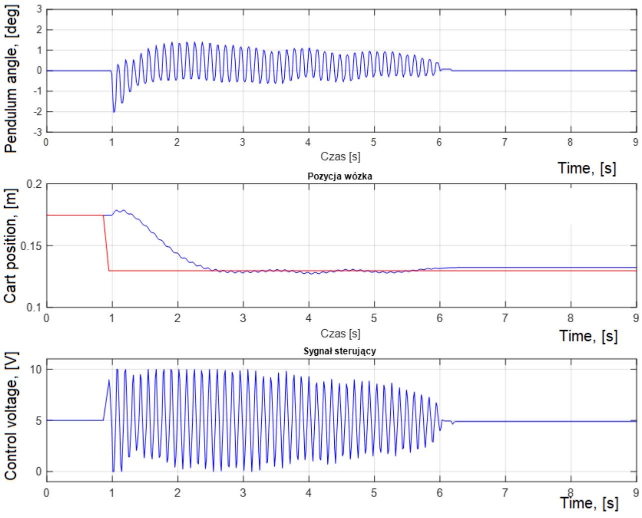


Fig. 17. Experimental results for short pendulum without additional load (step system response)

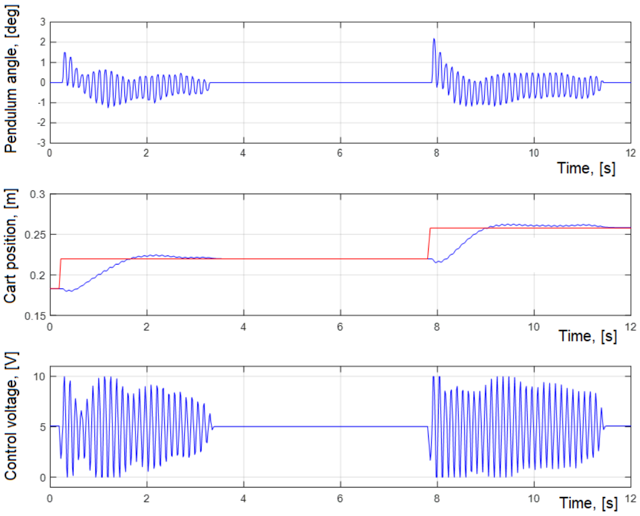


Fig. 18. Experimental results for short pendulum with additional load (step system response)

Next part of the experiment was step response by changing cart position. Fig. 15–18 present how system reacts to this kind of disturbance. The moment after cart set position changes, a pendulum is oscillating around its top vertical position, and at the same time, the cart is moving to the desired location. Also, this experiment resulted in average 3 seconds settling time and maximal deflection angle around 3°.

In Tab. 2, all the obtained results from the conducted experiments are shown. It contains setting times for external disturbance and step responses, maximal deflections and evolving error values for all tests.

Tab. 2. Control quality parameters

	Long pendulum without additional load	**Long pendulum with additional load**	**Short pendulum without additional load**	**Short pendulum with additional load**
Setting time (disturbance), [s]	3,12	3,84	1,58	2,25
Setting time (step response), [s]	3,41	2,14	5,67	4,02
Maximal deflection, [°]	3,05	2,83	3,36	3,69
Error, [cm]	0,16	0,29	0,31	0,17

6. CONCLUSIONS

The main purpose of the controlling inverted pendulum with pneumatic actuator was achieved. Difficulties during the development applied mainly to the programming part. One of them was reading pendulum angle using Simulink environment. The significant advantage of Arduino microcontroller is the possibility to control the inverted pendulum without a PC and any additional software.

The presented solution brings opportunity to research pneumatic drives in closed loop systems and thanks to that, the area of their application it can be broaden. Developed control algorithm can be successfully and easily adapted for inverted pendulum system with electric drive. Simultaneous use of MATLAB and microcontrollers allow relatively convenient verification of calculations without knowledge and deep analyse of C code. Developed amplifying device makes Arduino Due microcontroller much more applicable in automation industry. During the experiments, it was observed that cylinder mechanical friction is mainly responsible for pulling effect at the time of pendulum stabilization. Another big advantage of Arduino microcontroller is major cost reduction in comparison with the building control system based on DAQ cards. Pneumatically actuated inverted pendulum system would be much more robust if at least twice as long cylinder and bigger, heavier pendulum were used.

REFERENCES

1. **Ahangar-Asr H., Teshnehlab M., Mansouri M., Pazoki A. R.** (2011), A Hybrid Strategy for the Control of Rotary Inverted Pendulum, International Conference on Electrical and Control Engineering (ICECE), 16–18 September, Yichang, China, 5656–5659.
2. **Astrom K. J., Furuta K.** (2000), Swing up a pendulum by energy control, Automatica, Vol. 36, No. 2, 287–295.
3. **Avago Technologies** (2006), AEDS-962x for 150 LPI Ultra Small Optical Encoder Modules, May.
4. **Beater P.** (2007), Pneumatic Drives, System Design, Modelling and Control, Springer.
5. **Boubaker O.** (2012), The Inverted Pendulum: A Fundamental Benchmark in Control Theory and Robotics, IEEE International Conference on Education and e-Learning Innovations, Jul.
6. **Boubaker O.** (2013), The Inverted Pendulum Benchmark in Nonlinear Control Theory: A Survey, International Journal of Advanced Robotic Systems, Vol. 10, 233:2013.
7. **Bryson A. E., Ho Y. C.** (1969), Applied Optimal Control, Blaisdell.
8. **Hovingh A., Roon M.** (2007), Design and Control of an Inverted Pendulum, Western Michigan University, Department of Aeronautical and Mechanical Engineering.
9. **Hui L., Min Z., Chen G.** (2016), Cloud-Model PID Control of Double Inverted Pendulum Based on Information Fusion, 35th Chinese Control Conference, Jul. 27–29.
10. **MathWorks** (2017), Matlab Control System Toolbox User's Guide.
11. **Moore M. L., Musacchio J. T., Passino K. M.** (2001), Genetic adaptive control for an inverted wedge: experiments and comparative analyses, Engineering Applications of Artifcial Intelligence, Vol. 14, 1–14.
12. **Ozbek N. S., Efe M. O.** (2010), ”Swing up and Stabilization Control Experiments for a Rotary Inverted Pendulum - An Educational Comparison”, International Conference on Systems Man and Cybernetics (SMC), 10-13 October, Istanbul, Turkey, 2226–2231.
13. **Prasad L. B., Tyagi B., Gupta H. O.** (2014), Optimal Control of Nonlinear Inverted Pendulum System Using PID Controller and LQR: Performance Analysis Without and With Disturbance Input, International Journal of Automation and Computing 11(6), Dec., 661–670.
14. **Sang Y., Fan Y., Liu B.** (2011), Double Inverted Pendulum Control Based on Three-loop PID and Improved BP Neural Network, Second International Conference on Digital Manufacturing and Automation (ICDMA), 5-7 August, Zhangjiajie, Hunan, China, 456–459.
15. **Uszynski S., Ambroziak L., Kondratiuk M., Kulesza Z.** (2018), Air Consumption Analysis in Compressed Air Powered Vehicles, 23rd International Conference on Methods and Models in Automation & Robotics, MMAR 2018 8486124, 837–842.
16. **Wang H., Chamroo A., Vasseur C., Koncar V.** (2008), Stabilization of a 2-DOF Inverted Pendulum by a Low Cost Visual Feedback, American Control Conference, Seattle, 11–13 Jun.
17. **White W. N., Fales R. C.** (1999), Control of a Double Inverted Pendulum with Hydraulic Actuation: A Case Study. American Control Conference, San Diego, Jun.
18. **Zhangn X. L., Dai J. H., Cheng Y. T., Hao S., Li J. H.** (2017), Nonlinear control of triple inverted pendulum based on T-S cloud inference network, Control And Decision Conference (CCDC), 28-30 May, Chongqing, China, 3290–3295.
19. **Zhijun L., Chenguang Y., Liping F.** (2013), Advanced Control of Wheeled Inverted Pendulum Systems, Springer.

The research work has been carried out in the framework of work No. WZ/WM-IIM/1/2019 and financed from the funds for science by the Polish Ministry of Science and Higher Education.

DOI 10.2478/ama-2020-0014

NATURAL CONVECTION FROM FOUR CIRCULAR CYLINDERS IN ACROSS ARRANGEMENT WITHIN HORIZONTAL ANNULAR SPACE

Houssem LAIDOUDI*

*Laboratory of Sciences and Marine Engineering, Faculty of Mechanical Engineering,
USTO-MB, BP 1505, El-Menaouer, Oran, 31000, Algeria

houssem.laidoudi@univ-usto.dz

received 18 April 2020, revised 24 June 2020, accepted 26 June 2020

Abstract: Numerical investigation is accomplished to study the roles of governing parameters of natural convection on the fluid motion and heat transfer rate of four heated circular cylinders placed inside a circular enclosure of cold surface. The cylinders are positioned in across arrangement. The representative results are obtained within the ranges of initial conditions as: Prandtl number (Pr = 7.1 to 1000) and Rayleigh number (Ra = 10^3 to 10^5). The average Nusselt number of each inner cylinder is computed. The effects of thermal buoyancy strength on the fluid motion and temperature are also illustrated. It was found that the heat transfer rate of cylinders depends significantly on the position inside the enclosure. Moreover, the role of Prandtl number on flow and thermal patterns is negligible. The values of Nusselt number are also given, which can be useful for some engineering applications.

Key words: Natural convection, multiple cylinders, annular space, heat transfer, numerical simulation

1. INTRODUCTION

The heat transfer inside an annular space is an important phenomenon that can be widely encountered in many industrial applications such as heat exchangers, refineries, cooling small electronic devices, marine engineering applications, chemical processes and so on. Therefore, an important number of scientific researches have been devoted recently to understand the governing parameters on the fluid flow and thermal patterns. For example, Masoumi et al. (2019) studied the laminar natural convection between two horizontal cylinders. The cylinders were placed in a concentric manner. The fluid used for this work was yield stress fluid. The governing parameters of this research were limited for Ra = 10^3 to 10^6. Pandey et al. (2019) presented an extensive review of the roles of thermo-physical characteristics of the fluid and the geometrical forms of the inner bodies on natural convection inside square cavity. The studied characteristics are modulated by the dimensionless numbers of Rayleigh and Prandtl. Aly (2017) investigated the natural convection heat transfer from two circular cylinders placed in tandem arrangement within rectangular cavity. The annular space for this research was assumed to be porous. The research examined the porous characteristics and the cylinder dimensions on the rate of heat transfer. Hussein (2013) numerically examined the effects of geometrical modifications of outer square cavity on the heat transfer rate of circular cylinder placed in the centre of this outer cavity.

It can be seen that the heat transfer in annular space depends expensively on the geometrical configuration of heated or cold organs and the thermo-characteristics of the fluid. For this reason, several researchers have worked on these aspects in order to enhance the heat transfer rate in annular configurations.

Matin and Khan (2013) numerically studied the thermal buoyancy of natural convection in annular space. The studied geometry consisted of two horizontal concentric cylinders of circular shape. The computed domain was assumed to be filled with non-Newtonian power fluids. The controlling parameters of Prandtl and Rayleigh numbers were studied for Pr = 0.7 to 1000 and Ra = 10^3 to 10^5. The results showed that the first category of power-law fluids (shear thinning) having a tendency to improve the rate of heat transfer. On the other hand, the shear thickening fluids reduced the heat transfer rate. Abu-Nada et al. (2008) for the same the geometry of Matin and Khan (2013) used nanofluids instead of base-fluid. They confirmed that the nanofluids enhance the thermo-physical characteristics of base-fluid, and consequently, improve the heat transfer rate. Nada and Said (2019) found that there is a possibility to increase the heat transfer rate of single cylinder placed within cold outer cylinder. The solution was added to some radial fins to the wall of inner cylinder. El-Maghlany et al. (2016), Char and Lee (1998) and Nasiri et al. (2017) used the nanofluids as convective medium to transfer heat between two horizontal cylinders. The cylinders were arranged in an eccentric manner. They found that the eccentricity factor of inner cylinder increases the heat transfer rate. Iqbal et al. (2017), Arbaban and Salimpour (2014), Ha et al. (2004), Hadidi et al. (2020) also used the same idea to increase the effect of heat transfer. Laidoudi (2020) studied the natural convection in enclosed cold cavity in which two hot circular cylinders were placed in tandem. Laidoudi et al. (2020) also studied a numerical investigation on natural convection within an annular space of a special shape of horizontal inner cylinder. It was found that the thermo-physical proprieties of fluids have a limited effect on convective heat transfer. Eid (2011) performed an experimental study on free convection in annulus of elliptical cross-section. Kozlov (2018) numerically examined the effect of rotational vibrations on flow and thermal behaviours in annulus. Sheikhzadeh et al. (2013) numerically combined the effects of radial fins of inner cylinder with nanofluid characteristics to improve the heat transfer in an annulus of circular form.

From these mentioned papers and others such as Zhang et al. (1992), Kuehn and Goldstein (1976) and Ho et al. (1989), the analyses of fluid flow inside the annulus showed that the apparition of two loops and the form of these loops as well as their main centre depend mainly on the value of Rayleigh number, that is, the thermal buoyancy strength.

After this presentation of previous researches on laminar natural convection in annular space, it can be concluded that the heat transfer rate of inner cylinders is significantly affected by the shapes of the domain and thermo-characteristics of fluid. Although there are many researches that cover the buoyancy-driven flow in annular space, few papers studied the multiple arranged cylinders in circular enclosure under the effect of thermal buoyancy. For that reason, the present paper is an attempt to add some insights about the effects of mentioned parameters on the natural convection heat transfer from four circular cylinders placed in across arrangement within cold circular cylinder. The results are analysed and interpreted for the range of the following parameters: Ra = 10^3 to 10^5 and Pr = 7.1 to 1000. The evolutions of heat transfer rate are quantified by the value of average Nusselt number.

2. PHYSICAL DOMAIN AND MATHEMATICAL MODELLING

The studied problem is schematized in Fig. 1. It involves a cold circular enclosure (T_c) of a diameter H. Inside this circular cavity, four circular cylinders are placed in across arrangement. The cylinders are hot (T_h) and they have a similar diameter (d). This diameter (d) is defined by the blockage ratio B = d/H = 0.2. The gap distance between the centres of parallel cylinders is given by the value S = 0.5 H. Effectively, the studied domain is a symmetrical arrangement. The annular space between inner cylinders and the outer cavity is assumed to be filled with incompressible Newtonian fluid. Due to temperature difference between the surfaces of inner cylinders and outer cylinder that inducesthermal buoyancy force, which is the source of fluid motion.

These kinds of physical phenomena are mostly modelled by the governing equations of continuity, Navier-Stockes and energy. And the effect of buoyancy-driven flow is treated by Bousssinesq approximation.

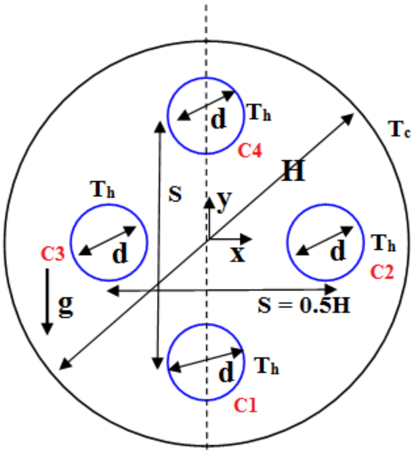


Fig. 1. Schematic view of studied geometry

The governing equations for tow-dimensional flow in Cartesian coordinate systems can be written in dimensionless form as:

$$\frac{\partial u^{\bullet}}{\partial x^{\bullet}}+\frac{\partial v^{\bullet}}{\partial y^{\bullet}}=0 \quad (1)$$

$$u^{\bullet}\frac{\partial u^{\bullet}}{\partial x^{\bullet}}+v^{\bullet}\frac{\partial u^{\bullet}}{\partial y^{\bullet}}=-\frac{\partial p^{\bullet}}{\partial x^{\bullet}}+\mathrm{Pr}\left(\frac{\partial^2 u^{\bullet}}{\partial x^{\bullet 2}}+\frac{\partial^2 u^{\bullet}}{\partial y^{\bullet 2}}\right) \quad (2)$$

$$u^{\bullet}\frac{\partial v^{\bullet}}{\partial x^{\bullet}}+v^{\bullet}\frac{\partial v^{\bullet}}{\partial y^{\bullet}}=-\frac{\partial p^{\bullet}}{\partial y^{\bullet}}+\mathrm{Pr}\left(\frac{\partial^2 v^{\bullet}}{\partial x^{\bullet 2}}+\frac{\partial^2 v^{\bullet}}{\partial y^{\bullet 2}}\right)+Ra.\mathrm{Pr}.T^{\bullet} \quad (3)$$

$$u^{\bullet}\frac{\partial T^{\bullet}}{\partial x^{\bullet}}+v^{\bullet}\frac{\partial T^{\bullet}}{\partial y^{\bullet}}=\left(\frac{\partial^2 T^{\bullet}}{\partial x^{\bullet 2}}+\frac{\partial^2 T^{\bullet}}{\partial y^{\bullet 2}}\right) \quad (4)$$

The number Ra and Pr are given by following expressions, respectively:

$$Ra=\frac{g\beta(T_h-T_c)d^3}{\nu\alpha},\ \mathrm{Pr}=\frac{\nu}{\alpha} \quad (5)$$

The written variables in governing equations refer to: *u* and *v* are velocity components along the *x* and y-directions, respectively; *p* is the pressure; and *T* is the temperature. These variables are written in dimensionless form as:

$$x^{\bullet}=\frac{x}{d},y^{\bullet}=\frac{y}{d},u^{\bullet}=\frac{ud}{\alpha},v^{\bullet}=\frac{vd}{\alpha},p^{\bullet}=\frac{pd^2}{\rho\alpha^2},T^{\bullet}=\frac{T-T_c}{T_h-T_c} \quad (6)$$

The studied fluid is defined by its thermo-physical characteristics, which are given by density *ρ*, kinematic viscosity *ν*, thermal diffusivity *α* and the expansion coefficient *β*. However, *g* refers to gravitational acceleration.

The average Nusselt number of each inner cylinder is calculated by integrating the local Nusselt number along the surface of the cylinder. The expressions for the Nusselt number are:

$$Nu_l=\left.\frac{\partial\phi}{\partial n}\right|_{wall} \quad (7)$$

$$Nu=\frac{1}{s}\int_0^s Nu_l\,ds \quad (8)$$

where *n* and *s* are the direction normal to the wall.

