A Quasilinearization Approach for Parameter Identification in a Nonlinear Model of Shape Memory Alloys

Pedro Morin* Rubén D. Spies*

Instituto de Desarrollo Tecnológico Para la Industria Química (INTEC)
Universidad Nacional del Litoral (UNL)
Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)
Güemes 3450 - 3000 Santa Fe
ARGENTINA

Abstract

The nonlinear partial differential equations considered here arise from the conservation laws of linear momentum and energy, and describe structural phase transitions (martensitic transformations) in one-dimensional Shape Memory Alloys (SMA) with non-convex Landau-Ginzburg free energy potentials. This system is formally written as a nonlinear abstract Cauchy problem in an appropriate Hilbert Space. A quasilinearization-based algorithm for parameter identification in this kind of Cauchy problems is proposed. Sufficient conditions for the convergence of the algorithm are derived in terms of the regularity of the solutions with respect to the parameters. Numerical examples are presented in which the algorithm is applied to recover the non-physical parameters describing the free energy potential in SMA, both from exact and noisy data.

Keywords: Parameter Identification, Shape Memory Alloys, Free Energy. Partial Differential Equations, Abstract Cauchy Problems, Quasilinearization.

1 Introduction

In this article the following one-dimensional nonlinear initial-boundary value problem is considered:

$$\rho u_{tt} - \beta\rho u_{xxt} + \gamma u_{xxxx} = f(x,t) + \left(2\alpha_2(\theta-\theta_1)u_x - 4\alpha_4 u_x^3 + 6\alpha_6 u_x^5\right)_x, \quad x \in (0,1),\ 0 \le t \le T \tag{1}$$

$$C_v\theta_t - k\theta_{xx} = g(x,t) + 2\alpha_2\theta u_x u_{xt} + \beta\rho u_{xt}^2, \quad x \in (0,1),\ 0 \le t \le T \tag{2}$$

$$u(x,0) = u_0(x),\ u_t(x,0) = v_0(x),\ \theta(x,0) = \theta_0(x), \quad x \in (0,1) \tag{3}$$

$$u(0,t) = u(1,t) = u_{xx}(0,t) = u_{xx}(1,t) = 0, \quad 0 \le t \le T \tag{4}$$

$$\theta_x(0,t) = \theta_x(1,t) = 0, \quad 0 \le t \le T \tag{5}$$

*The work of the authors was supported in part by CONICET, Consejo Nacional de Investigaciones Científicas y Técnicas of Argentina, UNL, Universidad Nacional del Litoral, Santa Fe, through project CAI+D 94-0016-004-023, and Fundación Antorchas through project A-13434/1 - 000133.

where the subscripts x and t denote partial derivatives with respect to x and t, respectively.

System (1)–(5) arises from the conservation laws governing the thermomechanical processes in one-dimensional Shape Memory Alloys (SMA). These processes are characterized by solid-solid phase transitions (martensitic transformations). Equations (1) and (2) reflect the conservation of linear momentum and energy, respectively. The functions and variables present in system (1)–(5) have the following physical meaning: $u(x,t)$ = transverse displacement, $\theta(x,t)$ = absolute temperature, C_v = specific heat, k = thermal conductivity coefficient, β = viscosity constant, $f(x,t)$ = distributed loads (input), $g(x,t)$ = distributed heat sources (input), T = prescribed final time, α_2, α_4, α_6, θ_1, γ are positive constants —depending on the material being considered— appearing in the free energy potential which is taken in the Landau-Ginzburg form

$$\Psi(\epsilon, \epsilon_x, \theta) = -C_v\theta \ln\left(\frac{\theta}{\theta_2}\right) + C_v\theta + C + \alpha_2(\theta-\theta_1)\epsilon^2 - \alpha_4\epsilon^4 + \alpha_6\epsilon^6 + \frac{\gamma}{2}\epsilon_x^2 \qquad (6)$$

where $\epsilon = u_x$ is the linearized shear strain. The constants θ_1 and θ_2 in (6) are two critical temperatures and C represents a fixed energy reference level. The body is assumed to be a simply supported unit-length beam thermally insulated at both ends. Other relevant boundary conditions can equally be considered.

The PDE's in (1)–(2) are coupled and nonlinear due to the terms coming from the partial derivatives of the free energy. The first equation can be regarded as a nonlinear hyperbolic equation in u while the second is a nonlinear parabolic equation in θ (for a detailed derivation of equations (1)–(2) see [34]).

