PERIODICA POLYTECHNICA SER. MECH. ENG. VOL. 45, NO. 1, PP. 75–86 (2001)

FUNDAMENTAL CHARACTERISTICS AND DESIGN METHOD FOR NICKEL-TITANIUM SHAPE MEMORY ALLOY

István MIHÁLCZ

Department of Precision Engineering and Optics
Budapest University of Technology and Economics
H–1521 Budapest, Hungary
Phone: +36-20-9460769, Fax: +36-1-4633787
E-mail: mihalcz@hotmail.com

Received: April 5, 2000

Abstract

Shape memory alloys (SMA), because of their unique mechanical characteristics and shape memory effect (SME), have been widely used as force and displacement actuators in many fields [DUERING et al, 1990]. In the industrial applications, it is necessary not only to calculate the mechanical response of the actuator in terms of recovery force or deformation, but also to evaluate its temporal characteristics, i.e., the actuation and reset times. This paper presents the fundamental characteristics of SMA and a complete design model, which requires a close connection between three models: a mechanical model to predict the response of the actuator to a given temperature increment, a thermal model to compute the temperature change in the device, and a continuum-mechanical model to predict the martensite fraction on the SMA. The methodology is applied to a linear wire actuator.

Keywords: shape memory alloys (SMA), shape memory effect (SME), martensitic transformation, NiTi, superelasticity.

1. History of Shape Memory Alloys

The first reported steps towards the discovery of the shape memory effect were taken in 1932 by A. ÖLANDER [R. GILBERTSON, 1994]. He discovered the pseudoelastic behavior of the Au-Cd alloy. GRENINGER and MOORADIAN observed in 1938 the formation and disappearance of a martensitic phase by decreasing and increasing the temperature of a Cu-Zn alloy. The basic phenomenon of the memory effect governed by the thermoelastic behavior of the martensite phase was widely reported a decade later by KURDJUMOV and KHANDROS in 1949 and also by CHANG and READ in 1951. In the early 1960s, BUEHLER and his co-workers at the U.S. Naval Ordnance Laboratory discovered the shape memory effect in an equiatomic alloy of nickel and titanium, which can be considered a breakthrough in the field of shape memory materials [BUEHLER et al., 1967]. This alloy was named Nitinol (***N****ickel–****T****itanium* ***N****aval* ***O****rdnance* ***L****aboratory*). Since that time, intensive investigations have been made to elucidate the mechanics of its basic behavior.

2. General Principles

A NiTi shape memory metal alloy can exist in two different temperature-dependent crystal structures (phases) called *martensite* (lower temperature) and *austenite* (higher temperature or parent phase). Several properties of austenite NiTi and martensite NiTi are notably different. When martensite NiTi is heated, it begins to change into austenite (*Fig. 1*). The temperature at which this phenomenon starts is called *austenite start temperature* (A_s). The temperature at which this phenomenon is complete is called *austenite finish temperature* (A_f). When austenite NiTi is cooled, it begins to change into martensite. The temperature at which this phenomenon starts is called *martensite start temperature* (M_s). The temperature at which martensite is again completely reverted is called *martensite finish temperature* (M_f) [BUEHLER et al., 1967].

The composition and metallurgical treatments have dramatic impacts on the above *transition temperatures*. From the point of view of practical applications, NiTi can have three different forms: martensite, stress-induced martensite (superelastic), and austenite. When the material is in its martensite form, it is soft and ductile and can be easy deformed. *Superelastic* NiTi is highly elastic (rubber-like), while austenitic NiTi is quite strong and hard (similar to titanium) (*Fig.2*). The NiTi material has all these properties, their specific expression depending on the temperature in which it is used. In *Fig. 1* M_d represents the highest temperature to strain-induced martensite and the grey area represents the area of optimal superelasticity.

2.1. Hysteresis

The temperature range for the martensite-to-austenite transformation, i.e. soft-to-hard transition that takes place upon heating is somewhat higher than that for the reverse transformation upon cooling (*Fig.1*). The difference between the transition temperatures upon heating and cooling is called *hysteresis*. Hysteresis is generally defined as the difference between the temperatures at which the material is in 50% transformed to austenite upon heating and in 50% transformed to martensite upon cooling. This difference can be up to 20–30 °C [BUEHLER et al., 1967].