Due to numerical considerations, appropriate boundary conditions are subjected to domain extremities as:

Around the outer cylinder; cold surface with no-slip boundary layer as:

$$u^{\bullet}=0,v^{\bullet}=0,T^{\bullet}=0 \quad (9)$$

On the surfaces of inner cylinders; hot surfaces with no-slip boundary layer as:

$$u^{\bullet}=0,v^{\bullet}=0,T^{\bullet}=1 \quad (10)$$

3. GRID INDEPENDENCY AND VALIDATION TESTS

The governing equations for present work are non-linear partial differential. Therefore, the suitable commercial code ANSY-CFX is used to solve them numerically. The software transforms the equations into matrix system under the aspect of finite-volume method. The convective terms of matrix system is solved by High resolution discretization scheme. However, SIMPLEC (Semi-Implicit Method for Pressure-Linked Equations-Consistent) is used for velocity-pressure coupling. The obtained results from these equations are considered when the relative error of the equations becomes less than 10^{-8} for continuity and momentum equations and 10^{-6} for energy equation.

The grid for present geometry is meshed by using the software Gambit. The elements of grid are non-uniform and unstructured. The elements of the grid are concentrated near the walls of cylinders where thermal and flow layers are very thin. Fig. 2 shows the grid for present investigation. The grid contains 74,962 elements. This value was considered after examining the grid indecency test on obtained results. This step is obviously represented in the results of Tab. 1.

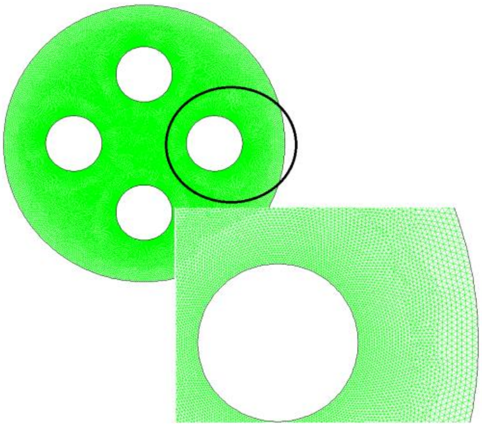


Fig. 2. Typical grid used for present investigation

Tab. 1. Results of grid independency test for Pr = 10 and Ra = 10^4

Mesh	Elements	Nu (C1)	Variation (%)
M1	10324	3.79345	0.74
M2	32980	3.76674	0.25
M3	74962	3.7575	-

This section also incorporates the validation test. This step is shown to prove the accuracy and precision of the used numerical method. To do so, we have executed the same geometry studied by Kuehn and Goldstein (1976) and Matin and Khan (2013). The comparison between all results is depicted in Fig. 3. In fact, Fig. 3 shows the variation of average Nusselt number versus Rayleigh number of a single circular cylinder placed at the centre of circular enclosure for Pr = 1. Fig. 3 shows a good agreement between all the results. The relative error between the present results and the previous works is almost less than 2%.

4. RESULT AND DISCUSSION

The present investigation is conducted to understand the roles of Rayleigh number and Prandtl number on laminar natural convection in annular space. The studied domain consists of four equal-sized cylinders in across arrangement confined in circular enclosure. The inner cylinders are heated with constant temperature (T_h) and the outer cylinder is kept cold with constant temperature (T_c). The computational domain is assumed to be filled with incompressible Newtonian fluid. The circulation of flow inside the domain and temperature distributions are depicted as representative streamline and isotherm contours. However, the average Nusselt number is graphically plotted versus the studied parameters.

Fig. 4 and 5 shows the isotherms and streamlines inside the annular space with different values of Rayleigh and Prandtl numbers. It can be noticed that the particles of the fluid around the heated cylinders become hot, and consequently, the fluid density in these regions becomes lighter, which leads to the movement of lighter regions towards the upper part of the cavity. On the other hand, the fluid particles by the curvature of cold cavity becomes heavier and they move down towards the lower part of circular enclosure. This successive movement of fluid generates two symmetrical loops in the annular space as it is showing Fig. 5 for Ra = 10^3.

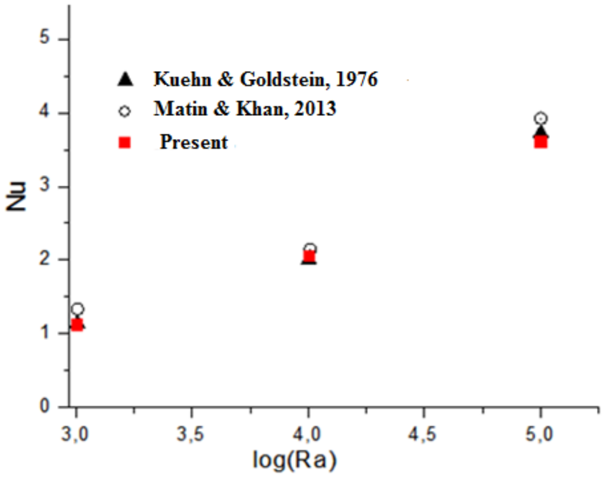


Fig. 3. Validation test

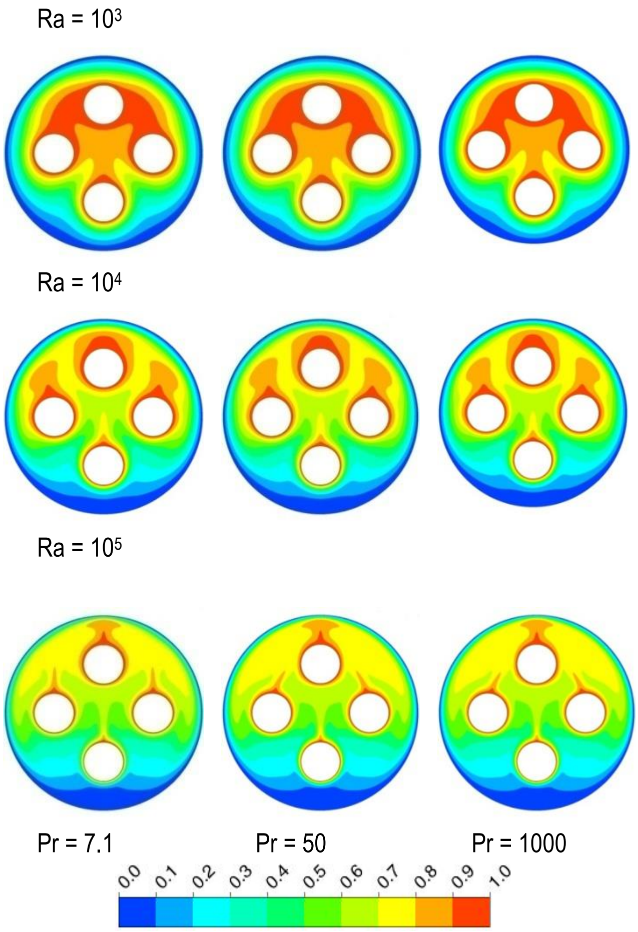


Fig. 4. Isotherm contours for three values of Prandtl and Rayleigh numbers

It is also shown that an increase in the value of Rayleigh number leads to an increase in the strength of thermal buoyancy, and as a consequence, the centre of main loops shifts upwards. Moreover, the increase in the value of Rayleigh number creates a complicated flow inside the domain. Two extra small loops appear above the lower cylinder for Ra = 10^5. For all cylinders, the lower part of cylinder has remarkable isotherms in comparison to the upper part. These thin regions of isotherms indicate that the thermal gradient is very important, which results in an important rate of heat transfer in these zones. Furthermore, the isotherms show that lateral distributions of isotherms decrease with an increase in the Rayleigh number, hinting that the heat transfer of inner cylinders increases with increasing Rayleigh number. The influence of Prandtl number on streamlines and isotherms is almost negligible. Finally, it can be expected that the lower cylinder (C1) has the highest value of heat transfer rate, whereas the upper cylinder (C4) has the lowest value of heat transfer rate. Also, the plumes that are shown on cylinders show the direction of the flow.

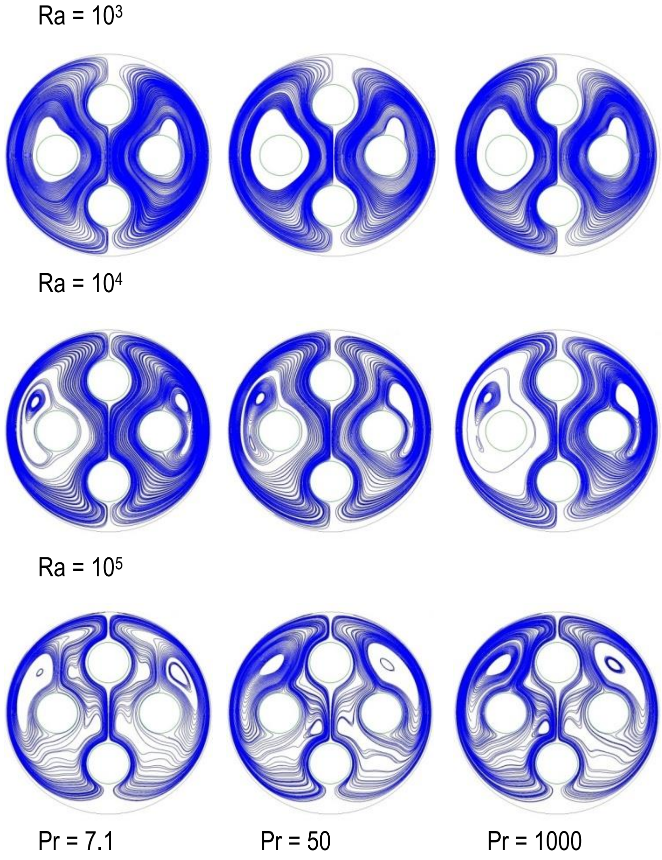


Fig. 5. Streamline contours for three values of Prandtl and Rayleigh numbers

In order to elucidate additionally the flow inside the annular space, Fig.6 depicts the evolution of dimensionless competent velocity profiles of y-direction along the line x for y =0. The velocity profiles indicate that an increase in the value of Rayleigh number increases the velocity. Also, the effect of Prandtl number is almost negligible. The negative values of velocity are close to the outer cylinder, whereas, the positive values are in the vicinity of inner cylinders. The maximum value of velocity is induced by cold source.

Fig. 7 shows the average Nusselt number of inner cylinders as a function of Prandtl number and Rayleigh number. As it was expected, the lower cylinder has the highest values of Nusselt number. Generally, as we move from the lower cylinder (C1) to the upper cylinder (C4), the average Nusselt number decreases. Increase in Rayleigh number increases the average Nusselt number of all cylinders. On the other hand, with an increase in Prandtl number, the values of Nusselt number remains constant for all cylinders.

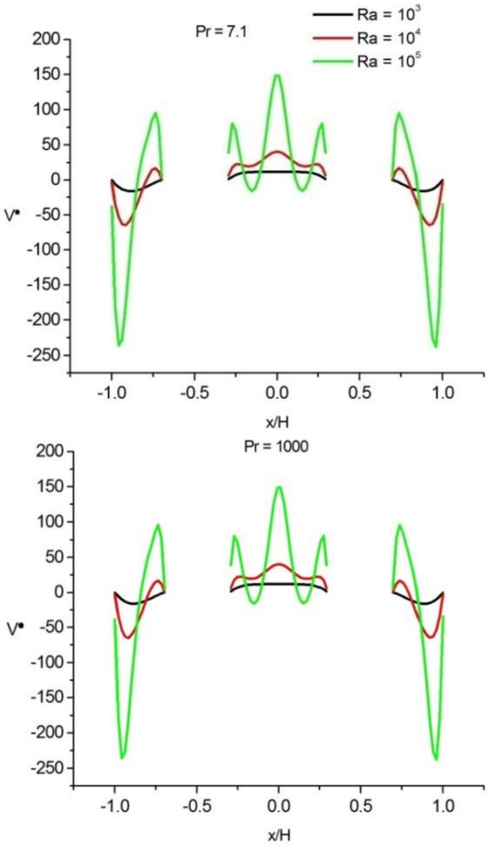


Fig. 6. Dimensionless velocity profiles along x-direction at y = 0 for different values of Ra and Pr

5. CONCLUSION

Numerical simulations are carried out to study the laminar natural convection in annular space. The domain consists of four circular cylinders, which are placed in across arrangement inside circular enclosure. The work tested the roles of Prandtl and Rayleigh numbers on the flow and thermal patterns. The obtained results of the present study let us draw some new points:

- The heat transfer rate of inner cylinders depends strongly on the position of cylinder.
- Increase in the value of Rayleigh number increases the heat transfer rate.
- The Prandtl number is a negligible parameter for the natural convection in annular space.
- As we move from the lower part to upper region of annular space increases the heat transfer rate.
- This arrangement of circular cylinders generate new extra loops above the lower cylinder.
- The upper part of inner cylinder has the lowest value of heat transfer rate, whereas, the bottom part of cylinder has the highest value of heat transfer.

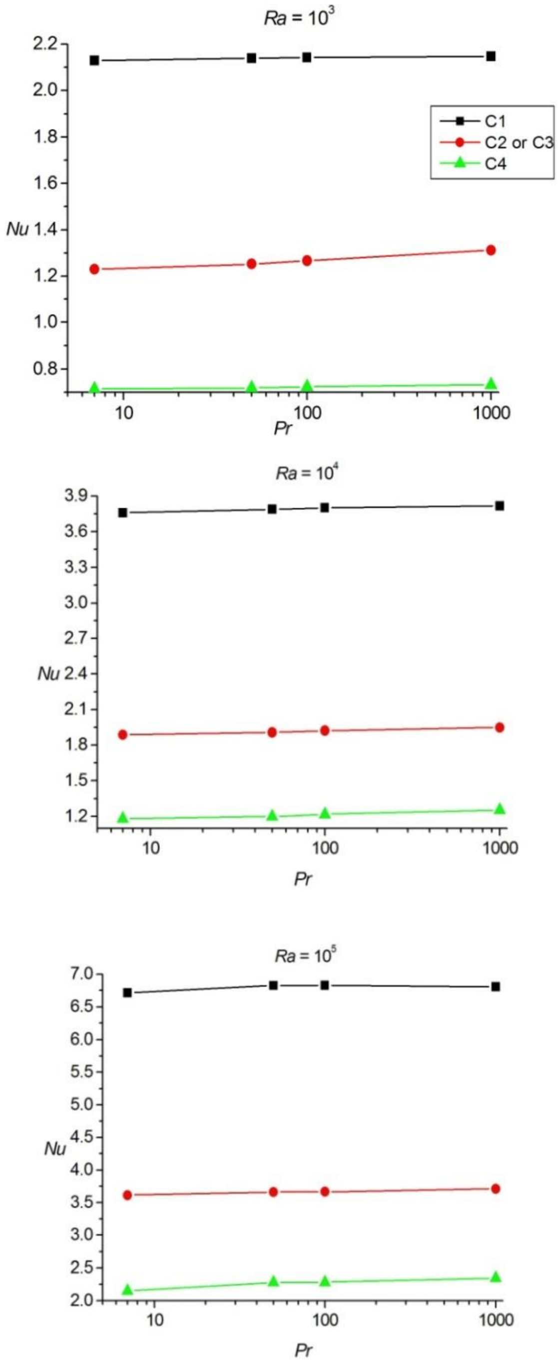


Fig. 7. Average Nusselt number versus Prandtl number for different cylinders and Rayleigh numbers

REFERENCES

1. **Abu-Nada E., Masoud Z., Hijazi A.**(2008), Natural convection heat transfer enhancement in horizontal concentric annuli using nanofluids, *International Communications in Heat and Mass Transfer*, 35, 657–665.
2. **Aly A.M.**(2017), Natural convection over circular cylinders in a porous enclosure filled with a nanofluid under thermo-diffusion effects, *Journal of the Taiwan Institute of Chemical Engineers*, 70, 88–103.
3. **Arbaban M., Salimpour M.R.**(2014), Enhancement of laminar natural convective heat transfer in concentric annuli with radial fins using nanofluids, *Heat Mass Transfer*, 47, 5181–5188.
4. **Char M., Lee G.C.**(1989), Maximum density effects on natural convection of micropolar fluids between horizontal eccentric cylinders, *Jnt. J. Engng Sci*., 36, 157–169.
5. **Eid E.I.** (2011) Experimental study of free convection in an elliptical annular enclosure in blunt and slender orientations, *Heat Mass Transfe*r, 47, 81–91.
6. **El-maghlany W., Abo-elazm, M., Shahata A., Eldrainy Y.**(2016), Mixed convection in an eccentric annulus filled by copper nanofluide, *Thermal Science*, 20, 1597–1608.
7. **HaM.Y., Kim J.G.** (2004), Numerical simulation of natural convection in annuli with internal fins, *KSME International Journal*, 18, 718–730.
8. **Hadidi H., Manshadi M.K.D., Kamali R.**(2020), Natural convection of power-law fluids inside an internally finned horizontal annulus. *Iranian Journal of Science and Technology, Transactions of Mechanical Engineering*, 44, 415–425.
9. **Ho C.J., Lin Y. H., Chen T.C.** (1989), A numerical study of natural convection in concentric and eccentric horizontal cylindrical annuli with mixed boundary conditions. *International Journal of Heat and Fluid Flow*, 10, 40–47.
10. **Hussein A.K.**(2013), Computational analysis of natural convection in a parallelogrammic cavity with a hot concentric circular cylinder moving at different vertical locations, *International Communications in Heat and Mass Transfer*, 46, 126–133.
11. **Iqbala Z., Syed K. S., Ishaq M.**(2017), Optimum configurations of annulus with triangular fins for laminar convection, *Thermal Science*, 21, 161–173.
12. **Kozlov N.** (2018),Steady flow in an annulus with a varying number of deflectors at rotational vibration, *Fluid Dynamic Research*, 50, 051402.
13. **Kuehn T.H., Goldstein R.J.**(1976), An experimental and theoretical study of natural convection in the annulus between horizontal concentric cylinders, *Journal of Fluid Mechanics*, 74, 695–719.
14. **Laidoudi H.** (2020), Buoynacy-driven flow in annular space from two circular cylinders in tandem arrangement, *Mutallurgical and Materials Engineering*, 26, 87–102.
15. **Laidoudi H., Helmaoui M., Bouzit M., Ghenaim A.** (2020) Natural convection of Newtonian fluids between two concentric cylinders of a special cross-sectional form, *Thermal Science*, 00, 00–00. 10.2298/TSCI200201190L.
16. **Masoumi H., Aghighi M.S., Ammar A.**(2019), Laminar natural convection of yield stress fluids in annular spaces between concentric cylinders, *International Journal of Heat and Mass Transfer*, 138, 1188–1198.
17. **Matin M.H., Khan W.A.**(2013), Laminar natural convection of non-Newtonian power-law fluids between concentric circular cylinders, *International Communications in Heat and Mass Transfer,* 43, 112–121.
18. **Nada S.A., Said M.A.**(2019), Effects of fins geometries, arrangements, dimensions and numbers on natural convection heat transfer characteristics in finned-horizontal annulus, *International Journal of Thermal Sciences*, 137, 121–137.
19. **Nasiri D., Dehghan A. A., Hadian M. R.**(2017), Conjugate natural convection between horizontal eccentric Cylinders, *Heat Mass Transfer*, 53, 799–811.
20. **Pandey S., Park Y. G., Ha M. Y.**(2019), An exhaustive review of studies on natural convection in enclosures with and without internal bodies of various shapes, *International Journal of Heat and Mass Transfer*, 138, 762–795.
21. **Sheikhzadeh G. A., Arbaban M., Mehrabian M. A.** (2013), Laminar natural convection of Cu-water nanofluid in concentric annuli with radial fins attached to the inner cylinder, *Heat Mass Transfer*, 49, 391–403
22. **Zhang H.L., Tao W.Q., Wu Q.J.**(1992), Numerical Simulation of natural convection in circular enclosures with inner polygonal cylinders, with confirmation by experimental results. *J. of Thermal Science*, 1, 249–258.