Although there are several representations for the free energy potential of pseudoelastic materials ([34], [16], [40], [38], [37]) the form (6) seems to be the simplest one which is able to reproduce several phenomena —such as hysteresis, shape memory and superelasticity— observed in real materials under different external thermomechanical actions. For values of θ close to θ_1, Ψ is a nonconvex function of ϵ and the stress-strain laws obtained from (6) are strongly temperature-dependent (see Figure 1). At low temperatures these curves exhibit an elasto-plastic behavior at small loads and a second elastic branch at large loads, which permits the body to withstand forces beyond the plastic yield, after which, subsequent unloading produces residual deformation. In the intermediate temperature range the behavior is superelastic, also called pseudoelastic. Here, a plastic yield is also found. However, loading beyond this plastic yield followed by complete unloading does not lead to residual deformation because of the existence of an intermediate elastic branch which the body reaches by creeping back after the load falls beyond a certain critical value. Finally, in the high temperature range the behavior is almost linearly elastic with higher modulus of elasticity for higher temperatures. Hysteresis loops are observed in the stress-strain curves at low and intermediate temperatures (see [34] and references therein).

Due to their unique characteristics, SMA have already found a broad spectrum of applications such as in the construction of orthodontic and other dental devices [4], heat engines, temperature switches and fuses, pipe coupling devices [14], hybrid composites [32] and several interesting applications in Medicine ([13], [14], [33]).

Since the discovery of NiTinol (a Nikel-Titanium alloy) by Buehler [23] in 1962 several mathematical models were proposed and studied ([1], [2], [3], [15], [16], [22], [25], [26], [27], [39]). Most of this models, however, were static and did not take into account the strong coupling between the mechanical and thermal properties, which is one of the distinguishing features possessed by SMA. It was not until recent years that mathematical models were able to deal with most of the unusual properties of SMA and, at the same time, to allow for the inclusion of boundary and distributed external actions that can be used as control variables ([28], [29], [40], [38], [36], [37], [34]). An extensive account on the recent advances in the mathematical modeling of SMA can be found in [9].

From a practical point of view it would be very important to find the values of all the parameters in (1)–(5) that "best fit" experimental data for a given material. This is called the parameter identification problem (ID problem in the sequel). Among all the constants appearing in (1)–(5), the parameters α_2, α_4, α_6, and θ_1 are non-physical and cannot be determined from laboratory experiments. In this article, a quasilinearization-based algorithm to recover these parameters, both from exact and noisy data, will be developed, its convergence to an optimal parameter will be proved and numerical results will be presented.

2 State-Space Formulation and Preliminaries

In this section, following the approach introduced in [34], the initial–boundary value problem (1)–(5) is formulated as an abstract Cauchy problem in an appropriate Hilbert space.

Define the admissible parameter set $\mathcal{Q} \doteq \{q = (\alpha_2, \alpha_4, \alpha_6, \theta_1) | q \in \mathbb{R}^4_+\}$, and the state space Z as the Hilbert space $H^1_0(0,1) \cap H^2(0,1) \times L^2(0,1) \times L^2(0,1)$ with the inner product

$$\left\langle \begin{pmatrix} u \\ v \\ \theta \end{pmatrix}, \begin{pmatrix} \tilde{u} \\ \tilde{v} \\ \tilde{\theta} \end{pmatrix} \right\rangle_Z \doteq \gamma \int_0^1 u''(x)\overline{\tilde{u}''(x)}\, dx + \rho \int_0^1 v(x)\overline{\tilde{v}(x)}\, dx + \frac{C_v}{k} \int_0^1 \theta(x)\overline{\tilde{\theta}(x)}\, dx.$$

Next, define the operator A on Z by

$$D(A) = \left\{ \begin{pmatrix} u \\ v \\ \theta \end{pmatrix} \in Z \;\middle|\; \begin{array}{l} u \in H^4(0,1),\ u(0) = u(1) = u''(0) = u''(1) = 0 \\ v \in H^1_0(0,1) \cap H^2(0,1) \\ \theta \in H^2(0,1), \theta'(0) = \theta'(1) = 0 \end{array} \right\}$$

and

$$A \begin{pmatrix} u \\ v \\ \theta \end{pmatrix} \doteq \begin{pmatrix} 0 & I & 0 \\ -\frac{\gamma}{\rho} D^4 & \beta D^2 & 0 \\ 0 & 0 & \frac{k}{C_v} D^2 \end{pmatrix} \begin{pmatrix} u \\ v \\ \theta \end{pmatrix}, \quad \text{for} \quad z = \begin{pmatrix} u \\ v \\ \theta \end{pmatrix} \in D(A),$$

where $D^n \doteq \frac{\partial^n}{\partial x^n}$.