2.2. Thermoelastic Martensitic Transformation

The unique behavior of NiTi is based on the temperature-dependent austenite-to-martensite phase transformation on an atomic scale, which is also called *thermoelastic martensitic transformation*. The thermoelastic martensitic transformation causing the shape recovery is a result of the need of the crystal lattice structure to accommodate to the minimum energy state for a given temperature [OTSUKA et al., 1998].

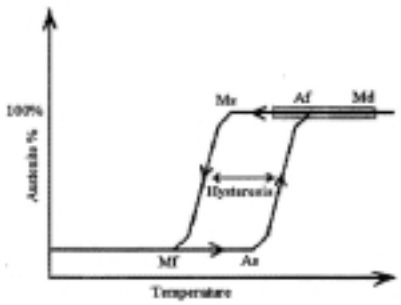

Fig. 1. Martensitic transformation and hysteresis

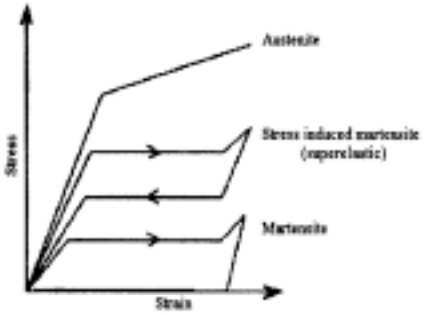

Fig. 2. Stress-strain behavior of different phases of NiTi at constant temperature

In NiTi, the relative symmetries between the two phases lead to a highly ordered transformation, where the displacements of individual atoms can be accurately predicted and eventually lead to a shape change on a macroscopic scale. The crystal structure of martensite is relatively less symmetric compared to that of the parent phase. If a single crystal of the parent phase is cooled below M_f, then martensite variants with a total of 24 crystallographically equivalent habit planes are generally created. There is, however, only one possible parent phase (austenite) orientation, and all martensitic configurations revert to that single defined structure and shape upon heating above A_f. The mechanism by which single martensite variants deform is called *twinning*, and it can be described as a mirror symmetry displacement of atoms across a particular atom plane, the twinning plane [BUEHLER et al., 1967].

While most metals can be deformed by slip or dislocation, NiTi responds to stress by simply changing the orientation of its crystal structure through the movement of twin boundaries.

A NiTi specimen will deform until it consists only of the correspondence variant, which produces maximum strain. However, deformation beyond this will result in classical plastic deformation by slip, which is irrecoverable and therefore has no 'memory effect'. If the deformation is halted midway, the specimen will contain several different correspondence variants. If such a specimen is heated above A_f, a parent phase with an orientation identical to that existing prior to the deformation is created from the correspondence variants in accordance with the lattice correspondences between the original parent phase and each variant (*Fig.3*). The austenite crystal structure is a simple cubic structure, while martensite has a more complex rhombic structure. This phenomenon causes the specimen to revert completely to the shape had before the deformation [GIL et al., 1998].

The above phenomenon is the basis of such special properties as the shape memory effect and superelasticity.

2.3. Shape Memory Effect

NiTi senses a change in ambient temperature and is able to convert its shape to a preprogrammed structure. While NiTi is soft and easily deformable in its lower temperature form (martensite), it resumes its original shape and rigidity when heated to its higher temperature form (austenite) (*Fig.3*). This is called the *one-way shape memory effect*. The ability of shape memory alloys to recover a preset shape upon heating above the transformation temperatures and to return to a certain alternate shape upon cooling is known as the *two-way shape memory effect*. Two-way memory is exceptional. There is also an *all-round shape memory effect*, which is a special case of the two-way shape memory effect [SHIMIZU et al., 1987].