sciendo

DOI 10.2478/ama-2020-0015

EFFECT OF PRESSURE FLUCTUATIONS ON THE TEMPERATURE DURING BRAKING

Katarzyna TOPCZEWSKA, Przemysław ZAMOJSKI

Faculty of Mechanical Engineering, Bialystok University of Technology, 45C Wiejska Str., 15-351 Bialystok, Poland

k.topczewska@pb.edu.pl, zamojski.przemyslaw@gmail.com

received 13 May 2020, revised 29 June 2020, accepted 1 July 2020

Abstract: The aim of this study is to develop the numerical–analytical model of frictional heating in a pad/disc system during braking including the pressure fluctuations, engendered by the pump in an anti-skid braking operation. For this purpose, the problem of motion and the one-dimensional thermal problem of friction for a semi-space/semi-space tribosystem were formulated and solved. Obtained solutions allow to calculate temperature distribution on the contact surface and inside the friction elements. Thermal analysis was performed for a metal–ceramic pad and a cast iron disc during one-time braking including the time-dependent, oscillating pressure. The influence of amplitude of pressure fluctuations on the temperature variations was investigated, especially on the value of maximum temperature achieved during braking.

Key words: Braking, Pressure fluctuations, Pad/disc system, Temperature

1. INTRODUCTION

Braking operation is performed by proper clamp of friction pads to the rotating disc. Therefore, the nominal value and variations of pressure determine the effectiveness of the friction process that occurs. In the case of severe, heavy-loaded braking processes, the spatial distribution of pressure on the contact surface can be assumed as uniform (Yevtushenko and Kuciej, 2012; Matysiak and Yevtushenko, 2001). However, the time profile of contact pressure can have various shapes, which has significant influence on the process.

Most known solutions for thermal friction problems have been obtained for a classic case of braking with constant deceleration, i.e. assuming that pressure is invariable in time (Fazekas, 1953; Belhocine and Bouchetara, 2012; Talati and Jalalifar, 2009; Nosko et al., 2012). More precise models use an exponential function to describe the increase of contact pressure at the beginning of braking to nominal value (Topczewska, 2018; Yevtushenko et al., 1999; Matysiak et al., 2002) or its simplified form – linear growth of pressure – and maintain the achieved nominal value to the end of braking (Topczewska, 2018; Yevtushenko and Grześ, 2015; Yevtushenko et al., 2019). In the article Topczewska (2018), the evolution of temperature on friction surface was found in a closed analytical form for a classic semi-space/semi-space system with regard to the time profiles of specific friction power corresponding to the exponential and linear increase of pressure. Another possible pressure variation during braking was taken into account indirectly by applying different profiles of specific friction power during single braking (Yevtushenko et al., 2019). In the above-mentioned article, the analytical solutions of one-dimensional heat conduction problem formulated for semi-space/semi-space system were found.

All of the above-mentioned problems concerned processes with monotonically increasing pressure during braking. However, the growth of contact pressure may also be accompanied by its fluctuations caused by anti-slip regulation, anti-lock braking system or pulse braking mode. Thermal problems of friction contact pressure fluctuations during braking were formulated and solved for a strip/semi-space (Yevtushenko et al., 2010) and semi-space/strip/semi-space tribosystems (Kuciej, 2011). In the latter articles, mainly, the influence of the time of pressure increase to nominal value and heat transfer through a contact surface at the temperature level were investigated.

In this paper, the distribution of temperature in the pad/disc system was found during braking with an exponential growth of pressure and its fluctuations. For this purpose, the initial problem of motion and the one-dimensional boundary value problem of heat conduction were considered. Performed numerical analysis allowed to establish the effect of amplitude of pressure oscillation on the contact surface during braking on the temperature distribution.

2. STATEMENT OF THE PROBLEM

One-time, rapid braking process of a pad/brake disc system was considered. The operating principle of this type of brake system is based on the friction pads' clamp on the surface of the rotating disc. As a result of friction, heat is generated on the contact surfaces. Investigations of temperature distribution in a heated friction element during braking at high sliding speed showed the following (Yevtushenko and Kuciej, 2012; Balakin and Sergienko, 1999):

1. Almost all frictional heat is absorbed by the pads and disc in an axial direction, i.e. perpendicular to the friction surface.
2. Significant gradients of temperature are present only at a short distance from the contact surface, the so-called effective depth of heat penetration (Chichinadze et al., 2010), which is much smaller than the actual thickness of friction elements.
3. The effect of convective heat exchange with the environment and wear are negligible (Yevtushenko et al., 2020; Topczewska et al., 2020). So, the whole energy generated during braking is converted into heat.

4. Thermal resistance on the interface of pad and disc is inconsiderable. So, thermal contact is perfect.

In order to determine the temperature distribution in friction elements, proper initial problem of motion and thermal problem of friction were formulated on the basis of the above assumptions (1–4).

3. THE INITIAL PROBLEM OF MOTION

The initial problem of motion for a pad/disc system during single braking from initial sliding velocity of the disc V_0 to the standstill position at t_s has the following form (Kuciej, 2012):

$$2W_0V_0^{-2}\frac{d}{dt}V(t)=-fp(t)A_k,\ 0\le t\le t_s\ , \tag{1}$$

$$V(0)=V_0, \tag{2}$$

where W_0 is the initial kinetic energy, f the coefficient of friction, $A_k=2\pi(r_e^2-r_i^2)$ the nominal contact surface, r_e, r_i the external and internal radii of a single friction surface and $p(t)$ is the contact pressure. Problem of motion (1), (2) can be written in the dimensionless form:

$$\tau_s^0\frac{d}{d\tau}V^*(\tau)=-p^*(\tau),0\le\tau\le\tau_s, \tag{3}$$

$$V^*(0)=1, \tag{4}$$

where the introduced dimensionless variables and parameters are:

$$\tau=\frac{k_1t}{d^2},\ \tau_s=\frac{k_1t_s}{d^2},\ \tau_s^0=\frac{k_1t_s^0}{d^2},\ t_s^0=\frac{2W_0}{fV_0p_0A_k},$$
$$V^*=\frac{V}{V_0},\ p^*=\frac{p}{p_0},\ d=\sqrt{3k_1t_s^0}, \tag{5}$$

where p_0 is the nominal value of pressure, k_1 the thermal diffusivity of the disc material and d is the effective depth of heat penetration. Based on equation (3), taking into account of the initial condition (4), the dimensionless sliding velocity can be found using the equation:

$$V^*(\tau)=1-\frac{1}{\tau_s^0}\int_0^\tau p^*(s)ds\ ,0\le\tau\le\tau_s. \tag{6}$$

Taking into account the growth of contact pressure at the beginning of braking, and also fluctuations of its value due to anti-lock braking system, the dimensionless time-dependent pressure function $p^*(\tau)$ is (Yevtushenko et al., 2010; Kuciej, 2011) calculated as:

$$p^*(\tau)=\left[1-\exp\left(\frac{-\tau}{\tau_m}\right)\right][1+a\sin(\omega\tau)],0\le\tau\le\tau_s, \tag{7}$$

where t_m is the time of pressure increase, $\tau_m=k_1t_md^{-2}$ the dimensionless time of pressure increase, a the dimensionless amplitude of pressure oscillations and ω is the dimensionless frequency of pressure oscillations. With the analytical solution of equation (6) including pressure (7), the dimensionless sliding speed was found:

$$V^*(\tau)=1-\frac{\tau}{\tau_s^0}+\frac{\tau_m}{\tau_s^0}\left[1-\exp\left(\frac{-\tau}{\tau_m}\right)\right]-\frac{a}{\tau_s^0}V_a^*(\tau),\ 0\le\tau\le\tau_s,$$
$$V_a^*(\tau)=\frac{1}{\omega}[1-\cos(\omega\tau)]+\frac{1}{\omega^2+\tau_m^{-2}}\Big\{\exp\left(\frac{-\tau}{\tau_m}\right)[\omega\cos(\omega\tau)+$$
$$+\tau_m^{-1}\sin(\omega\tau)]\Big\}-\omega. \tag{8}$$

The stop condition $V^*(\tau_s)=0$ allows to calculate numerically the time of braking.

For constant friction coefficient f, the specific friction power is the product of sliding velocity (8) and pressure (7):

$$q(t)=q_0q^*(\tau),\ q_0=fp_0V_0,$$
$$q^*(\tau)=\left[1-\exp\left(\frac{-\tau}{\tau_m}\right)\right][1+a\sin(\omega\tau)]\,\langle1-\frac{\tau}{\tau_s^0}+\frac{\tau_m}{\tau_s^0}\Big[1-$$
$$-\exp\left(\frac{-\tau}{\tau_m}\right)\Big]-\frac{a}{\tau_s^0}V_a^*(\tau)\rangle,\ 0\le\tau\le\tau_s. \tag{9}$$

Substituting $a=0$, the obtained relations (7)–(9) take the forms:

$$p^*(\tau)=1-\exp\left(\frac{-\tau}{\tau_m}\right),\ V^*(\tau)=1-\frac{\tau}{\tau_s^0}+\frac{\tau_m}{\tau_s^0}\left[1-\exp\left(\frac{-\tau}{\tau_m}\right)\right],$$
$$q^*(\tau)=\left[1-\exp\left(\frac{-\tau}{\tau_m}\right)\right]\langle1-\frac{\tau}{\tau_s^0}+\frac{\tau_m}{\tau_s^0}\left[1-\exp\left(\frac{-\tau}{\tau_m}\right)\right]\rangle,$$
$$0\le\tau\le\tau_s, \tag{10}$$

which correspond to braking with monotonically increasing pressure without fluctuations.

As the time of pressure growth approaches zero, $\tau_m\to0$, from equations (7) to (10), the following equations were obtained:

$$p^*(\tau)=1+a\sin(\omega\tau),$$
$$V^*(\tau)=1-\frac{\tau}{\tau_s^0}-\frac{a}{\tau_s^0}\Big\{\frac{1}{\omega}[1-\cos(\omega\tau)]-\omega\Big\},$$
$$q^*(\tau)=[1+a\sin(\omega\tau)]\,\langle1-\frac{\tau}{\tau_s^0}-\frac{a}{\tau_s^0}\Big\{\frac{1}{\omega}[1-\cos(\omega\tau)]-\omega\Big\}\rangle,$$
$$0\le\tau\le\tau_s. \tag{11}$$

And the simpliest typical braking process with uniform pressure can be found for $a=0$ and $\tau_m\to0$:

$$p^*(\tau)=1,\ V^*(\tau)=1-\frac{\tau}{\tau_s^0},\ q^*(\tau)=1-\frac{\tau}{\tau_s^0},$$
$$0\le\tau\le\tau_s=\tau_s^0. \tag{12}$$

4. THE HEAT PROBLEM OF FRICTION

To establish the temperature field $T_i(z,t)$ in the brake disc $(i=1)$ and pad $(i=2)$, the thermal problem of friction was considered. Due to the symmetry of the system, we assume that the friction processes occurring on both friction surfaces of the brake disc are the same; so, the model was prepared for half of this system. Based on assumptions (1–4), half of the disc and the brake pad was simplified by a system of two semi-limited bodies. The orientation of semi-space/semi-space system considered here is given relative to a Cartesian frame of reference $Oxyz$ (Fig. 1).

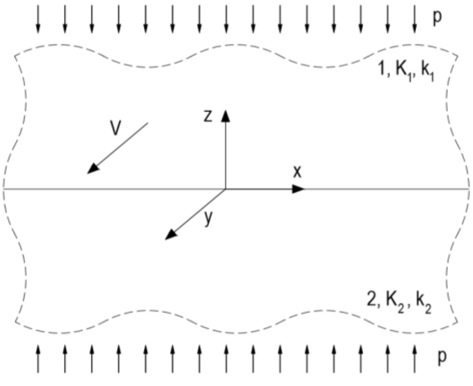


Fig. 1. Scheme of the problem

On the contact surface, $z = 0$; the frictional heat is generated and absorbed by elements along the z-axis direction. It was assumed that the friction materials are homogeneous and their thermal properties do not change during braking.

The boundary value problem of heat conduction was formulated for a semi-space/semi-space system in the dimensionless form:

$$\frac{\partial^2}{\partial \zeta^2} T_1^*(\zeta,\tau) = \frac{\partial}{\partial \tau} T_1^*(\zeta,\tau), \zeta > 0, 0 < \tau \le \tau_s, \tag{13}$$

$$\frac{\partial^2}{\partial \zeta^2} T_2^*(\zeta,\tau) = \frac{1}{k^*}\frac{\partial}{\partial \tau} T_2^*(\zeta,\tau), \zeta < 0, 0 < \tau \le \tau_s, \tag{14}$$

$$K^* \frac{\partial}{\partial \zeta} T_2^*(\zeta,\tau)\Big|_{\zeta=0} - \frac{\partial}{\partial \zeta} T_1^*(\zeta,\tau)\Big|_{\zeta=0} = q^*(\tau), 0 < \tau \le \tau_s, \tag{15}$$

$$T_1^*(0,\tau) = T_2^*(0,\tau) = T^*(\tau), 0 < \tau \le \tau_s, \tag{16}$$

$$T_i^*(\zeta,\tau) \to 0, |\zeta| \to \infty, 0 < \tau \le \tau_s, i = 1{,}2, \tag{17}$$

$$T_i^*(\zeta,0) = 0, |\zeta| \ge 0, i = 1{,}2, \tag{18}$$

$$\zeta = \frac{z}{d}, K^* = \frac{K_2}{K_1}, k^* = \frac{k_2}{k_1}, T_0 = \frac{q_0 d}{K_1}, \tau = \frac{k_1 t}{d^2}, \tau_s = \frac{k_1 t_s}{d^2},$$

$$T_i^* = \frac{T_i - T_a}{T_0},\ d = \sqrt{3k_1 t_s^0},\ i = 1{,}2\ , \tag{19}$$

where T_a is the initial temperature.

Solution to the problem (13)–(19) was found based on the following Duhamel's theorem (Ozisik,1993).

$$T_i^*(\zeta,\tau) = \int_0^\tau q^*(s)\frac{\partial}{\partial \tau} T_i^{*(0)}(\zeta,\tau - s)ds,$$

$$|\zeta| \ge 0, 0 \le \tau \le \tau_s, i = 1{,}2, \tag{20}$$

where (Carlslaw and Jaeger, 1959)

$$T_i^{*(0)}(\zeta,\tau) = 2\gamma\sqrt{\tau}\,\mathrm{ierfc}[Z_i(\zeta,\tau)],\ \zeta \ge 0, 0 \le \tau \le \tau_s,$$

$$Z_1(\zeta,\tau) = \frac{\zeta}{2\sqrt{\tau}}, \zeta \ge 0, Z_2(\zeta,\tau) = \frac{\zeta}{2\sqrt{\tau k^*}}, \zeta \le 0,$$

$$\gamma = (1+\varepsilon)^{-1}, \varepsilon = K^*(k^*)^{-0.5} \tag{21}$$

is the solution to the problem (13)–(19) with a constant intensity of frictional heat flux $q^*(\tau) = 1$ in the boundary condition (15).
Taking into account the partial derivative of function (21) (Abramowitz and Stegun, 1972)

$$\frac{\partial}{\partial \tau} T_i^{*(0)}(\zeta,\tau) = \gamma[\pi(\tau - s)]^{-0.5}\exp\left[-Z_i{}^2(\zeta,\tau - s)\right] \tag{22}$$

and the obtained specific friction power $q^*(\tau)$ (9), we determined the temperature fields for the braking process considered here:

$$\begin{aligned} T_i^*(\zeta,\tau) = {} & \gamma\pi^{-0.5}\int_0^\tau \left[1 - \exp\left(\frac{-s}{\tau_m}\right)\right][1 + a\sin(\omega\tau)]\,\langle 1 - \frac{s}{\tau_s^0} + \\ & + \frac{\tau_m}{\tau_s^0}\left[1 - \exp\left(\frac{-s}{\tau_m}\right)\right] - \\ & - \frac{a}{\tau_s^0}\left\{\frac{1}{\omega}[1 - \cos(\omega s)] + \frac{1}{\omega^2 + \tau_m^{-2}}\left\{\exp\left(\frac{-s}{\tau_m}\right)[\omega\cos(\omega s) + \right.\right. \\ & \left.\left. + \tau_m^{-1}\sin(\omega s)]\right\} - \omega\right\}\rangle\,(\tau - s)^{-0.5}\exp\left[-Z_i{}^2(\zeta,\tau - s)\right]ds, \end{aligned}$$

$$0 \le \tau \le \tau_s, -\infty < \zeta < \infty,\ i = 1{,}2. \tag{23}$$

In the case of braking, when the time of pressure increase is close to zero, $\tau_m \to 0$ (11), the solution to the thermal friction problem (13)–(19) has the following form:

$$\begin{aligned} T_i^*(\zeta,\tau) = {} & \frac{\gamma}{\sqrt{\pi}}\int_0^\tau [1 + a\sin(\omega\tau)]\,\langle 1 - \frac{\tau}{\tau_s^0} - \frac{a}{\tau_s^0}\left\{\frac{1}{\omega}[1 - \cos(\omega s)] - \right. \\ & \left. - \omega\right\}\rangle\,(\tau - s)^{-0.5}\exp\left[-Z_i{}^2(\zeta,\tau - s)\right]ds, \end{aligned}$$

$$0 \le \tau \le \tau_s, -\infty < \zeta < \infty,\ i = 1{,}2. \tag{24}$$

Analytical integration of the formulas (23) and (24) is impossible. Therefore, it was done by means of the adaptive quadrature integrator, which handles singularities – procedure QAGS from a package of numerical integration QUADPACK (Piessens et al., 1983).