The following standing hypothesis will be assumed on the functions $f(x,t)$ and $g(x,t)$.

(H1). *For each fixed $t \geq 0$, the functions $f(x,t)$, $g(x,t)$ are in $L^2(0,1)$ and there exist nonnegative functions $K_f(x)$, $K_g(x) \in L^2(0,1)$ such that*

$$|f(x,t_1) - f(x,t_2)| \leq K_f(x)|t_1 - t_2|, \quad \text{and} \quad |g(x,t_1) - g(x,t_2)| \leq K_g(x)|t_1 - t_2|$$

for all $x \in (0,1)$, $t_1, t_2 \in [0,T]$.

Also, define $z_0(x) = \begin{pmatrix} u_0(x) \\ v_0(x) \\ \theta_0(x) \end{pmatrix}$ and $F(q,t,z) : \mathcal{Q} \times \mathbb{R}^+_0 \times D(Z) \to Z$ by $D(Z) = H^2(0,1) \times H^1(0,1) \times H^1(0,q)$, and $F(q,t,z) = \begin{pmatrix} 0 \\ f_2(q,t,z) \\ f_3(q,t,z) \end{pmatrix}$, where

$$\rho f_2(q,t,z)(x) = f(x,t) + \left(2\alpha_2(\theta - \theta_1)u_x - 4\alpha_4 u_x^3 + 6\alpha_6 u_x^5\right)_x,$$
$$C_v f_3(q,t,z)(x) = g(x,t) + 2\alpha_2 \theta u_x v_t + \beta \rho v_t^2.$$

With this notation, the IBVP (1)–(5) can be formally written as the following nonlinear Cauchy problem in the Hilbert space Z:

$$\frac{d}{dt}z(t) = Az(t) + F(q,t,z), \quad 0 \le t \le T \qquad z(0) = z_0. \tag{7}$$

The following preliminary results can be easily derived from theorems 3.7 and 3.11 in [34] with only slight modifications in order to take into account for the different boundary conditions being considered here. Since the modifications needed are trivial and not important for the goals pursued by this article, details are not given here.

Theorem 1 ([34]). *Let the operator $A : D(A) \subset Z \to Z$ be as previously defined. Then*

(i) *A is dissipative;*

(ii) *The adjoint A^* of A is also dissipative and is given by $D(A^*) = D(A)$, and*

$$A^* \begin{pmatrix} u \\ v \\ \theta \end{pmatrix} \doteq \begin{pmatrix} 0 & -I & 0 \\ \frac{\gamma}{\rho}D^4 & \beta D^2 & 0 \\ 0 & 0 & \frac{k}{C_v}D^2 \end{pmatrix} \begin{pmatrix} u \\ v \\ \theta \end{pmatrix}$$

(iii) *The operator A has pure point spectrum $\sigma_p(A)$ given by*

$$\sigma_p(A) = \left\{\lambda_n^+\right\}_{n=1}^{\infty} \cup \left\{\lambda_n^-\right\}_{n=1}^{\infty} \cup \left\{\alpha_n\right\}_{n=0}^{\infty},$$

where

$$\lambda_n^{+,-} = \sqrt{\mu_n}\left(-r \pm \sqrt{r^2-1}\right), \quad \alpha_n = -\frac{k}{C_v}n^2\pi^2 \quad \text{and} \quad \mu_n = \frac{\gamma n^4 \pi^4}{\rho}, \quad r = \frac{\beta\sqrt{\rho}}{2\sqrt{\gamma}}.$$

(iv) *The operator A generates an analytic semigroup of contractions $T(t)$ on Z.*

Theorem 2 ([34], Local existence of solutions). *Let $q \in \mathcal{Q}$ and let A be as defined above. Then for any initial data $z_0 \in D(A)$ there exists $t_1 = t_1(z_0)$ such that the IVP (7) has a unique classical solution $z(t;q) \in C([0,t_1) : Z) \cap C^1((0,t_1) : Z)$.*

The following result, regarding the smoothness of the solution with respect to the parameter can be found in [12].

Theorem 3 ([12]). *Let A and $F(q,t,z)$ be as defined above. Then the mapping $q \to z(t;q)$ is Fréchet differentiable and its derivative $z_q(t;q)$ is locally Lipschitz continuous as a mapping from $\mathcal{Q}$ into $L^\infty(0,T : Z)$.*

3 Parameter Identification

In this section, the parameter identification problem is stated and a quasilinearization-based algorithm is proposed for the identification of the non-physical parameters appearing in system (1)–(5).