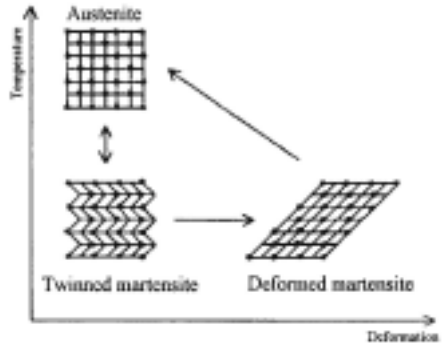

Fig. 3. Transformation from the austenite to the martensite phase and shape memory effect

2.4. Superelasticity

Superelasticity (or *pseudoelasticity*) refers to the ability of NiTi to return to its original shape upon unloading after a substantial deformation. This is based on *stress-induced martensite* formation. The application of an outer stress causes martensite to form at temperatures higher than M_s. The macroscopic deformation is accommodated by the formation of martensite. When the stress is released, the martensite transforms back into austenite and the specimen returns back to its original shape (*Fig.4*). Superelastic NiTi can be strained several times more than ordinary metal alloys without being plastically deformed, which reflects its rubber-like behavior. It is, however, only observed over a specific temperature area. The highest temperature at which martensite can no longer stress induced is called M_d. Above M_d NiTi alloy is deformed like ordinary materials by slipping. Below A_s, the material is martensitic and does not recover. Thus, superelasticity appears in a temperature range from near A_f and up to M_d. The largest ability to recover occurs close to A_f [DUERING et al., 1996].

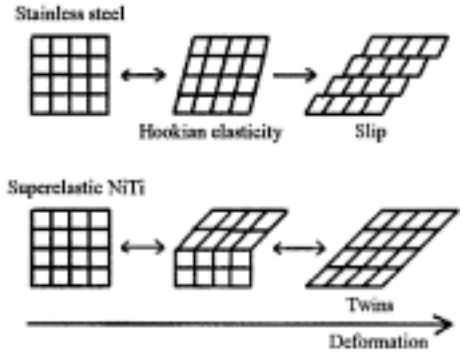

Fig. 4. Schematic presentation of lattice structure changes caused by outer stress in stainless steel or superelastic NiTi alloy

2.5. Limitations of Shape Memory and Superelastic Behavior

About 8% strain can be recovered by unloading and heating. Strain above the limiting value will remain as a permanent plastic deformation. The operating temperature for shape memory devices must not move significantly away from the transformation range, or else the shape memory characteristics may be altered. A NiTi shape memory must be deformed at a temperature below A_s . Moreover, the deformation limit determined by distinctive design (sharp angles, etc.) and the intrinsic strain tolerance of NiTi material must not be disregarded [OTSUKA et al., 1998].

2.6. Mechanical Properties of NiTi

Generally, there are two basic mechanical demands for the material and design of the mechanisms. Service stresses must be safely below the yield strength of the material, and in cyclic loads the service stress must be kept below the fatigue limit.

The mechanical properties of NiTi depend on its phase state at a certain temperature [BUEHLER et al. 1967] (*Fig. 3*). The common mechanical properties of martensitic and austenitic NiTi are presented in *Table 1*. NiTi has an ability to be highly damping and vibration attenuating below A_s . For example, when a martensic NiTi ball is dropped from a constant height, it bounces only slightly over half the height reached by a similar ball dropped above the A_f temperature. The low elastic modulus of NiTi might provide benefits in specific applications. NiTi has unique high fatigue and ductile properties, which are also related to its martensitic transformation. NiTi is a non-magnetic alloy. Electrical resistance and acoustic damping also change when the temperature changes [MIHÁLCZ, 1998].

Table 1. Selected mechanical properties of NiTi

	Austenite	Martensite
Ultimate tensile strength (MPa)	800–1500	103–1100
Tensile yield strength (MPa)	100–800	50–300
Modulus of elasticity (GPa)	70–110	21–69
Elongation at failure (%)	1–20	Up to 60

2.7. Effect of Alloy Composition

It is feasible to vary the critical transition temperatures either by small variations of the Ti/Ni composition or by substituting metallic cobalt for nickel. Lowering of A_f is possible by adding nickel. If nickel is added above 55.6 Wt%, a stable second phase (Ti-Ni3) forms and the NiTi properties are lost. To avoid this problem, the cobalt substitution can be used to lower the transition temperature. The properties of NiTi can also be greatly modified by mechanical working and through heat treatment (time and temperature) [BUEHLER et al., 1967].

2.8. Fabrication

Solid NiTi alloys are manufactured by a double vacuum melting process, to ensure the quality, purity and properties of the material. After the formulation of raw materials, the alloy is vacuum induction melted (1400°C). After the initial melting, the alloy transition temperature must be controlled due to the sensitivity of the transition temperature to small changes in the alloy chemistry. This is followed by vacuum arc remelting to improve the chemistry, homogeneity and structure of the alloy. Double-melted ingots can be hot-worked (800°C) and cold-worked to a wide range of product sizes and shapes.