The analytical solution to the considered problem (13)–(19) without taking into account the pressure oscillations for $a = 0$ has been succesfully found in the paper Topczewska (2018). Using the same method (20)–(22) for functions (10), the following formula has been obtained:

$$\begin{aligned} T^*(\tau) = \gamma\sqrt{\tau}\Bigg[& \left(1 + \frac{\tau_i}{2\tau_s^0} - \frac{2}{3}\frac{\tau}{\tau_s^0}\right)\frac{2}{\sqrt{\pi}} - \left(1 - \frac{\tau}{\tau_s^0} + \frac{3}{2}\frac{\tau_i}{\tau_s^0}\right)D\left(\sqrt{\frac{\tau}{\tau_i}}\right) + \\ & + \frac{\tau_i}{\tau_s^0}D\left(\sqrt{\frac{2\tau}{\tau_i}}\right)\Bigg], 0 < \tau \le \tau_s, \end{aligned} \tag{25}$$

where

$$D(x) = \frac{2}{\sqrt{\pi}}\frac{\exp(-x^2)}{x}\int_0^x \exp(s^2)\,ds. \tag{26}$$

However, this solution, (25) and (26), allows to establish temperature evolution only of the friction surface $\zeta = 0$, where the temperature level is the highest.

Also, the full analytical solution of this problem (13)–(19) has been determined for the uniform braking mode in the paper Yevtushenko et al. (2019):

$$\begin{aligned} T_i^*(\zeta,\tau) = {} & \frac{4}{3}\gamma\sqrt{\tau}\frac{\tau}{\tau_s}\langle\left\{3\frac{\tau_s}{\tau} - 2\left[1 + Z_i{}^2(\zeta,\tau)\right]\right\}\mathrm{ierfc}[Z_i(\zeta,\tau)] + \\ & + Z_i(\zeta,\tau)\mathrm{erfc}[Z_i(\zeta,\tau)]\rangle, \end{aligned}$$

$$|\zeta| \ge 0, 0 \le \tau \le \tau_s, i = 1{,}2, \tag{27}$$

which allows to calculate temperature both on the contact surface and at any depth $|\zeta| \ge 0$ inside the friction elements.

5. NUMERICAL ANALYSIS

Distribution of temperature during braking was designated for the friction pair consisting of the pad made of metal–ceramic composite FMC-11 and the brake disc made of cast iron ChNMKh. Properties of these materials and input parameters to perform analysis are presented in Table 1.

Table 1. Input parameters (Yevtushenko et al., 2020; Kuciej, 2012)

Thermal conductivity of disc, K_1 $(\mathrm{W/mK})$	51
Thermal conductivity of pad, K_2 $(\mathrm{W/mK})$	34.3
Thermal diffusivity of disc, k_1 $(\mathrm{mm^2/s})$	14
Thermal diffusivity of pad, k_2 $(\mathrm{mm^2/s})$	15.2
Initial velocity, V_0 $(\mathrm{m/s})$	23.8
Initial kinetic energy, W_0 (k)	103.54
Nominal contact pressure, p_0 (MPa)	0.607
Friction coefficient, f $(-)$	0.27
Time of pressure increase, t_m (s)	0.5
Ambient temperature, T_a $(°\mathrm{C})$	20
External radius, r_e (mm)	37.5
Internal radius, r_i (mm)	26.5
Pressure fluctuation amplitude, a $(-)$	0.1
Pressure fluctuation frequency, ω $(-)$	100

Based on the values included in Table 1, the following parameters were computed: the time of braking $t_s = 12.43$ s, the whole area of friction pair contact $A_k = 4.42 \times 10^{-3}$ m^2, effective depth of heat penetration in the brake disc $d = 22.45$ mm and the coefficient of heat flux partition $\gamma = 0.607$.

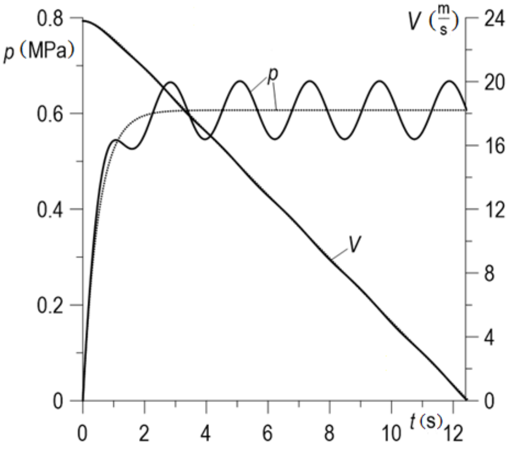


Fig. 2. Changes of contact pressure p and sliding speed V during braking on account of pressure fluctuations for $a = 0.1$ (solid lines), (7), (8) and without for $a = 0$ (dashed lines) (6), (10)

Changes in contact pressure p and sliding speed V during braking are shown in Fig. 2. Solid lines present the results obtained for oscillations with amplitude $a = 0.1$ (7), (8), and dashed lines show the results obtained excluding the pressure fluctuations for $a = 0$ (10). In the first case (for $a = 0.1$), at the beginning of the process, the pressure rapidly increases. After that, it oscillates around the nominal value p_0 with a constant amplitude to the end of braking $t = t_s$. However, the sliding velocity monotonically decreases in time until the stop moment. The time profiles of speed in both cases are almost convergent (Fig. 2).

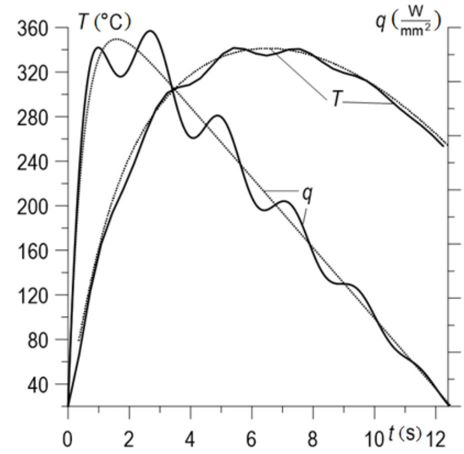


Fig. 3. Changes of specific friction power q (9) and temperature T on the contact surface during braking: including the pressure fluctuations for $a = 0.1$ (solid lines) (23) and excluding for $a = 0$ (dashed lines) (24)

Pressure oscillations with $a = 0.1$ noticeably affect specific friction power q (9) and temperature T (23), (24) on the contact surface $z = 0$, which is demonstrated in Fig. 3 by solid lines. The amplitude of their fluctuations is the greatest near the time moment of reaching their maximum values $q_{\max} = 3.46$ W/mm^2 and $T_{\max} = 339.5$°C at time moments $t = 2.66$ s and $t = 6.64$ s, respectively, and after that it decreases until the end of braking. Temperature oscillations on the friction surface are much less intense than the fluctuations of specific friction power.

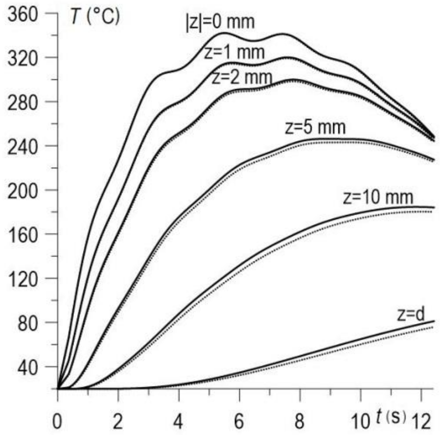


Fig. 4. Variations of temperature T (23) on the contact surfaces $z = 0$ and on selected distance from z in the disc $(i = 1)$ (solid lines) and pad $(i = 2)$ (dotted lines)

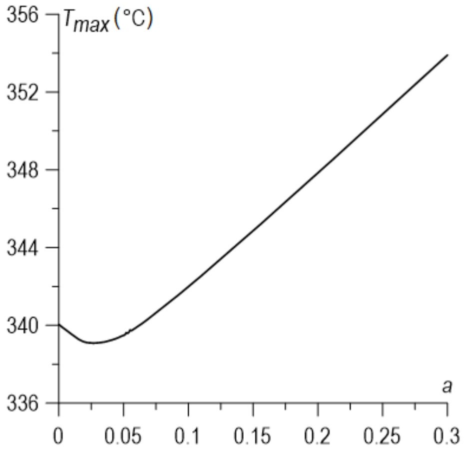


Fig. 5. Dependence of maximum temperature $T_{\max}$ attained during braking on the value of dimensionless amplitude a

Changes in temperature T on the interface and few depths z inside the brake disc $(i = 1)$ (solid lines) and pad $(i = 2)$ (dotted lines) were calculated based on formula (23) and are presented in Fig. 4. The temperature values on the contact surfaces $|z| = 0$ are the highest and the same for both friction pair elements (16). Due to similarity of thermal properties of friction materials (Table 1), the values of achieved temperature at selected distances in both elements are close to each other. The greater the distance from the friction surface z, the more noticeable are the differences between the temperature in the disc and the pad. As the distance from the friction surface increases, the temperature level decreases and its maximum values are reached later. The influence of pressure oscillations on temperature is especially visible on the friction surfaces $z = 0$, and moving away from this surface, the fluctuation amplitude decreases. Below the depth $z = 5$ mm, this effect is not noticeable both in the disc and the brake pad.

Figure 5 presents the maximum temperature values achieved on the contact surface during braking for different values of fluctuations amplitude a. When braking without contact pressure oscillation $a = 0$, the maximum temperature value is $T_{\max} = 340°\mathrm{C}$. The lowest value of maximum temperature $T_{\max} = 339°\mathrm{C}$ is attained for small pressure fluctuations with an amplitude $a = 0.03$. However, further increasing the amplitude value causes increase in maximum temperature to $T_{\max} = 353.9°\mathrm{C}$ at the most intense oscillations for an amplitude $a = 0.3$.

6. CONCLUSIONS

Numerical analysis of temperature distribution in a pad/disc brake system was performed based on the solution to the boundary value problem of heat conduction formulated with oscillating pressure. Calculations were performed for the frictional pair such as composite pad and cast iron disc. Results were compared with data obtained excluding the pressure fluctuations. The following conclusions were made:

1. Pressure fluctuations do not influence the sliding speed of the brake disc, and hence the time of braking.
2. The temperature in the disc and pad oscillates only in the near distance from friction surface and with a much lower intensity than the corresponding fluctuations in pressure. The amplitude of temperature oscillations changes over time during braking and achieves the highest value in the middle of process, despite the fact that pressure fluctuates with a constant amplitude.
3. Increase in the level of amplitude of pressure fluctuations causes a slight increase in the value of maximum temperature attained on the friction surface.

REFERENCES

1. **Abramowitz M., Stegun I.A.** (1972), *Handbook of Mathematical Functions with Formulas, Graphs and Tables*, National Bureau of Standards, Washington.
2. **Balakin V. A., Sergienko V. P.** (1999), *Thermal calculations of brakes and units of friction*, IMMS NANB, Gomel.
3. **Belhocine A., Bouchetara M.** (2012), Thermal analysis of a solid brake disc. *Appl. Therm. Eng.*, Vol. 32, 59–67.
4. **Carlslaw H.S., Jaeger J. C.** (1959), *Conduction of Heat in Solids*, 2nd ed.Clarendon Press, Oxford.
5. **Chichinadze A.V., Kozhemyakina V.D., Suvorov A.V.** (2010), Method of temperature-field calculation in model ring specimens during bilateral friction in multidisc aircraft brakes with the IM-58-T2 new multipurpose friction machine, *J. Friction. Wear*, Vol. 31, No. 1 23-32.
6. **Fazekas G.A.G.** (1953), Temperature gradients and heat stresses in brake drums, *SAE Trans*. Vol. 61, 279–284.
7. **Kuciej M.** (2011), Accounting changes of pressure in time in one-dimensional modeling the process of friction heating of disc brake. *Int. J. Heat Mass Transfer*, 54 (1–3), 468–474.
8. **Kuciej M.** (2012), *Analytical Models of Transient Friction Heating*, Oficyna Wydawnicza Politechniki Białostockiej, Białystok (in Polish).
9. **Matysiak S. J., Yevtushenko A. A., Ivanyk E. G.** (2002), Contact temperature and wear of composite friction elements during braking. *Int. J. Heat Mass Transfer*, Vol. 45, No. 1, 193-199.
10. **Matysiak S.J., Yevtushenko A.A.** (2001), On heating problems of friction, *J. Theor. Appl. Mech*., Vol. 39, No. 3, 577–588.
11. **Nosko A.L., Mozalev V.V., Nosko A.P., Suvorov A.V., Lebedeva V.N.** (2012), Calculation of temperature of carbon disks of aircraft brakes with account of heat exchange with the environment, J. Friction. Wear, Vol. 33, No. 4, 233–238.
12. **Ozisik M.N**. (1993), *Heat conduction*, 2nd Ed. Wiley: New York.
13. **Piessens R., de Doncker-Kapenga E., Überhuber C.W., Kahaner D.K.** (1983), *QUADPACK: A Subroutine Package for Automatic Integration,* Springer-Verlag, Berlin.
14. **Talati F., Jalalifar S.** (2009), Analysis of heat conduction in a disk brake system, *Heat Mass Transfer*, Vol. 45, 1047–1059.
15. **Topczewska K.** (2018), Influence of the Time of Increase in Contact Pressure in the Course of Braking on the Temperature of a Pad–Disc Tribosystem, *Materials Science*, Vol. 54, No. 2, 250-259.
16. **Topczewska K., Schlattmann J., Abdullah O.I.** (2020), Temperature and thermal stresses distributions in a dry friction clutch. *J. Theor. Appl. Mech*., Vol. 58, No. 2, 351–360.
17. **Yevtushenko A., Grześ P.** (2015), Maximum temperature in a three-disc thermally nonlinear braking system, *Int. J. Heat Mass Transfer*, Vol. 68, 291–298.
18. **Yevtushenko A., Kuciej M.** (2012), One–dimensional thermal problem of friction during braking: The history of development and actual state, *Int. J. Heat Mass Transfer,* Vol. 55, 4148–4153.
19. **Yevtushenko A., Kuciej M., Topczewska K.** (2019), Effect of the temporal profile of the friction power on temperature of a pad-disc brake system, *J. Theor. Appl. Mech*., Vol. 57, No. 2, 461-473.
20. **Yevtushenko A., Kuciej M., Topczewska K.** (2020), Some theoretical model for determining the temperature field of a multi-disk brake. *Adv. Mech. Eng.*, Vol. 12, No. 1, 1–15.
21. **Yevtushenko A.A., Ivanyk E.G., Yevtushenko O.O.** (1999), Exact formulae for determination of the mean temperature and wear during braking. *Heat Mass Transfer*, Vol. 35, No. 2, 163-169.
22. **Yevtushenko A.A., Kuciej M., Yevtushenko O.O.** (2010), Influence of the pressure fluctuations on the temperature in pad/disc tribosystem. *Int. J. Heat Mass Transfer*, Vol. 37, No. 8, 978–983.

DOI 10.2478/ama-2020-0016

FRACTIONAL VECTOR-ORDER h-REALISATION OF THE IMPULSE RESPONSE FUNCTION

Ewa PAWŁUSZEWICZ

Faculty of Mechanical Engineering, Department of Mechatronics Systems and Robotics,
Bialystok University of Technology, Wiejska 45c, 15-351 Białystok, Poland

e.pawluszewicz@pb.edu.pl

received 12 March 2020, revised 30 June 2020, accepted 3 July 2020

Abstract: The problem of realisation of linear control systems with the $\boldsymbol{h}-$difference of Caputo-, Riemann–Liouville- and Grünwald–Letnikov-type fractional vector-order operators is studied. The problem of existing minimal realisation is discussed.

Key words: Realisation, Linear system, Fractional vector-order system, $\boldsymbol{h}-$Markov parameters

1. INTRODUCTION

In engineering experiments, in many cases, only information about inputs and measurements of the investigated process are available. So, a relationship between these variables is needed. It is a question of the possible systems that provide a good description of observed system's input–output behaviour. This leads to the crucial idea of the realisation problem. In fact, the realisation of an input–output map describing a system's behaviour means finding a dynamical state-space system with input and output, which can be reproduced, when initialised at some state for the given (input–output) behaviour. In Bartosiewicz and Pawluszewicz (2006), classical conditions for existing realisation for continuous time or/and discrete time linear systems were generalised to any time domain. The next natural question is whether this problem can be extended to a more general case of differential/difference order, i.e. on systems defined by fractional order operators. The term fractional basically implies all non-integer numbers. In fact, in nature, there are many processes that can be more accurately modelled using fractional differ-integrals (see, e.g. in Ambroziak et al., 2016; Das, 2008; Koszewnik et al., 2016; Sierociuk et al., 2013; Wu et al., 2015). The rapid development of computer techniques has caused the parallel investigations in the field, among others, combinatorics tools and difference equations. This is the reason that in modelling of real phenomena, a generalisation of nth order differences to their fractional forms and the state-space description of control systems in discrete time are used (see, e.g. Bastos et al., 2011; Oprzedkiewicz and Gawin, 2016; Podlubny 1999).

The goal of this study is to construct a state-space fractional vector-order representation of an abstract input–output map and to give conditions under which such representation exists. To achieve this aim, fractional order $\mathrm{h}-$differences of Caputo-, Riemann–Liouville- and Grünwald–Letnikov-type operators are considered. Taking into account their properties (Mozyrska et al., 2013), the state-space description of the system's behaviours is presented in terms of these operators parallel. The main result may be seen as an extension of the classical realisability criterion saying that an abstract input–output map has a state-space realisation if and only if the Markov parameters satisfy a recurrence relation (see Sontag, 1998; Zabczyk, 2008). To achieve this aim, $\mathrm{h}-$Markov parameters for the input–output map are defined. It is shown that the input–output map has a state-space fractional vector-order $\mathrm{h}-$realisation in finite number of steps if and only if the $\mathrm{h}-$Markov parameters satisfy the linear recursion equation. The obtained relation is similar to the one given in the classical case; it extends the classical result to the fractional case. Generally, the obtained realisation is not unique; but under certain minimality or redundancy requirements, it can be what is a desirable property in practice. In fractional order case, some aspects of realisation problem were raised in Bettayeb et al. (2008), where the concept of the structured realisation index was introduced.

The paper is organised as follows. After introducing Mittag-Leffler function (Section 2.1) and fractional order difference operators (Section 2.2), the extension of controllability and observability conditions for fractional vector-order systems is presented (Section 3). In the next step, conditions of existing state-space fractional vector-order realisation are considered. As the last step, the problem of existing minimal fractional vector-order realisation is discussed (Section 4).

2. PRELIMINARIES

2.1. Discrete Mittag-Leffler function

Let α be any number and s any integer. Then:

$$\binom{\alpha}{s} = \begin{cases} 0 & for\ s<0 \\ 1 & for\ s=0 \\ \dfrac{\alpha(\alpha-1)\ldots(\alpha-s+1)}{s!} & for\ s>0 \end{cases}$$

denotes the classical binomial coefficient. Denote the family of binomial functions by φ_μ parametrised by $\mu>0$ as:

$$\varphi_\mu(n) = \begin{cases} \binom{n+\mu-1}{n} & for\, n \in N_0 \\ 0 & for\, n < 0. \end{cases} \tag{1}$$

If "$*$" denotes a convolution operator, then $\left(\varphi_\mu * \bar{x}\right)(n) := \sum_{s=0}^{n} \binom{n-s+\mu-1}{n-s} \bar{x}(s)$, where $\bar{x}(s) := x(a+sh)$.

The discrete two-parameter Mittag-Leffler function is defined as (Mozyrska and Wyrwas, 2015; Mozyrska et al., 2017):

$$E_{(\alpha,\beta)}(\lambda,n) := \sum_{k=0}^{\infty} \lambda^k \varphi_{k\alpha+\beta}(n-k). \tag{2}$$

If A is $\mathrm{n} \times \mathrm{n}$ dimensional matrix with constant coefficients, then $E_{(\alpha,\beta)}(A,n) := \sum_{k=0}^{\infty} A^k \binom{n-k+k\alpha+\beta-1}{n-k}$ and $\varphi_{k\alpha+\beta}(n-k) = 0$ for $n < k$. If $\alpha = \beta$, then $E_{(\alpha,\alpha)}(A,n) = \sum_{k=0}^{\infty} A^k \binom{n-(k+1)(\alpha-1)}{n-k}$. It is easy to check that even if there are matrices B and C such that $A = BC$, then $E_{(\alpha,\beta)}(BC,n) \neq E_{(\alpha,\beta)}(B,n) E_{(\alpha,\beta)}(C,n)$.

Let $A \in R^{n\times n}$ be a diagonalisable matrix with eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_{m_r}$ of multiples $m_1, m_2, \dots, m_r$, respectively, such that $\sum_{\mathrm{i}=1}^{\mathrm{r}} \mathrm{m_i} \leq \mathrm{n}$. Suppose that the function $f(\lambda)$ is well defined on the spectrum of the matrix A. Then a function of the matrix A is given by the Lagrange–Sylvester's interpolation formula (Gantmacher, 1959; Kaczorek, 1998) as:

$$f(A) = \sum_{i=1}^{r} [Z_{i1} \tfrac{df(\lambda)}{d\lambda} |_{\lambda=\lambda_i} + \dots + Z_{i1} \tfrac{df^{(m_i-1)}(\lambda)}{d\lambda^{(m_i-1)}} |_{\lambda=\lambda_i}] \tag{3}$$

with coefficients:

$$Z_{ij} = \sum_{k=j-1}^{m_i-1} \tfrac{(A-\lambda_i I_n)^k \Phi(A)}{(k-j+1)!(j-1)!} \tfrac{df^{k-j+1}}{d\lambda^{k-j+1}} \left[\tfrac{1}{\Phi(\lambda)}\right] |_{\lambda=\lambda_i} \tag{4}$$

for $j = 1, \dots, m_r$, $\Phi(A) = \prod_{j=1}^{m_r} \left(A - \lambda_j I_n\right)^{m_i}$, $\Phi(\lambda) = \prod_{j=1}^{m_r} \left(\lambda - \lambda_j\right)^{m_i}$ and I_n – identity matrix of dimension $n \times n$.

Theorem 1 (Kaczorek, 2017): Let $\Phi(\lambda) = \det[\lambda I_n - f(A)] = \lambda^n + a_{n-a}\lambda^{n-a} + \dots + a_1\lambda + a_0$, where $f(A)$ is given by (3), be the characteristic polynomial of matrix A. Then $f(A)$ satisfies its characteristic equation $[f(A)]^n + a_{n-1}[f(A)]^{n-1} + \dots + a_1 f(A) + a_0 I_n = 0$.

Proposition 2: Let $\Phi(\lambda) = \det\left[\lambda I_n - E_{(\alpha,\beta)}(A,k)\right] = \lambda^n + a_{n-a}\lambda^{n-a} + \dots + a_1\lambda + a_0$ be the characteristic equation of the Mittag-Leffler function (2). Then matrix $\Phi(k) = E_{(\alpha,\beta)}(A,k)$ satisfies its characteristic equation $\left[E_{(\alpha,\beta)}(A,k)\right]^n + a_{n-1}\left[f E_{(\alpha,\beta)}(A,k)\right]^{n-1} + \dots + a_0 I_n = 0$.

Proof: The reasoning using the Lagrange–Sylvester's formula (3), based on Kaczorek (2017), is the same as for one-parameter function Mittag-Leffler given in Pawluszewicz and Koszewnik (2019). □

From Proposition 2, immediately it follows that $E_{(\alpha,\beta)}(A,n) := \sum_{k=0}^{n} A^k \binom{n-k+k\alpha+\beta-1}{n-k}$.

2.2. Fractional h −difference operators

Let h be a positive real number. For any real a, let $(hN)_a = \{a, a+h, a+2h, \dots\}$. Consider a function $x: (hN)_a \to R$. The forward h-difference operator is classically defined as $(\Delta_h x)(t) = \frac{x(t+h)-x(t)}{h}$. The n-fold application n of operator Δ_h, i.e. $\Delta_h^n := \Delta_h \circ \dots \circ \Delta_h$, for any natural n, leads to $(\Delta_h^n x)(t) = h^{-n} \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} x(t+kh)$. Additionally, we have $(\Delta_h^0\, x)(t) := x(t)$. The fractional h −sum of order $\alpha > 0$ for a function $x: (hN)_a \to R$ is defined by:

$$\left({}_a\Delta_h^{-\alpha} x \right)(t) := h^\alpha (\varphi_\alpha * \bar{x})(n),$$

where $t = a + (\alpha + n)h$ for any natural n.

Let $\alpha \in (0,1]$. The *Caputo-type* h *−difference operator* ${}_a\Delta_{h,*}^{\alpha}$ of order α for a function $x: (hN)_a \to R$ is defined as (Mozyrska and Girejko, 2013):

$$({}_a\Delta_{h,*}^{\alpha} x)(t) := ({}_a\Delta_h^{-(1-\alpha)} (\Delta_h x))(t) \tag{5}$$

for any $t \in (hN)_{a+(1-\alpha)\,h}$. If $\alpha = 1$, then $({}_a\Delta_{h,*}^{\alpha=1} x)(t) = (\Delta_h\, x)(t)$ for any $t \in (hN)_a$. Note that $({}_a\Delta_{h,*}^{\alpha} x)(t) = h^{-\alpha}(\varphi_{1-\alpha} * \Delta_{h=1}\bar{x})(n)$ for any $t = a + (1-\alpha)h + nh$ and $\bar{x}(n) = x(a+nh)$.

The *Riemann–Liouville-type fractional* h *−difference operator* ${}_a\Delta_h^{\alpha}$ of order $\alpha \in (0,1]$ for a function $x: (hN)_a \to R$ is defined as (Bastos et al., 2011; Fereira and Torres, 2011):

$$\left({}_a\Delta_h^{\alpha} x \right)(t) := (\Delta_h\left({}_a\Delta_h^{-(1-\alpha)} x \right))(t),$$

where $t \in (hN)_{a+(1-\alpha)h}$.

The last operator we are considering is the *Grünwald–Letnikov-type fractional* h *−difference operator* ${}_\mathrm{a}\widetilde{\Delta}_\mathrm{h}^\alpha$ of a real order α, defined for a function $x: (hN)_a \to R$ as (Mozyrska et al., 2013):

$$({}_a\widetilde{\Delta}_h^{\alpha} \mathrm{x})(t) := \sum_{s=0}^{\frac{t-a}{h}} a_s^{(\alpha)}\, x(t-sh),$$

where $a_s^{(\alpha)} = (-1)^s \binom{\alpha}{s} \frac{1}{h^\alpha}$. If $a = (\alpha-1)h$. Then

$$({}_0\widetilde{\Delta}_h^{\alpha} \mathrm{y})(t+h) = \left({}_a\Delta_h^{\alpha} x \right)(t), \tag{6}$$

where $x(t) = y(t-a)$ for $t \in (hN)_a$ (Mozyrska et al., 2013). Also, in Mozyrska et al. (2013), it was shown that for $\alpha \in (0,1]$,

$$({}_a\Delta_{h,*}^{\alpha} x)(t) = \left({}_a\Delta_h^{\alpha} x \right)(t) - \tfrac{x(a)}{h^\alpha} \binom{\frac{t-a}{h}}{-\alpha} \tag{7}$$

for $t \in (hN)_{a+(1-\alpha)h}$. Taking into account relations (6) and (7), one can use the common symbol defined by its values:

$$({}_a\Upsilon_h^{\alpha} x)(t) = \begin{cases} ({}_a\Delta_{h,*}^{\alpha} x)(t)\; or \left({}_a\Delta_h^{\alpha} x \right)(t) & for\, a = (\alpha-1)h \\ ({}_a\widetilde{\Delta}_h^{\alpha} \mathrm{x})(t+h) & for\, a = 0. \end{cases}$$

Recall that the single-sided Z −transform of a sequence $\{y(n)\}_{n\in N_0}$ is a complex function $Y(z)$ given by $Y(z) := Z[y](z) = \sum_{k=0}^{\infty} \frac{y(k)}{z^k}$, where z is a complex variable for which series $\sum_{k=0}^{\infty} \frac{y(k)}{z^k}$ converges absolutely.

Proposition 3 (Mozyrska and Wyrwas, 2015): Let $a \in R$ and $\alpha \in (0,1]$. Define $y(n) := ({}_a\Upsilon_h^{\alpha} x)(t)$, where $t \in (hN)_{a+(1-\alpha)h}$ and $t = a + (1-\alpha)h + nh$. Then:

$$Z[\left({}_a\Upsilon_h^{\alpha} x)(t)\right](z) = z\left(\tfrac{hz}{z-1}\right)^{-\alpha} (X(z) - x(a)), \tag{8}$$

where $X(z) = Z[\bar{x}](z)$, $\bar{x}(n) := x(a+nh)$ and $\beta = \alpha$ for the Riemann–Liouville- or Grünwald–Letnikov-type h −difference operators and $\beta = 1$ for the Caputo-type h −difference operator, and $a = \alpha - 1$ for the Riemann–Liouville- or Caputo-type operators and $a = 0$ for the Grünwald–Letnikov-type operator.

Proposition 4 (Mozyrska and Wyrwas, 2015): Let $\alpha \in (0{,}1]$. Then $Z\left[E_{(\alpha,\beta)}(\lambda,\cdot)\right](z) = \left(\frac{z}{z-1}\right)^{\beta}\left(1-\frac{\lambda}{z}\left(\frac{z}{z-1}\right)^{\alpha}\right)^{-1}$, where $|z| > 1$ and $|z-1|^{\alpha}|z|^{1-\alpha} > |\lambda|$. Additionally $\beta = \alpha$ for the Riemann–Liouville- or Grünwald–Letnikov-type h −difference operators and $\beta = 1$ for the Caputo-type h −difference operator.

3. LINEAR FRACTIONAL VECTOR-ORDER SYSTEMS

Let us consider the following common form of vector-order $\alpha = (\alpha_1, \dots, \alpha_p)$, $\alpha_i \in (0{,}1]$, $i = 1, \dots, p$, linear control systems initialised at time $t_0 \in (hN)_{t_0}$:

$$({}_{t_0}\Upsilon_h^{\alpha} x)(t) = Ax(t + t_0) + Bu(t) \tag{9a}$$

$$y(t) = Cx(t + t_0) \ , \tag{9b}$$

where $x\colon (hN)_{t_0} \to R^p$ denotes a state vector, $y\colon (hN)_0 \to R^r$ an output vector, $u\colon (hN)_0 \to R^m$ a control, and $A \in R^{p\times p}$, $B \in R^{p\times m}$ and $C \in R^{r\times p}$ are real stationary matrices. Equation (9a) defines the dynamics of system (9) and Equation (9b) its output. Since matrices A, B, C for given $\alpha = (\alpha_1, \dots, \alpha_p)$ completly determinate system (9), shortly we will say that this system is described by the triple (A, B, C) for a given α. From definitions of fractional h-difference operators $({}_a\Delta_{h,*}^{\alpha} x)(t)$, $({}_a\Delta_h^{\alpha} x)(t)$ and $({}_a\widetilde{\Delta}_h^{\alpha} \mathrm{x})(t + h)$, it follows that dynamics (9a) can be rewritten as:

$$({}_{t_i}\Upsilon_h^{\alpha_i} x)(t_i) = h^{\alpha_i}\sum_{j=1}^{p} A_{ij}x_j\left(t_i + t_{0_i}\right) + h^{\alpha_i}\sum_{\iota=1}^{m} B_\iota u\_\iota(t_i)$$

for $i = 1, \dots, p$. Let $g^{\rho}(t) := g(t-h)$ for any $t \in (hN)_{t_0}$. For matrix $A \in R^{p\times p}$, define

$$E_{(\alpha,\beta)}(A,p) \coloneqq diag\{E_{(\alpha_i,\beta_i)}(A,p) \colon i = 1, \dots, p\}. \tag{10}$$

Example 5: Let $A = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$, $p = 2$ and $\alpha = (\alpha_1, \alpha_2)$, such that $\alpha_1 = \beta_1 = 0{,}5$ and $\beta_2 = 0{,}5\alpha_2 = 0{,}5$. Then, for a positive h, we have:

$$E_{(\alpha,\beta)}(A,2) = diag\left\{E_{(\alpha_1,\beta_1)}(A,2), E_{(\alpha_2,\beta_2)}(A,2)\right\} =$$

$$diag\{\begin{pmatrix} \frac{3}{8} + (h^{0,5}-1)h^{0,5} & (1-2h^{0,5})h^{0,5} \\ 0 & \frac{3}{8} + (h^{0,5}-1)h^{0,5} \end{pmatrix}, \begin{pmatrix} \frac{3}{8} + (h^{0,25}-1)h^{0,25} & \left(\frac{3}{8} - 2h^{0,25}\right)h^{0,25} \\ 0 & \frac{3}{8} + (h^{0,25} - \frac{3}{4})h^{0,25} \end{pmatrix}\}.$$

Lemma 6: Let $\alpha = (\alpha_1, \dots, \alpha_p)$, $\alpha_i \in (0{,}1]$, $i = 1, \dots, p$ and $\bar{A} = diag\{h^{-\alpha_i}A \colon i = 1, \dots, p\}$, $\bar{B} = \begin{pmatrix} H^{-1}B & 0_{p\times m} & \cdots & 0_{p\times m} \end{pmatrix}^T$, $H := diag(h^{-\alpha_i} \colon i = 1, \dots, p)\}$. Dynamics of system (9) together with initial state $x_0 = \left(x(t_{0_1}) \quad \dots \quad x\left(t_{0_p}\right)\right)^T = (x_{0_1} \quad \cdots \quad x_{0_p})^T = x_0 \in R^p$, $t_{0_i} = (\alpha_i - 1)h$, $i = 1, \dots, p$, and fixed controls u_ι, $\iota = 1, \dots, m$ has the unique solution:

$$x(t) = E_{(\alpha,\beta)}\left(\bar{A}, \frac{t-t_0}{h}\right)x_0 + \left(E_{(\alpha,\alpha)}^{\rho}(\bar{A},\cdot) * \bar{B}\bar{u}\right)\left(\frac{t-t_0}{h}\right),$$

where $\bar{u}\left(\frac{t-t_0}{h}\right) = h^{\alpha}u(t)$, and $\beta = 1$ for the fractional h −difference of Caputo-type operator, $\beta = \alpha$ for the fractional h −differences of Riemman–Liouville- and Grünwald–Letnikov-type operators.