Porous NiTi can be made by sintering or using self-propagating high temperature synthesis, also called ignition synthesis. The possibility to make composite SMA products (combination with polymers) is under investigation.

2.9. Programming

The use of the one-way shape memory or superelastic property of NiTi for a specific application requires a piece of NiTi to be molded into the required shape. The characteristic heat treatment is then done to set the specimen to its final shape. The heat treatment methods used to set shapes in both the shape memory and the superelastic forms of NiTi are similar. Adequate heat treatment parameters

(temperature and suitable time) are needed to set the shape and the properties of the item [OTSUKA et al., 1998]. They must usually be determined experimentally for the requirements of each desired part. Rapid cooling of some kind is preferred, such as water quenching or rapid air-cooling.

Shape memory effect (SME) training or stress-induced martensite (SIM) training can make the two-way shape memory training procedure. In *SME training*, the specimen is cooled below M_f and bent to the desired shape. It is then heated to a temperature above A_f and allowed freely to take its austenite shape. The procedure is repeated 20–30 times, which completes the training. The sample now assumes its programmed shape upon cooling under M_f and to another shape when heated above A_f.

In *SIM training*, the specimen is bent just above M_s to produce the preferred variants of stress-induced martensite and then cooled below the M_f temperature. Upon subsequent heating above the A_f temperature, the specimen takes its original austenitic shape. This procedure is repeated 20–30 times.

3. The Model

The SMA is used in wire form and is activated electrically (with current or supply source). The design of electrically activated SMA actuator requires the calculation of the mechanical response of the SMA element, as well as an estimation of actuation times.

The mechanical analysis of SMA actuator is based on a planar version of the non-linear phenomenological model. The envelope of the thermo-mechanical behaviour for a given SMA is shown in *Fig. 5*, which is delimited by three loading planes (elastic austenite, transformation loading and elastic martensite). This envelope is built from a set of stress-strain isothermal curves.

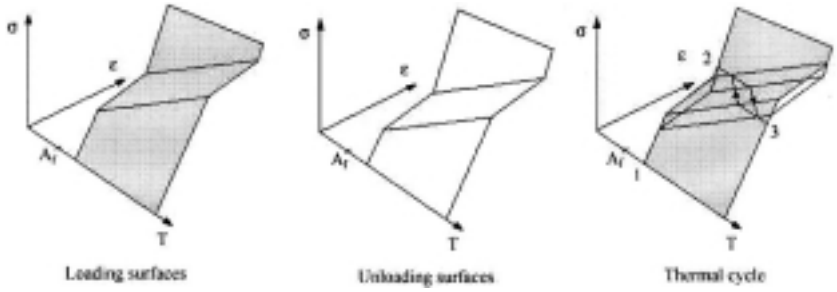

Fig. 5. The linearized material model

An example of such a curve in the superelastic domain is given in *Fig.6*. The linear dependence of the characteristic transformation stresses versus temperature is illustrated in *Fig.7*.

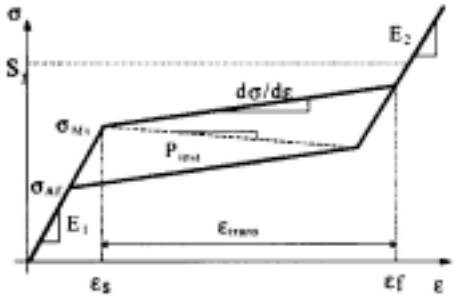

Fig. 6. Bilinear approximation of a superelastic curve

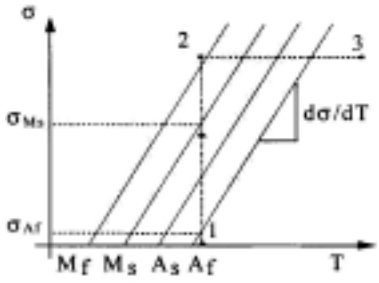

Fig. 7. Approximation of the state diagram

The transformation can be induced thermally or mechanically. For example the sequence starts by creating 100% stress-induced martensite in the wire (trajectory 1–2 in *Figs.* 5 and 7). Then the complete martensite-austenite transformation is realized by heating the wire with electric current. After turning off the electric current, the SMA wire is cooled by natural convection and the reverse transformation occurs (trajectory 3–2 in *Figs.* 5 and 7). In conclusion temperature and stress variation can be imposed in a sequential way in order to model complex super-thermoelastic cycles.