Proof: Taking the Z −transform of both sides of Equation (9a), from Proposition 3 it follows that:

$$zh^{-\alpha_i}\left(1-\tfrac{1}{z}\right)^{\alpha_i}\left(X_i(z) - \left(\tfrac{z}{z-1}\right)^{\beta_i} x_i(t_{0_i})\right) = \sum_{j=1}^{p} A_{ij}X_j(z) + \sum_{\iota=1}^{m} B_\iota U_\iota(z), \tag{11}$$

where $Z\left[\bar{x}_i\left(\frac{t_i-t_{0_i}}{h}\right)\right](z) = X_i(z)$, $U_\iota\left[u_\iota\left(\frac{t_i-t_{0_i}}{h}\right)\right](z) = U\iota(z)$ for $i = 1, \dots, p$ and $\iota = 1, \dots, m$. Denoting $X(z) = (X_1(z) \quad \dots \quad X_p(z))^T$, $\Lambda_\alpha = diag\left\{\left(\frac{z}{z-1}\right)^{\alpha_i} \colon i = 1, \dots, p\right\}$ and $\Lambda_\beta = diag\left\{\left(\frac{z}{z-1}\right)^{\beta_i} \colon i = 1, \dots, p\right\}$, equation (11) can be rewritten as:

$$X(z) = \left(I_n - \frac{1}{z}\Lambda_\alpha HA\right)^{-1}\Lambda_\beta x_0 + \frac{1}{z}\left(I_n - \frac{1}{z}\Lambda_\alpha HA\right)^{-1}\Lambda_\alpha HBU(z).$$

If we put $F_1(z) = \left(I_n - \frac{1}{z}\Lambda_\alpha HA\right)^{-1}\Lambda_\beta$ and $F_2(z) = \left(I_n - \frac{1}{z}\Lambda_\alpha HA\right)^{-1}\Lambda_\alpha$, then $X(z) = F_1(z)x_0 + F_2(z)HB$. So, $\bar{x}(n) = Z^{-1}[X(z)]\left(\frac{t-t_0}{h}\right) = Z^{-1}[F_1(z)]\left(\frac{t-t_0}{h}\right)x_0 + Z^{-1}[F_2(z)U(z)]\left(\frac{t-t_0}{h}\right)$.

Since H, Λ_α and Λ_β are diagonal matrices, then by Proposition 4, one has $x_i\left(t_i + t_{0_i}\right) = E_{(\alpha_i,\beta_i)}\left(h^{-\alpha_i}A, \frac{t-t_0}{h}\right)x_0 + \left(E_{(\alpha_i,\alpha_i)}^{\rho}\left(\overline{h^{-\alpha_\iota}A},\cdot\right) * h^{-\alpha_i}B\bar{u}\right)\left(\frac{t-t_0}{h}\right)$. Taking into account (10), one obtains thesis. □

By $J_0(m)$, let us denote the set of all sequences $U = (u_0, u_1, \dots)$, where $u_n := u(t) = u(nh + t_0) \in \Omega$, $t \in (hN)_{t_0}$. Then, $\gamma(t + t_0, x_0, U) := x(t + t_0)$ will denote the state forward trajectory of system (9), i.e. a solution which is uniquely defined by the initial state x_0 and the control sequence $U \in J_0(m)$. The reachable set from the given initial state x_0 in q steps, denoted as $R_q(x_0)$, is the set of all states to which the given system can be steered from x_0 in q steps by the control sequence $U \in J_0(m)$, i.e. $R_q(x_0) := \{x \in R^p \colon x = \gamma(q, x_0, U), U \in J_0(m)\}$ with $R_0(x_0) := \{x_0\}$. Then, the set $R(x_0) := \bigcup_{q\in N_0} R_q(x_0)$ is the set of all states reachable from x_0.

Definition 7: System (9) is *locally controllable in* q *steps* from x_0 if there exists a neighbourhood $V \subset R^n$ of x_0, such that $V \subset R_q(x_0)$. System (9) is *globally controllable from* x_0 *in* q *steps* if $R_q(x_0) = R^p$.

Proposition 8: Let $\alpha = (\alpha_1, \dots, \alpha_p)$ with $\alpha_i \in (0,1]$, $i = 1, \dots, p$. Then system (9) is controllable in q steps if and only if the rank of controllability matrix $Q_q = \begin{pmatrix} \bar{\boldsymbol{B}} & E_{(\alpha,\alpha)}(\bar{A},1)\bar{B} & \dots & E_{(\alpha,\alpha)}(\bar{A},q-1)\bar{B} \end{pmatrix}$ is full, i.e. $rankQ_q = p$.

Proof: The result is the consequence of Lemma 6. The reasoning is similar to the one in a scalar fractional order case in Mozyrska et al. (2017). □

From the Rank Matrix Theorem and Proposition 8, it follows that $rankQ_q = p$ if and only if $q = p$. So, any state $x_0 \in R^p$ can be steered to a final state $x_f \in R^n$ in no more then p steps.

Definition 9: System (9) is *observable in* q *steps* if from the control sequence $U = \begin{pmatrix} u(0) \\ u(1) \\ \vdots \\ u((q-1)h) \end{pmatrix}$ and the output sequence $Y = \begin{pmatrix} y(0) \\ y(1) \\ \vdots \\ y((q-1)h) \end{pmatrix}$ it is possible to determinate uniquely initial state x_0 of the given system.

Proposition 9: Let $\alpha = (\alpha_1, \dots, \alpha_p)$ with $\alpha_i \in (0,1]$, $i = 1, \dots, p$. Then, system (9) is observable in q steps if and only if the rank of observability matrix $W_q = \begin{pmatrix} \bar{C} \\ \bar{C}E_{(\alpha,\beta)}(\bar{A},1) \\ \vdots \\ \bar{C}E_{(\alpha,\beta)}(\bar{A},q-1) \end{pmatrix}$, where $\bar{C} := \begin{pmatrix} C & 0_{r\times p} & \dots & 0_{r\times p} \end{pmatrix}$ is full, i.e. $rankW_q = p$.

Proof: The result is the consequence of Lemma 6. The reasoning is similar to the one in a scalar fractional order case in Mozyrska et al. (2017). □

From the Rank Matrix Theorem and Proposition 10, it follows that $rank\, W_q = p$ if and only if $q = p$. So, based on the knowledge of control and output measurable sequences U and Y, respectively, the initial state $x_0 \in R^p$ can be uniquely determined in no more then p steps.

Controllable and observable triple (A, B, C) is called *canonical triple.*

4. REALISATION PROBLEM OF THE GIVEN IMPULSE

Consider system (9) with the initial state $x(t_0) = x_0$, vector-order $\alpha = (\alpha_1, \dots, \alpha_p)$, $\alpha_i \in (0,1]$, $i = 1, \dots, p$ and given positive h. Observe that for any input $u: (hN)_{t_0} \to R^m$ and $t \geq t_0, t \in (hN)_{t_0}$, the following holds:

$$y(t) = \sum_{s=0}^{q=\frac{t-t_0}{h}} \Psi_\Lambda(t-sh)u(sh) = (\Psi_\Lambda * u)(t), \tag{12}$$

where $\Psi_\Lambda(t) = CE_{(\alpha,\beta)}(A, \frac{t-t_0}{h})B$, and $\beta = 1$ for the fractional h –difference of Caputo-type operator, $\beta = \alpha$ for the fractional h –differences of Riemman–Liouville- and Günwald–Letnikov-type operators. Function Ψ_Λ is called the *impulsive response* of system (9). Formula (12) defines the relation $S_{\Psi,q}$ between the input u and output y in q steps of the given control system, i.e.:

$$S_{\Psi,q}(u) = y. \tag{13}$$

Map $S_{\Psi,q}$ is called the *(q step) input–output map* of the considered system. Observe that between the impulsive response and the input–output map, there is a mutually inverse correspondence.

Suppose that $S_{\Psi,q}$ is an abstract q –steps input–output map acting on the input function u as

$$S_q(u) = \sum_{s=0}^{q=\frac{t-t_0}{h}} \Psi(t-sh)u(sh) = (\Psi * u)(t), \tag{14}$$

where map $\Psi: t \mapsto \Psi(t)$ is defined for all $t \in (hN)_{t_0}$. The problem is: *find a fractional vector-order state-space representation of map* S_q *in q steps.* In other words, for a chosen real positive h, we are looking for a linear fractional vector−order $\alpha = (\alpha_1, \dots, \alpha_p)$ control system (A, B, C), such that maps S_q and $S_{\Psi,q}$ coincide.

For the given abstract input–output map $S_q(u) = (\Psi * u)(t)$, define h –Markov parameters as

$$M_n^h := \Psi(nh + t_0) = \Psi(t),\ t \in (hN)_{t_0}. \tag{15}$$

Sequence $M^h = \{M_n^h : n \in N_0\}$ with elements M_n^h given by (15) will be called *h-Markov sequence* and its elements as *h-Markov parameters.*

Theorem 11: Let $h > 0$ and $\alpha = (\alpha_1, \dots, \alpha_p)$ with $\alpha_i \in (0,1]$, $i = 1, \dots, p$. Function $\Psi(t) = \Psi(nh + t_0) = \sum_{n=0}^{\infty} M_n^h$ is an impulsive characteristic of the fractional vector-order $\alpha = (\alpha_1, \dots, \alpha_p), \alpha_i \in (0,1]$, $i = 1, \dots, p$ system given by triple (A, B, C) if and only there are natural $\tilde{p}$ and real $a_0, a_1, \dots, a_{\tilde{p}-1}$, such that the following recursive relation holds:

$$M_{\tilde{p}+j}^h + a_{\tilde{p}-1}M_{\tilde{p}+j-1}^h + \dots + a_1 M_{1+j}^h + a_0 M_j^h = 0. \tag{16}$$

for $j = 0,1,2,3, \dots$.

Proof: Suppose that the input–output map (14) is a realisation of the fractional vector-order α system (A, B, C). Denote $\overline{\Psi}(n) := \Psi(nh + t_0)$ for any $n \in N$. So, there is α_i, $i = 1, \dots, p$, such that $\overline{\Psi}_i(p) = \bar{C}E_{(\alpha_i,\beta_i)}(\bar{A}, p)\bar{B}$ for some $\bar{A} := \mathrm{diag}\{\mathrm{h}^{-\alpha_i}\mathrm{A} : \mathrm{i} = 1, \dots, \mathrm{p}\}$, $\bar{B} := \begin{pmatrix} H^{-1}B & 0_{p\times m} & \dots & 0_{p\times m} \end{pmatrix}^{\mathrm{T}}$, $\bar{C} := \begin{pmatrix} C & 0_{r\times p} & \dots & 0_{r\times p} \end{pmatrix}$ with $H = diag\{h^{-\alpha_i} : i = 1, \dots, p\}$. Thus, $M_p^h = \overline{\Psi}_i(p) = \bar{C}E_{(\alpha_i,\beta_i)}(\bar{A}, p)\bar{B}$. By Proposition 2 and formula (10) for any natural j, the following holds: $\bar{C}\left[E_{(\alpha,\beta)}(\bar{A}, p)\right]^{\tilde{p}+j}\bar{B} + a_{p-1}\bar{C}\left[E_{(\alpha,\beta)}(\bar{A}, p)\right]^{\tilde{p}+j-1}\bar{B} + \dots + a_0\bar{C}\left[E_{(\alpha,\beta)}(\bar{A}, p)\right]^{j}\bar{B} = 0.$

Hence, (16) is fulfilled.

Now suppose that (16) holds for the given $\alpha = (\alpha_1, \dots, \alpha_p)$ with $\alpha_i \in (0,1]$, $i = 1, \dots, p$. Then, for

$$A = \begin{pmatrix} 0_{p\times p} & 0_{p\times p} & \dots & -a_0 h^{\alpha_1} I_p \\ h^{\alpha_2} I_p & 0_{p\times p} & \dots & -a_1 h^{\alpha_2} I_p \\ 0_{p\times p} & h^{\alpha_3} I_p & \dots & -a_1 h^{\alpha_2} I_p \\ \vdots & \vdots & \dots & \vdots \\ 0_{p\times p} & 0_{p\times p} & \dots & -a_p h^{\alpha_p} I_p \end{pmatrix} \tag{17}$$

and

$$B = \begin{pmatrix} H^{-1} \\ 0_{p\times p} \\ \vdots \\ 0_{p\times p} \end{pmatrix} \tag{18}$$

$$C = \begin{pmatrix} M_0^h & M_1^h & \dots & M_{p-1}^h \end{pmatrix}, \tag{19}$$

one obtains $M_p^h = CE_{(\alpha,\beta)}\left(A, p = \frac{t-t_0}{h}\right)B$. Hence, matrices A, B, C given by (17)–(19) define a realisation of the map $S_{\Psi,q}$. □

Example 12: Suppose that h is any real positive number, $t_0 = 0$ and $\alpha = (\alpha_1, \alpha_2)$, such that $\alpha_i \in (0,1]$, $i = 1,2$ and $\alpha_1 = \alpha_2$. Let $\Psi(t) = \Psi(nh) = \sum_{k=0}^{n=\frac{t}{h}} h^{-2\alpha_1} \binom{-k\alpha_1 - \beta_1}{n-k}$. Therefore, $M_n^h = \sum_{k=0}^{n} h^{-2\alpha_1} \binom{-k\alpha_1 - \beta_1}{n-k}$. Then, $M_{2+j}^h - M_{1+j}^h + \left(1 - h^{-\alpha_1} - \frac{\beta(\beta_1+1)}{2} h^{-\alpha_1}\right) M_j = 0$. The realisation for fractional h –differences of Riemman–Liouville- and Grünwald–Letnikov-type operators is given by the matrices $A = \begin{pmatrix} 0 & \frac{\alpha_1(\alpha_1+1)}{2} h^{2\alpha_1} \\ h^{\alpha_1} & -1 \end{pmatrix}$, $B = \begin{pmatrix} h^{-\alpha_1} \\ 0 \end{pmatrix}$ and $C = \begin{pmatrix} h^{-\alpha_1} & h^{-\alpha_1}(1 - 2\alpha_1) \end{pmatrix}$, and for fractional h –difference of Caputo-type operator, it is given by $A = \begin{pmatrix} 0 & h^{2\alpha_1} \\ h^{\alpha_1} & -1 \end{pmatrix}$, $B = \begin{pmatrix} h^{-\alpha_1} \\ 0 \end{pmatrix}$ and $C = \begin{pmatrix} h^{-\alpha_1} & -\alpha_1 h^{-\alpha_1} \end{pmatrix}$.

For a given h –Markov sequence M^h and positive integers s, ν, the block matrix

$$H_{s\nu}(M^h) = \begin{pmatrix} M_1^h & M_2^h & \dots & M_\nu^h \\ M_2^h & M_3^h & \dots & M_{\nu+1}^h \\ \vdots & \vdots & \ddots & \vdots \\ M_\nu^h & M_{\nu+1}^h & \dots & M_{s+\nu-1}^h \end{pmatrix} \tag{20}$$

is called the *Hankel matrix* associated with the sequence M^h.

Proposition 13: Let the triple (A, B, C) be described by (9a)–(9b). Then,

1. (A, B, C) realises M^h in q steps if and only if $W_q Q_q = H_{s\nu}(M^h)$ for all $s, \nu \in N$.
2. If (A, B, C) realises M^h in q steps and (A, B, C) is a canonical triple, then $rank\ H_{s\nu}(M^h) = p$ for $s, \nu \geq p$.

Proof: The result follows directly from propositions 8 and 10. □

In general, realisations are not unique. From a practical point of view, it is good to have such realisation for which the state-space has the possible minimal dimension, i.e. it is good to have a *minimal realisation*. This property is not easy for checking, but classically it is equivalent to the fact that triple (A, B, C) realising the h –Markov sequence should be canonical.

Theorem 14: If there is a realisation of h –Markov sequence M^h, then it is the canonical realisation.

Proof: The idea of the proof comes from Bartosiewicz and Pawluszewicz (2006). Suppose that system (9) is not controllable in a finite number of steps. So, there exist a natural number p_1 and a nonsingular matrix $P \in R^{p \times p}$, such that $P^{-1}AP = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}$ with $A_{11} \in R^{p_1 \times p_1}$, $A_{12} \in R^{p_1 \times (p-p_1)}$ and $A_{22} \in R^{(p-p_1) \times (p-p_1)}$. Let $\bar{A}$ and $\bar{B}, \bar{C}$ be defined as in Lemma 6. So, by Proposition 2, it follows that $P^{-1}E_{(\alpha,\beta)}(\bar{A}, p)P = \begin{pmatrix} E_{(\alpha,\beta)}(\bar{A}_{11}, p) & E_{(\alpha,\beta)}(\bar{A}_{12}, p) \\ 0 & E_{(\alpha,\beta)}(\bar{A}_{22}, p) \end{pmatrix}$. Also, $P\bar{B} = \begin{pmatrix} \bar{B} \\ 0 \end{pmatrix}$ with $\bar{B}_1 \in R^{p_1 \times m_1}$. So, $\bar{C}E_{(\alpha,\beta)}(\bar{A}, p)\bar{B} = \bar{C}PE_{(\alpha,\beta)}(\bar{A}, p)P^{-1}\bar{B} = \bar{C}_1 PE_{(\alpha,\beta)}(\bar{A}_{11}, p)P^{-1}\overline{B_1}$ for some matrix $\bar{C}$. So, system (9) is controllable in a finite number of steps.

The reasoning that system (9) is observable is the same. □

Corollary 15: A realisation of h –Markov sequence M^h is minimal if and only if it is canonical.

Proof: The implication " ⇒ " is the direct consequence of Proposition 13. The implication " ⇐ " follows from Theorem 14. □

5. CONCLUSIONS

The problem of realisation of the impulsive response function for fractional vector-order discrete time linear control systems was considered. It is shown that an abstract input–output map has a state-space realisation if and only if the h –Markov parameters satisfy the recurrence relation given by (16). This result extends the classical realisability criterion to fractional order systems. The description of state-space representation of input–output map is given in terms of fractional vector-order h –differences of Caputo-, Riemann–Liouville- and Grünwald–Letnikov-type operators. It is shown that the minimal fractional vector-order realisation exists if and only if triple $(\mathrm{A}, \mathrm{B}, \mathrm{C})$ defining the state-space system is controllable and observable. Obtained results are illustrated by an academic example. Further work will focus on practical implementation of the obtained results in physical systems, including automatic control systems. Therefore, in a natural way also, more practical examples will appear.

REFERENCES

1. **Ambroziak L., Lewon D., Pawluszewicz E**. (2016), The use of fractional order operators in modeling of RC-electrical systems, *Control & Cybernetics*, 45(3), 275--288
2. **Bartosiewicz Z., Pawluszewicz E**. (2006), Realizations of linear control systems on time scales, *Control and Cybernetics*, 35(4), 769-786.
3. **Bastos, N.R.O., Ferreira, R.A.C., Torres, D.F.M.** (2011), Necessary optimality conditions for fractional difference problems of the calculus of variations, *Discrete and Continuous Dynamical Systems*, Vol.29(2), 417-437.
4. **Bettayeb M., Djennoune S., Guermah S., Ghanes M**. (2008), Structural properties of linear discrete-time fractional order systems, *Proc. of the 17th World Congres The Federation of Automatic Control*, Seoul, Korea, 15262—15266.
5. **Das S.** (2008), *Functonal Fractional Calculus for System Identyfication and Controls*, Springer.
6. **Ferreira, R.A.C., Torres, D.F.M.** (2011) Fractional h –difference equations arising from the calculus of variations, *Applicable Analysis and Discrete Mathematics*, 5(1), 110-121.
7. **Gantmacher F.R**. (1959), *The Theory of Matrices*, Chelsea Pub. Comp., London.
8. **Kaczorek T.** (1998), *Vectors and Matrices in Automation and Electrotechnics*, WNT Warsaw 1998 (in Polish).
9. **Kaczorek T**. (2017), Cayley-Hamilton theorem for fractional linear systems, *Theory and Applications of Non-integer Order Systems*, LNEE 407, 45-55, Springer.
10. **Koszewnik A., Nartowicz t., Pawluszewicz E.** (2016), Fractional order controller to control pump in FESTO MPS® PA Compact Workstation, *Proc. of the Int.Carpathian Control Conference*, 364-367.
11. **Mozyrska D., Girejko E**. (2013), Overview of the fractional h –differences operator, in: Almeida A., Castro L., Speck FO. (eds) Advances in Harmonic Analysis and Operator Theory, *Operator Theory: Advances and Applications*, 229, Birkhäuser, Basel.
12. **Mozyrska D., Girejko E., Wyrwas M.** (2013), Comparision of h –difference fractional operatots, *Theory and Appication of Non-integer Order systems*, LNEE 257, 191-197, Springer.

13. **Mozyrska D., Wyrwas M., Pawluszewicz E.** (2017), Local observability and controllability of nonlinear discrete-time fractional order systems based on their linearisation, *International Journal of Systems Science*, 48(4), 788--794
14. **Mozyrska, D., Wyrwas, M.** (2015), The Z-Transform Method and Delta Type Fractional Difference Operators, *Discrete Dynamics in Nature and Society*, Article ID 852734, 12 pages
15. **Oprzędkiewicz K., Gawin E**. (2016), A non integer order, state space model for one dimensional heat transfer process, *Archives of Control Sciences*, 26(2), 261-275
16. **Pawluszewicz E., Koszewnik A.** (2019), Markov parameters of the input-output map for discrete-time fractional order systems, *Proc. of the 24th Int.Conf. on Methods and Models in Automation and Robotics MMAR'2019*, Miedzyzdroje, Poland
17. **Podlubny I.** (1999), *Fractional differential systems*, Academic Press, San Diego
18. **Sierociuk D., Dzieliński A., Sarwas G., Petras I., Podlubny I., Skovranek T.** (2013), Modelling heat transfer in heterogenous media using fractional calculus, Phylosophical Transaction of the Royal Society A-Mathematical, *Physical and Engineering Sciences*, Soc A 371: 20120146, 10 pages
19. **Sontag E.** (1998), *Mathematical Control Theory*, Springer – Verlag
20. **Wu G.Ch., Baleanu D., Zeng S.D., Luo W.H.** (2016), Mittag-Leffler function for discrete fractional modelling, *Journal of King Saud University - Science*, 28, 99—102
21. **Wu, D.Baleanu, Zeng S.D., Deng Z.G.** (2015), Discrete fractional diffusion equation, *Nonlinear Dynamics*, 80, 281–286
22. **Zabczyk J.** (2008), *Mathematical Control Theory*, Birkhaüser.

DOI 10.2478/ama-2020-0017

PREDICTIVE NEURAL NETWORK IN MULTIPURPOSE SELF-TUNING CONTROLLER

Oleksiy BONDAR[*]

[*]Electronic Microwave Devices Department, Institute of Radio Astronomy of the National Academy of Sciences of Ukraine,
4 Mystetstv St., Kharkiv, 61002, Ukraine

bondarrian@gmail.com

received 16 April 2020, revised 20 July 2020, accepted 22 July 2020

Abstract: A very important problem in designing of controlling systems is to choose the right type of architecture of controller. And it is always a compromise between accuracy, difficulty in setting up, technical complexity and cost, expandability, flexibility and so on. In this paper, multipurpose adaptive controller with implementation of artificial neural network is offered as an answer to a wide range of tasks related to regulation. The effectiveness of the approach is demonstrated by the example of an adaptive thermostat. It also compares its capabilities with those of classic PID controller. The core of this approach is the use of an artificial neural network capable of predicting the behaviour of controlled object within its known range of parameters. Since such a network, being trained, is a model of a regulated system with arbitrary precision, it can be analysed to make optimal management decisions at the moment or in a number of steps. Network learning algorithm is backpropagation and its modified version is used to analyse an already trained network in order to find the optimal solution for the regulator. Software implementation, such as graphical user interface, routines related to neural network and many other, is done using Java programming language and Processing open-source integrated development environment.

Keywords: Artificial neural network, adaptive regulator, backpropagation algorithm, system modelling

1. INTRODUCTION

Nowadays, the use of artificial neural networks in control processes and other activities is a very popular area of research. And this is understandable, since these structures are a good tool in the modelling of complex systems, for which it is difficult to find a simple mathematical solution. It should also be remembered that artificial neural networks are capable of providing high quality data processing even in the conditions of their incompleteness. Such a predicting model is successfully being using for weather forecasting in short term for some localities (Wica et al., 2019). It can also be useful to provide neural classification mechanism in genetics researches (Liu et al., 2019; MacLean, 2019). The usage of artificial neural network (hereinafter referred to as ANN) is a reasonable solution for power flow regulators (Ma et al., 2018). Also, ANN can be used as a tool for detecting stable equivalent series resistance (ESR) in voltage regulator characterization (Zaman et al., 2018), in mechatronic hydraulic drive regulation (Burennikov et al., 2017) or autopilot (Zhao et al., 2018). However, researchers are particularly curious about the possibility of using artificial neural networks in the automatic tuning of PID regulators (Ayomoh and Ajala, 2012; Hernández-Alvarado et al., 2016; Pirabakaran and Becerra, 2002; Zhang et al., 2016; Du et al. 2018; Han et al., 2017). The method proposed in this paper excludes the PID section from the controller system. As will be shown, a trained neural network with a multi-step error estimation module is sufficient for high quality control of a wide range of systems providing flexible controls. Moreover, the network does not require any specific or detailed data, rather accurate system data in the regulatory range. To demonstrate the principle operation of the multipurpose controller as thermostat, a software model of the solid body temperature under the influence of external factors was created. Such an example is simple to understand and easily portable to a real thermostat or to another type of system. For the software implementation, Processing integrated development environment (IDE) was selected, which uses a Java programming language. It speeds up and facilitates both writing and debugging of programs for which it is important to have as many graphical evaluation options as possible.

2. DESCRIPTION OF MULTIPURPOSE NEURONAL NETWORK-BASED CONTROLLER

In order to understand the principle of operation of a multipurpose adaptive controller, it is necessary to consider its work in stages with a detailed analysis of the work of each element. We list these stages. Primary training of predictive ANN, during which the regulator acquires the ability to assess the state of regulated object in the future, based on the state of that object at the present moment. Further, the already trained ANN is embedded in a multi-step planning mechanism, where, thanks to the neural network's ability to approximate functions, we obtain a model of the object's behaviour depending on the actions of the controller itself on a certain number of steps forward. The predicted sequence of states of the control object over time is used to estimate the error relative to the desired control result for each of these states. Depending on the approach chosen in the assessment, these error values are used to optimize the controller's actions in the appropriate way. First of all, we will look at the core of the whole system – one-step predictive artificial neural network.

2.1. One-step predictive artificial neural network

The module is so called Feedforward Backpropagation Neural Network. Quite traditional architecture of network with neurons (knots) grouped into layers (see Fig.1). There are three types of layers: input, hidden, and output. The input layer is a set of corresponding signals that the module must process to obtain the output signal. In this case, the following signals are: RV – regulated value, RA – regulator action. Regulated value is the parameter that should be controlled such as temperature (for thermostats), speed or torque (for motor controller), distance, pressure or voltage and so on. Regulator action is any relevant action that affects a regulated value such as change of voltage, PWM, power, or any other parameter that the regulator may manipulate during control.

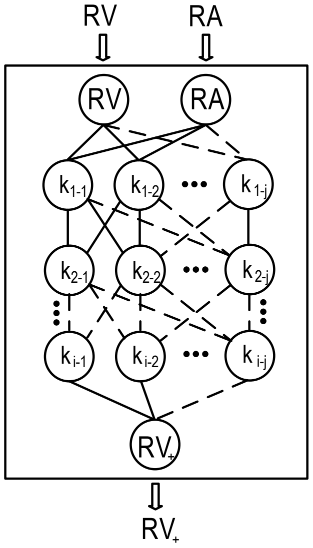


Fig. 1. The scheme of one-step predictive ANN

Hidden layers can contain a certain number of neurons in each layer and consist of a certain number of layers. The number of hidden layers depends on the complexity of the behaviour of the control object, but in most cases, one hidden layer is sufficient (Heaton, 2008). The number of neurons in each layer is defined as a balance between the precision with which the neural network as a model is able to reproduce the behaviour of the system on which it is trained and the complexity of implementing such a system in a software sense. Each additional neuron in the layer increases the accuracy with which the system can make predictions as a whole, but it also increases the computational power requirements of the equipment on which this network must operate. For each neuron of the hidden or output layers, the internal state of the neuron is first calculated based on the signals coming from the previous layer. As it was stated, the prediction module trains with the backpropagation algorithm, but there are no restrictions on using other algorithms such as genetic or imperialist competitive or others (Elsisi, 2019). Equation 1 shows how to calculate the internal state of the hidden or output neuron in layer with i neurons and previous layer with l neurons. All other neurons in this layer are calculated similarly by replacing the corresponding index i with the desired one.

$$state_{k_i} = \left(\sum\nolimits_l (w_{li} \cdot out_l)\right) + b_{k_i} \tag{1}$$

where: $state_{k_i}$– internal state of neuron in a row of 1 to i of hidden or output layer, l – the number of neurons in the previous layer, w_{li} – the weight of the connection between the calculated neuron and the output of neuron in a row of 1 to l from the previous layer (out_l), b_{k_i} – offset of the calculated neuron, makes it possible for the internal state of the neuron not to be 0 when the input signals are 0. After that, the output of the neuron is calculated. To do this, the resulting internal state of the neuron must be passed through the activation function. There are different variants of the activation function, but in this paper, a linear activation function is used for the output neuron, and a fast sigmoid (https://stackoverflow.com/questions/10732027/fast-sigmoid-algorithm) for the hidden layers. The linear activation function simply transmits to the output of the neuron its internal state and allows the output value of the neuron to not be limited by any asymptotes in its range. Which is useful for a network whose output values may reach values that other activation functions do not allow. The sigmoid, as a function of activating the hidden layers, on the contrary, limits the output values in the range from 0 to 1 and provides nonlinearity in the operation of the network. Nonlinearity is necessary where the behaviour of the system to be simulated cannot be reduced to the sum of linear functions, which is any more or less complex system. The fast sigmoid (see Eq. 2) is close to the original but significantly reduces the time to calculate the activation function since it does not require a floating point exponent.

$$out_{k_i} = \begin{cases} \frac{0.5}{1+state_{k_i}{}^2}\ , \ when \ \ state_{k_i} < 0 \\ 1 - \frac{0.5}{1+state_{k_i}{}^2}\ , \ when \ \ state_{k_i} \geq 0 \end{cases} \tag{2}$$

where: out_{k_i} – the output value of the neuron in a row of 1 to i of some hidden layer (from 1 to j, see Fig.1).

The output layer of this model consists of one neuron and forms the final result of the network. The output RV_+ itself is the prediction of a regulated value (RV) on the next step of regulation. It is worth noting that such a network may also contain some additional inputs and outputs, which on the one hand allow to expand the number of parameters being monitored for regulation, and on the other hand create additional opportunities for optimizing the controller's actions. But it should also be remembered that increasing the number of inputs and outputs of the network requires an increase in its complexity, and thus, the requirements for hardware to process it (Heaton, 2008).

2.2. Multi-step planning mechanism

After considering the one-step predictive ANN as a basic element of the planning mechanism, we move on to the multi-step mechanism as a whole.

As can be seen from Fig. 2, the mentioned mechanism is composed as a sequence of one-step forecasts made using the previously described neural network. Moreover, since the trained network is capable of making predictions for the original managed object within the range of values in which it was trained, only one forecasting network is needed for any long-range forecast in time. The first step in forecasting is based on the current data of regulated value (RV_0) and regulator action (RA_0). After this, the

DOI 10.2478/ama-2020-0017

forecast for the first adjustment step RV_1 becomes the input value for the prediction of the next step RV_2, and so on, down to some step n which limits the planning horizon. Increasing the planning horizon on one hand improves the dynamic characteristics of the controller, such as stability and the absence of overshoots in operation, on the other hand, it increases the time to calculate all the steps. Regarding the values of future control actions ($RA_1 \ldots RA_{n-1}$), prior to optimization, they may be equal to the current control or may be random in a certain range. Further, all the results predicted at each step are sent for estimation of a regulation error in the corresponding module.

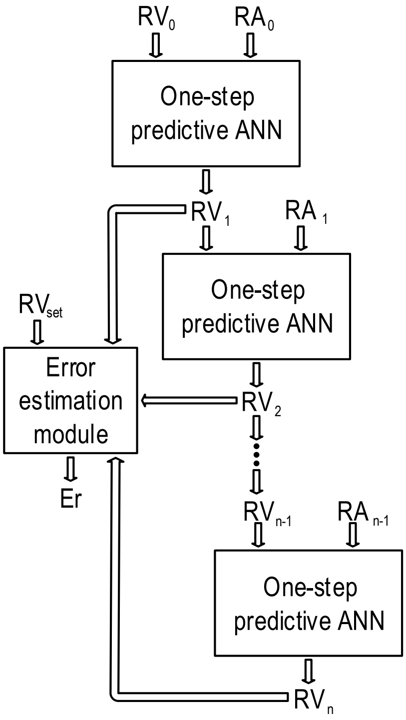


Fig. 2. The scheme of multi-step planning mechanism

2.3. Error estimation module

As the proposed multi-step controller must be able to make current decisions based on certain predictions of both behaviour of the system it manages and its actions, it must also adequately assess how its actions are approaching in the future achievement of the goal, avoiding oscillations and overshoots. This function is performed by the error estimation module. Taking into account the target value to be reached by the control object (RV_{set}), the controller calculates an error for each predicted step. The equation to calculate the regulation error in the first step of forecasting is:

$$Er_1 = \frac{(RV_1 - RV_{set})^2}{2} \tag{3}$$

where: Er_1 – regulation error in the first step of forecasting.

It should also be remembered that each new step of the prediction is based on the previous one, and the error of the neural network modelling of the source system will accumulate. Therefore, the regulator should not equally evaluate the result at each step. Equation 4 shows how it calculates the values of the regulation errors in the range from 2nd to n-th steps.

$$Er_n = \frac{(RV_n - RV_{set})^2}{2 \cdot (n-1) \cdot d} \tag{4}$$

where: Er_n – regulation error in the n-th step of forecasting, n – step number, d – initial depreciation, RV_n – regulated value in the n-th step of forecasting. As we can see, from the second step, the significance of each subsequent regulation error decreases in arithmetic progression. The initial depreciation is the factor in how many times the second forecast is less important for the regulator than the first one. This approach helps to increase the stability of the regulator and indicate that the first step is the most important to control because it influences further events by domino effect.

3. THE CONTROL ALGORITHM

After a schematic description of the proposed regulator, we proceed to the disclosure of the algorithm of its operation. First of all, let's say that the given controller can work on different algorithms of optimization of control action. This paper describes one such algorithm. A common feature of such algorithms is their similarity to the backpropagation method in neural network training. They are also iterative and use a gradient descent method with dividing into the forward and backward pass. But if backpropagation in a step-by-step training changes the parameters of the neural network, then the control algorithms below do not change the network itself, but use it to detect a correlation between the action of the controller and the error of regulation. Mathematically, the optimization step in general is shown in Equation 5.

$$RA' = RA - \lambda \cdot \frac{\partial Er}{\partial RA} \tag{5}$$

where: RA' – regulator action after one optimisation step, λ – optimization rate, Er – predicted regulation error (depends on the optimization strategy). Thus, at each step of the optimization of the control action, it changes by a value proportional to the instantaneous speed of change of the predicted regulation error with the change of the regulatory action. Moreover, λ should be small enough to ensure the smoothness and accuracy of the process, but not too small, as this will require a large number of iterations for a successful result. Further work is to calculate $\frac{\partial Er}{\partial RA}$. According to the chain rule, this task can be divided into two simpler ones (see Eq. 6).

$$\frac{\partial Er}{\partial RA} = \frac{\partial Er}{\partial RV} \cdot \frac{\partial RV}{\partial RA} \tag{6}$$

where: RV – predicted regulated value. So, calculation of $\frac{\partial Er}{\partial RV}$ allows the error estimation module. And the forecasting network allows to calculate $\frac{\partial RV}{\partial RA}$. But for this, we need to decide on a strategy by which we will evaluate regulatory errors and optimize the regulation action. So, let's describe the strategy of optimization of the maximum predicted regulation error.

3.1. Maximum Predicted Error Reduction Strategy (MAPERS)

The essence of the Maximum Predicted Error Reduction Strategy (hereinafter referred to as MAPERS) is to look for that step in the planning mechanism that predicts more regulation error than others, then change the planned action in this step to the side, which should reduce the error in that step (see Equation 5).

As can be seen from Fig. 3, in each forward pass of the optimization cycle, the multi-step planning mechanism generates a chain of forecasts from RV_1 to RV_n, which together with RV_{set} go to the input of the error estimation module. Here, the error of regulation of each of the predicted steps is calculated separately by Equations (3) and (4), after which the step with the maximum error is selected from all the steps (Er_{max}). The Er_{max} computation ends the forward pass and the back pass begins, the purpose of which is to change the step with the predicted Er_{max} to reduce it. The way of back pass is shown in the Fig. 4 (arrow pointing from Er_{max} to RA_{max}).