The continuummechanical model is based on the TANAKA's constitutive equation [TANAKA, 1991; BRINSON, 1996]:

$$\sigma - \sigma_0 = E(\xi) \cdot \varepsilon - E(\xi_0) \cdot \varepsilon_0 + \Theta \cdot (T - T_0) + \Omega(\xi) \cdot \xi_S - \Omega(\xi_0) \cdot \xi_{S0}, \quad (1)$$

where σ is the stress in the wire, ε is the strain, T is the temperature, E is the Young's modulus of the material, Θ is the thermal coefficient of expansion, ξ is the martensitic fraction in material, and Ω is the transformation tensor.

The transformation tensor must be expressed as:

$$\Omega(\xi) = -\varepsilon_L \cdot E(\xi). \quad (2)$$

The modulus of the material has been taken in practice to be calculated by a rule of mixture values:

$$E(\xi) = \xi \cdot E_M + (1 - \xi) \cdot E_A, \quad (3)$$

where ε_L is the maximum residual strain for the given SMA material. The elongation ε is a function of σ, ξ and T. The martensite fraction ξ is a memory term, and depends on the current loading, the temperature and on the previous martensite fraction [LIANG, 1993]. Many authors proposed a mathematical expression for

martensite fraction. It was used one proposed by LIANG:

$$\begin{aligned} \xi &= \frac{1-\xi_A}{2}\cos\left[a_M(T-M_f)+b_M\sigma\right]+\frac{1+\xi_A}{2} \qquad \text{for cooling,} \\ \xi &= \frac{\xi_M}{2}\cos[a_A(T-A_s)+b_A\sigma]+1 \qquad \text{for heating,} \end{aligned} \tag{4}$$

where a_A, a_M, b_A, b_M are material constants.

The thermal model describes the current temperature of the wire, which depends on the convection heat coefficient (the heat radiation and the conduction is neglected). The heat balance is:

$$\frac{\mathrm{d}Q}{\mathrm{d}t}+\dot{Q}_{\text{conv}}+\dot{Q}_{\text{rad}}=0. \tag{5}$$

The temperature of the wire was determined with the following equation:

$$T = Fo^{-\frac{r}{Th}}\cdot e^{-Fo(Bi+\varepsilon)}, \tag{6}$$

where Fo is the Fourier number, Th is the Thring number, Bi is the Biot number, r is the emission constant, and ε represents the strain of the wire.

The actuator's time constant is:

$$T_A = \frac{c\cdot\rho\cdot\chi^2}{\lambda(Bi+\varepsilon)}, \tag{7}$$

where c represents the heat capacity, ρ is the medium density of SMA, χ represents the geometrical factor of cross section.

The transformation temperatures (M_s, M_f, A_s, A_f) depend on the stress in the wire. This dependence rule is calculated with the following equation (see *Fig.7*):

$$T_{\text{transformation}} = T_{\sigma=0}+\frac{\Delta\sigma}{k}, \qquad \text{where } k = 9\ldots11\ \text{N/mm}^2\text{K}^0. \tag{8}$$

4. The Control of SMA Actuator

The dynamic behaviour of the actuator differs in every point of the state (martensitic volume fraction ξ, temperature T, mechanical stress σ). An essential failure of the actuator system is the distinctive hysteresis of the martensite volume fraction ξ over T and σ. If the actuator leaves the outer hysteresis loop (HL) it joins into the first order reverse path (RP). If the actuator changes its direction of motion once more, the order of the RP increases. If the actuator is heated, for example, until it joins an RP of lower order, it adopts the order of this path.