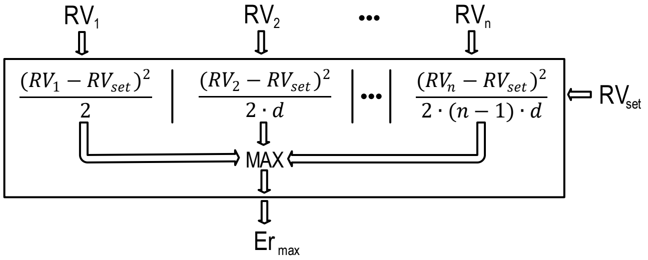


Fig. 3. Schematic diagram of the error estimation module for the Maximum Predicted Error Reduction Strategy (forward pass)

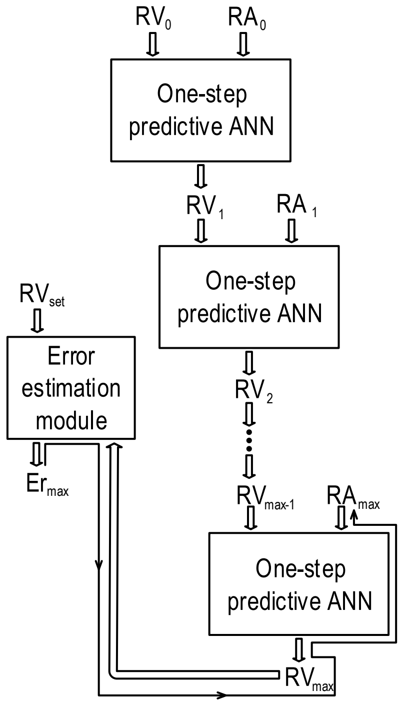


Fig. 4. Schematic diagram of the way of back pass for MAPERS

Therefore, Equation 5 for MAPERS can be rewritten according to the Fig. 4 as:

$$RA'_{max} = RA_{max} - \lambda \cdot \frac{\partial Er_{max}}{\partial RA_{max}} \tag{7}$$

where: RA_{max} – regulator action in step with maximum value of predicted regulation error.

We can also rewrite Equation 6 for MAPERS as:

$$\frac{\partial Er_{max}}{\partial RA_{max}} = \frac{\partial Er_{max}}{\partial RV_{max}} \cdot \frac{\partial RV_{max}}{\partial RA_{max}} \tag{8}$$

where: RV_{max} – predicted regulated value in step with maximum value of predicted regulation error.

To calculate $\frac{\partial Er_{max}}{\partial RV_{max}}$, we need to refer to Equations 3 and 4, mentioned earlier. After calculations, we get the corresponding derivatives:

$$\frac{\partial Er_{max}}{\partial RV_{max}} = \begin{cases} RV_1 - RV_{set}, \; if \; Er_{max} = Er_1 \\ \frac{RV_n - RV_{set}}{(n-1) \cdot d}, \; if \; Er_{max} = Er_n \; and \; n > 1 \end{cases} \tag{9}$$

Now to calculate $\frac{\partial RV_{max}}{\partial RA_{max}}$, we have to dive into the work of the one-step predictive ANN with the algorithm of backpropagation of the effect of the signal. Since this algorithm is involved in the regulator strategy, we are considering that it is worth exploring it in detail.

3.2. Backpropagation of the Effect of the Signal(BES)

Having a trained neural network, we can use it to determine the direction and value in which the output parameter will change when the input is changed. The algorithm of Backpropagation of the Effect of the Signal (hereinafter referred to as BES) resembles a backpropagation training algorithm but works without changing the network settings. Instead, it allows to calculate $\frac{\partial out}{\partial in}$ for the specific status of the inputs and outputs of this network. And that is just what we have left to do to complete Equation 8. Refer to Fig. 1 for an explanation of the BES algorithm, but somewhat simplify the network structure. Let's leave one input (no matter which one) and one output. Hidden neurons are located in two layers, two in each. This network configuration makes it possible to trace possible signal propagation options and can be easily scaled to any number of layers with any number of neurons in each.

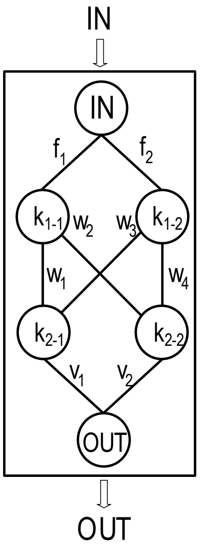


Fig. 5. The scheme of simplified one-step predictive ANN in BES algorithm

First of all, it should be understood that this scheme provides for three types of connections between elements. Namely, input-hidden neuron, hidden neuron-hidden neuron and hidden neuron-output neuron. Therefore, a description of these three types of communication enables a general description of the scheme. For simplicity, each type of communication has its own letter in the scheme: f for input-neuron, w for neuron-neuron, v for neuron-output. Although they correspond to the weights for Equation 1. The BES algorithm starts with a forward pass through the network, in order to establish all the internal states and outputs of its nodes. After that, the back passage begins, where the effect of network input on the output must be calculated. Using a chain rule, find the effect on the output signal of a neuron from the last hidden layer, for example, k_{2-1}:

$$\frac{\partial out}{\partial state_{k_{2-1}}} = \frac{\partial out}{\partial out_{k_{2-1}}} \cdot \frac{\partial out_{k_{2-1}}}{\partial state_{k_{2-1}}} = \mu_{k_{2-1}} \quad (10)$$

where: $\mu_{k_{2-1}}$– effect of the neuron k_{2-1} state on the output signal. As for neuron k_{2-2}, its effect is calculated in the same way. Remembering Equations (1) and (2), we unpack the derivatives.

$$\mu_{k_{2-1}} = out_{k_{2-1}}\left(1 - out_{k_{2-1}}\right) \cdot v_1 \quad (11)$$

As we go further, we come across a neuron-neuron type connection, so we write down an equation describing the effect of neuron k_{1-2} on the output signal (with k_{1-1} everything is the same).

$$\frac{\partial out}{\partial state_{k_{1-2}}} = out_{k_{1-2}}\left(1 - out_{k_{1-2}}\right) \cdot \left(\mu_{k_{2-1}} \cdot w_3 + \mu_{k_{2-2}} \cdot w_4\right) (12)$$

Finally, we pack the influence of the neurons of the first hidden layer and write the equation for the input signal.

$$\frac{\partial out}{\partial in} = \frac{\partial out}{\partial state_{k_{1-1}}} \cdot \frac{\partial state_{k_{1-1}}}{\partial in} + \frac{\partial out}{\partial state_{k_{1-2}}} \cdot \frac{\partial state_{k_{1-2}}}{\partial in} \quad (13)$$

Solving derivatives, we have:

$$\frac{\partial out}{\partial in} = \mu_{k_{1-1}} \cdot f_1 + \mu_{k_{1-2}} \cdot f_2 \quad (14)$$

After solving this problem in a simplified version, we can generalize its solution to allow search of $\frac{\partial out}{\partial in}$ in a network with an arbitrary number of hidden layers and with an arbitrary number of neurons in each (but in all hidden layers identical).

For the input of the neural network (see Fig. 1), the effect equation will be:

$$\frac{\partial out}{\partial in} = \sum_j \left(\mu_{k_{1-j}} \cdot f_j\right) \quad (15)$$

where: $\mu_{k_{1-j}}$ – effect on output of the neuron of the first hidden layer which makes neuron in range from 1 to j, f_j – weight of the connection between k_{1-j} and input.

For hidden layers (except the last):

$$\mu_{k_{(i-1)-j}} = out_{k_{(i-1)-j}}\left(1 - out_{k_{(i-1)-j}}\right) \cdot \sum_j \left(\mu_{k_{i-j}} \cdot v_{i-j}\right) \quad (16)$$

where: $k_{(i-1)-j}$ and k_{i-j} – neurons from adjacent layers (i - the number of hidden layers) with j neurons in each, v_{i-j} – weight of connection between them.

For the last hidden layer:

$$\mu_{k_{i-j}} = out_{k_{i-j}}\left(1 - out_{k_{i-j}}\right) \cdot w_j \quad (17)$$

where: k_{i-j} – neuron from the last hidden layer, w_j – weight of connection between k_{i-j} and output.

So, using the BES algorithm, we are able to complete the calculation of the Equation 8 and, accordingly, Equation 7. Multiple repetition of MAPERS gradually reduces the largest values of the regulation error throughout the planning horizon. As a result, we get an array of optimized predicted regulator actions, including RA_0 (see Fig. 4), which is the action that will be sent for execution by the next control cycle.

In the next section, we will consider the implementation of the previously described multipurpose neuronal network-based regulator on the example of a thermostat.

4. TEMPERATURE CONTROLLER BASED ON SELF-TUNING PREDICTIVE ANN

To demonstrate the operation of the specified multipurpose controller, run it within the task of regulating the temperature of a particular object. Thus, in Fig. 2, we replace the regulated value (RA) with the temperature of the regulated object (T), and at the site of the regulator action (RA), there will be a heater power (P) that is able to heat the control object. Our control object will be virtual, so let's specify the equation according to which it functions.

$$T = T_a + \frac{P}{m} \cdot \left(1 - e^{-b \cdot t}\right) + \left(T_0 - T_a\right) \cdot e^{-b \cdot t} \quad (18)$$

where: T_a – ambient temperature, m and b – parameters that make object temperature inertia (taking into account mass and volume), t – time, T_0 – initial temperature. With each change of heater power (P), time (t) is reset to zero and T_0 equals the current temperature value (T). In this way, the attenuation processes are restarted. Set the temperature range in which the controller should operate from 0 to 50$°C$. In this range, it shall provide a regulation accuracy of up to 0.5$°C$. Since this paper describes the processes in the simulated environment, to simplify the demonstration of the principles, the error of temperature measurement is neglected.

One hidden layer with 30 neurons is sufficient to achieve the specified accuracy of controller.

4.1. Training of one-step predictive artificial neural network

As a training set, we will use the reaction of the control object on full power of the heater with consistent cooling.

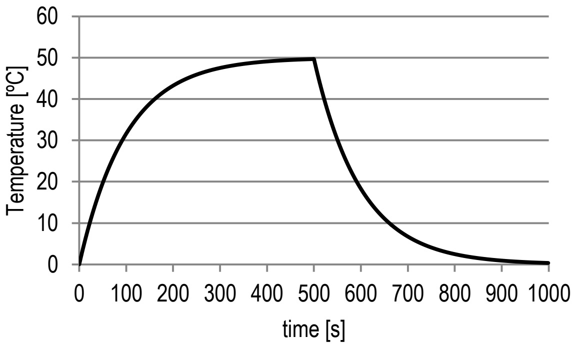


Fig. 6. Training set of temperature for ANN

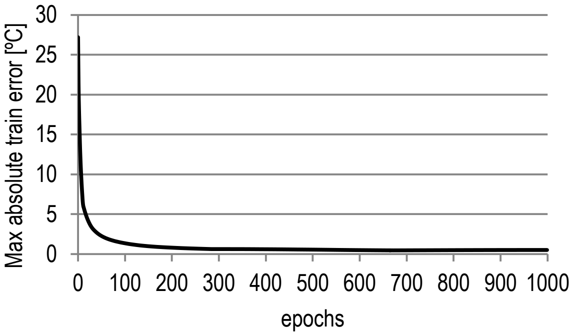


Fig. 7. Training process of ANN

Fig. 6 shows that for half of the training set points (500), the object is heated at P = 100%, after which the heater switches off (P = 0) and the object cools the other half of the set to ambient temperature ($T_a = 0°C$). Regarding the shape of the curve, it can be changed, the slope can be reduced and split into a larger number of time intervals with different heater power. It is not important for training but may be important for maintaining the control object.

As can be seen from Fig. 7, the ANN training process is non-linear. The first 100 epochs reduce the maximum absolute prediction error for the training set to 2 degrees. And after 1000 epochs, the level of the specified error decreases under 0.5 degrees. This means that the network is able to operate with the desired accuracy within the specified range. After training, the regulator is ready to go. Now let's compare it with a PID controller.

4.2. The result of the operation of the self-tuned temperature controller

To evaluate the performance of the previously described controller (with MAPERS), consider it together with a regular PID controller. Fig. 8 shows the response of the system to the change in the desired temperature of the control object (T_{set}).

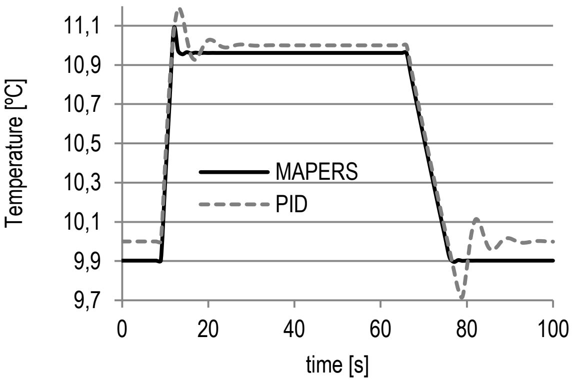


Fig. 8. Comparison of adaptive controller operation with classic PID

At the beginning of measurement, T_{set} is equal to 10 degrees, after which it switches to 11 degrees and after the transition processes returns to 10 degrees. So, we can see the response of regulators to the single step up and down. From what we have seen, we can conclude that the self-tuning controller is capable of operating in a given range with a given accuracy ($0.5°C$), and even significantly outperforming it (the maximum absolute error for steady state during measurements was $0.097°C$). It is worth recalling that in modelling, we ignore measurement errors and focus on the behaviour of regulators under given conditions. Also, the self-tuning controller performed better than the manual tuned PID controller in dynamics, avoiding oscillations and large overshoot. At the same time, the PID controller was more accurate in steady state. It is possible to increase the accuracy of the self-tuning controller by increasing the number of hidden layer neurons and increasing the learning time. But the main advantage of an adaptive controller, as opposed to the manual tuned PID, is automatic tuning. With having a training set that characterizes the behaviour of the control object, we can build a controller with arbitrary complexity and precision without human intervention. The given example of a temperature regulator is the most primitive and clear, therefore, it is necessary to emphasize the possibility for the described adaptive regulator to operate much more complex, multidimensional processes (such as autopilot, industrial control systems, automotive onboard systems, etc.) with the involvement of many input and output signals. The use of a classic PID controller in such processes may not be appropriate and sometimes possible.

5. CONCLUSIONS

The research work aimed to consider a control system that may be an alternative to classic PID controllers in tasks that require automation of the controller setup process. A multipurpose neuronal network-based controller, discussed in this paper, may be that kind of system. We first looked at its structure and the main modules that make it up with a detailed description of the operation of each of these modules and the system as a whole. After a general analysis, we proceeded to solve a specific problem, which was the synthesis of an adaptive thermostat. The synthesized adaptive thermostat showed the set accuracy and stability of control at the simulation level. The success of this task, showed the value of the original model, its strengths, weaknesses and possible ways to improve it. At the same time, it is necessary to continue studying the capabilities of the described regulator on the examples of real control processes. It is worth to study the possible strategies (beyond MAPERS) for optimizing the control action of the regulator.

REFERENCES

1. **Ayomoh M. K. O., Ajala M. T.** (2012), Neural Network Modeling of a Tuned PID Controller, *European Journal of Scientific Research,* 71, 283–297.
2. **Burennikov Y., Kozlov L., Pyliavets V., Piontkevich O.** (2017), Mechatronic Hydraulic Drive with Regulator, Based on Artificial Neuron Network, *IOP Conference Series: Materials Science and Engineering,* 209(1):012071.
3. **Du X., Wang J., Jegatheesan V., Shi G.** (2018), Dissolved Oxygen Control in Activated Sludge Process Using a Neural Network-Based Adaptive PID Algorithm, *Applied Sciences,* 8(2):261, DOI: 10.3390/app8020261.
4. **Elsisi M.** (2019), Design of neural network predictive controller based on imperialist competitive algorithm for automatic voltage regulator, *Neural Computing and Applications,* 31, 5017–5027.

5. **Han G., Fu W., Wang W., Wu Z.** (2017), The Lateral Tracking Control for the Intelligent Veicle Based on Adaptive PID Neural Network, *Sensors,* 17(6):1244, DOI: 10.3390/s17061244.
6. **Heaton J.** (2008), *Introduction to Neural Networks with Java*, Heaton Research Inc., St. Louis.
7. **Hernández-Alvarado R., García-Valdovinos L.G., Salgado-Jiménez T., Gómez-Espinosa A., Fonseca-Navarro F.** (2016), Neural Network-Based Self-Tuning PID Control for Underwater Vehicles, *Sensors,* 16(9), 1429, https://doi.org/10.3390/s16091429.
8. https://stackoverflow.com/questions/10732027/fast-sigmoid-algorithm (08.02.2018)
9. **Liu B., Hussami N., Shrikumar A., Shimko T., Bhate S., Longwell S., Montgomery S., Kundaje A.** (2019), A multi-modal neural network for learning cis and trans regulation of stress response in yeast, arXiv:1908.09426.
10. **Ma H., Lang S., Wellßow W.** (2018) Fallback Solution for a Low-Voltage Regulator Control using Artificial Neural Networks, *CIRED 2018 Ljubljana WS*, http://dx.doi.org/10.34890/413.
11. **MacLean D.** (2019), A convolutional neural network for predicting transcriptional regulators of genes in Arabidopsis transcriptome data reveals classification based on positive regulatory interactions, bioRxiv 618926.
12. **Pirabakaran K., Becerra V.M.** (2002), PID autotuning using neural networks and model reference adaptive control, IFAC Proceedings, 35, 451–456.
13. **Wica M., Witkowsk M., Szumiec A., Ziebura T.** (2019), Weather forecasting system with the use of neural network and backpropagation algorithm, *Proceedings of the International Conference on Data Engineering and Communication Technology,* 2468, 37–41, DOI: 10.1007/978-981-10-1675-2_62.
14. **Zaman M.H.M., Marzuki M.M., Hannan M.A., Hussain A.** (2018), Neural Network Based Prediction of Stable Equivalent Series Resistance in Voltage Regulator Characterization, *Bulletin of Electrical Engineering and Informatics,* 7, 134–142, DOI: 10.11591/eei.v7i1.857.
15. **Zhang Z., Ma C., Zhu R.** (2016), Self-Tuning Fully-Connected PID Neural Network System for Distributed Temperature Sensing and Control of Instrument with Multi-Modules, *Sensors,* 16(10):1709, DOI: 10.3390/s16101709.
16. **Zhao D., Yang T., Ou H., Zhou H.** (2018), Autopilot Design for Unmanned Surface Vehicle based on CNN and ACO, *International Journal of Computers Communications & Control,* 13(3), 429–439, DOI: 10.15837/ijccc.2018.3.3236.