A precise examination of the parameter model shows that a differentiation in outer HL, first order RP and higher order RP is sufficient [PRITSCHOW, 1998]. The

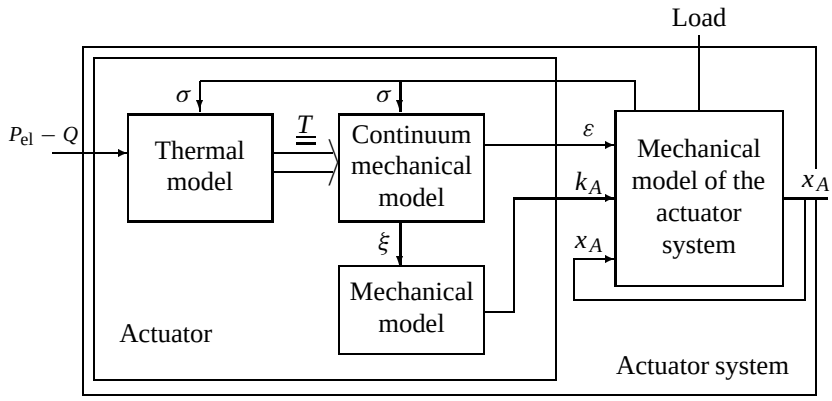

Fig. 8. The complete model diagram

eight identification points are sufficient to find out the order of the RP, on which the actuator is situated and the actuator moves. The actuator state is identified if one of the eight points (see *Fig. 9*) is recognised. The logical condition for the detection of the eight points is shown in *Table 2*.

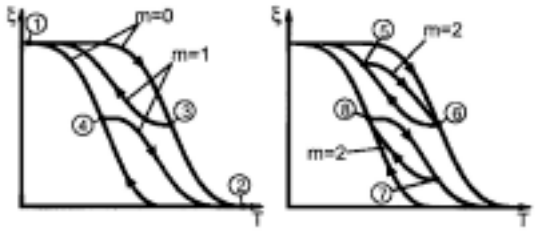

Fig. 9. The identification points [PRITSCHOW, 1998]

where $\dot{\xi}$ represents the elongation or contraction of the actuator (direction of motion), m is the order of the actual RP, $\xi \neq \xi_d$ comparison between detected ξ_d and actual ξ, h and c heating, and cooling process.

The transfer function of the closed loop:

$$W(s) = \frac{W_c \cdot W_p}{1 + W_c \cdot W_p} = \frac{A_c \cdot A_p(T_I \cdot s + 1)}{T_I \cdot s(T_I \cdot s + 1) + A_c \cdot A_p(T_I \cdot s + 1)}. \qquad (9)$$

The error can be evaluated by the integrated square error (ISE). After the minimalization of ISE the optimal control parameters can be found: T_I corresponds to the time constant T_A of the actuator and the amplification factor A_c of the controller is equivalent to the inverse gain parameter of the plant A_p. The signal flow diagram (*Fig.10*) shows the structure of the controlled loop and the adaptive controller.

Table 2. The logical condition of the eight points [PRITSCHOW, 1998]

	$\dot{\xi}$	m	$\xi \neq \xi_d$	Direction	Value of ξ	Description
1	≤ 0	–	–	–	$\xi = 1$	Actuator on HL, heating
2	≥ 0	–	–	–	$\xi = 0$	Actuator on HL, cooling
3	≥ 0	=1	$\xi < \xi_d$	h	–	Actuator leaves HL, cooling
4	≤ 0	=1	$\xi > \xi_d$	c	–	Actuator leaves HL, heating
5	≤ 0	≥ 2	$\xi > \xi_d$	c	–	Actuator leaves RP $(m-1)$, heating
6	< 0	≥ 2	$\xi < \xi_d$	–	–	Actuator on RP m and joins RP $(m-2)$ or HL, heating
7	≥ 0	≥ 2	$\xi < \xi_d$	h	–	Actuator leaves RP $(m-1)$, cooling
8	> 0	≥ 2	$\xi > \xi_d$	–	–	Actuator on RP m and joins RP $(m-2)$ or HL, cooling

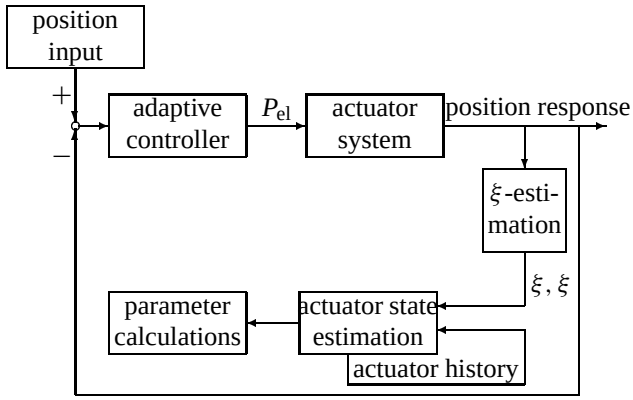

Fig. 10. The signal flow diagram

5. Conclusion

This paper presents a complete design model for linear SMA actuator based on the mixture of three models: thermal, mechanical and continuummechanical model. In the paper an adaptive control system is also presented. For the simulation the visual programming language Labview was used.

The disadvantage of the adaptive controller can be described by the inaccurate approximations (the actuator systems are modelled as an PT_1 element) in the calculation of the optimal control parameters. Applying more exact methods the accuracy will increase.

The actuator state estimation is problematic, because the martensite fraction ξ is not measurable; it is available only by approximation.

Acknowledgements

This work was carried out at the Department of Precision Engineering and Optics, Technical University of Budapest. The author wishes to thank the National Technical Research Fund (OTKA: F 026127) for the financial support.

References

[1] BRINSON, L. C. – HUANG, M. S.: Simplification and Comparison of Shape Memory Alloy Constitutive Models, *Journal of Intelligent Material Systems and Structures*, **7** (1996), pp. 108–114.

[2] BUEHLER, William J. – WANG, Frederick E.: A Summary of Recent Research on the NITINOL Alloys and their Potential Application in Ocean Engineering, *Ocean Engineering*, **1**, (1967) pp. 105–120, Pergamon Press.

[3] DUERING, T. W. – STOCKEL, D. – KEELEY, A.: *Actuator and Work Production Devices, Engineering Aspects of Shape Memory Alloys*, T.W. Duering, K.N. Melton, D. Stockel, and C.M. Wayman (eds), Butterworth-Helnemann, London, (1990) pp. 181–194. ISBN 0-750-61009-3.

[4] GILBERTSON, Roger G.: *Muscle Wires Project Book*, Mondo-Tronics Inc., San Anselmo, 1994, ISBN 1-879896-13-3.

[5] GILL, F. A. – PLANELL, J. A.: In Vitro Thermomechanical Ageing of Ni-Ti Alloys, *Journal of Biomaterial Application*, 1998/12, pp. 237–248.

[6] LIANG, C. – ROGERS, C. A.: Design of Shape Memory Alloy Actuators, *Journal of Mechanical Design*, **114** June 1992, pp. 223–230.

[7] MIHÁLCZ, I. – ILIE, Z. E.: Using Electrical Resistance Variation of Shape Memory Alloys for Transformation Monitoring, 9th *International DAAAM Symposium Intelligent Manufacturing, Automation and Networking*, 22-24 Oct.1998, Cluj-Napoca, Romania, pp. 215–216, ISBN 3-901509-08-9.

[8] OTSUKA, K. – WAYMAN, C. M.: *Mechanism of Shape Memory Effect and Superelasticity, Shape Memory Materials*, Cambridge University Press, Cambridge, (1998) pp. 27–48.

[9] PRITSCHOW, G. – KEHL, G.: Ähnlichkeitsansätze zur Gestaltung von Formgedächtnis-Aktorsystemen, *2. Polisch-Deutscher Workshop 'Werkzeuge der Mechatronik'* 14-15 Mai 1998, Ilmenau, Deutschland, 1998, ISBN 3-00-002556-1.

[10] PRITSCHOW, G. – KEHL, G. – SEIDL, J.: Control and Positioning Behaviour of SMA-actuated Multi-axes Manipulators, *Actuator'98, 6th International Conference on New Actuators*, Bremen, Germany, 17-19 June, 1998., Editor: Hubert Borgmann, Messe Bremen Gmbh, 1998, ISBN 3-933339-00-6.

[11] SHIMIZU, K. – TADAKI, T.: *Shape Memory Effect Mechanism*, Funakubo edition, Shape Memory Alloys, Gordon and Breach Science, New York, (1987) pp. 1–60.

[12] TANAKA, K. et al.: Analysis of Cyclic Effect on Uniaxial Deformation in Shape Memory Alloys under Mechanical and/or Thermal Loads, *JSME, International J.*, Series 1, **34**, (1990), pp. 83–89.