Given $n \in \mathbb{N}$ let Y denote the space $\mathbb{R}^n$ and $\mathcal{C}$ be a bounded linear operator from Z into Y, $\mathcal{C} \in \mathcal{L}(Z,Y)$. The operator $\mathcal{C}$ shall be referred to as the "observation operator". Let $\hat{z}_i \in Y$, be "observations" at times t_i, $i = 1, 2, \ldots, m$ of the process described by the

IVP (7). The *"parameter identification problem"* (ID problem in the sequel) associated to problem (7) and the observations $\{\hat{z}_i\}_{i=1}^m$ is:

$$\begin{aligned} \text{(ID)}: \quad & \text{find } q^* \in \mathcal{Q} \text{ that minimizes the error criterion} \\ & J(q) \doteq \frac{1}{2} \sum_{i=1}^{m} \|\mathcal{C} z(t_i; q) - \hat{z}_i\|_Y^2 . \end{aligned} \tag{8}$$

The following recursive algorithm is proposed.

Step 1: Given an estimate q^k of q^*, approximate $z(t;q)$ by its first order Taylor expansion about q^k, i.e., let $z^{k+1}(t;q) \doteq z(t;q^k) + z_q(t;q^k)\left(q - q^k\right)$.

Step 2: Define the modified error criterion by

$$\begin{aligned} J^k(q) &\doteq \frac{1}{2} \sum_{i=1}^{m} \left\|\mathcal{C} z^{k+1}(t_i; q) - \hat{z}_i\right\|_Y^2 \\ &= \frac{1}{2} \sum_{i=1}^{m} \left\|\mathcal{C}\left[z(t_i; q^k) + z_q(t_i; q^k)\left(q - q^k\right)\right] - \hat{z}_i\right\|_Y^2 . \end{aligned}$$

Step 3: Next, define q^{k+1} to be a minimizer of the modified error criterion $J^k(q)$. In order to find q^{k+1}, differentiate $J^k(q)$, set the result equal to zero and solve for q. Finally, call this solution q^{k+1}, replace k with $k+1$ and repeat Step 1.

Observe that, unless $z_q(t_i; q^k) = 0$, for all $i = 1, 2, \ldots, m$ the functional $J^k(q)$ is strictly convex and therefore, there exists only one solution of $D_q(J^k(q)) = 0$ and this solution is a minimizer. Also, the condition $D_q(J^k(q)) = 0$ is satisfied if and only if

$$\sum_{i=1}^{m} \left\langle \mathcal{C}\left[z_q(t_i; q^k)h\right], \mathcal{C}\left[z_q(t_i; q^k)(q - q^k)\right]\right\rangle_Y = -\sum_{i=1}^{m} \left\langle \mathcal{C}\left[z_q(t_i; q^k)h\right], \mathcal{C} z(t_i; q^k) - \hat{z}_i\right\rangle_Y$$

for every $h \in \mathbb{R}^4$.

Now let $M(t;q)$ and $M(t;q)^*$ denote the matrix associated to the linear transformation $\mathcal{C} z_q(t;q)$ and its transposed, respectively. Then, by defining

$$D(q) \doteq \sum_{i=1}^{m} M(t_i; q)^* \left[M(t_i; q)\right],$$

the algorithm reduces to

$$q^{k+1} = E(q^k) \doteq q^k - \left[D(q^k)\right]^{-1} \sum_{i=1}^{m} M(t_i; q^k)^* \left[\mathcal{C} z(t_i; q^k) - \hat{z}_i\right]$$

whenever $\left[D(q^k)\right]^{-1}$ exists.

4 Convergence of the Quasilinearization Algorithm

In this section the convergence of the algorithm introduced in the previous section is studied. Sufficient conditions for the convergence of the algorithm are presented. The following two preliminary lemmas will be needed.

Lemma 4. *Let $t \in [0,T]$ be fixed and $M(t;q)$, $q \in \mathcal{Q}$, as previously defined. Then the mapping $q \to M(t;q)$ is continuous from $\mathcal{Q} \to \mathcal{L}(\mathbb{R}^4, Y)$. Moreover, for any $q \in \mathcal{Q}$, there exist positive constants η_q and $\mathcal{L}_q$ depending on t such that*

$$\|M(t;q) - M(t;\tilde{q})\| \le \mathcal{L}_q \,|q - \tilde{q}|\,, \qquad \forall \tilde{q} \text{ such that } |\tilde{q} - q| < \eta_q.$$

Proof. Let $t \in [0,T]$ be fixed. By Theorem 3, for all $q \in \mathcal{Q}$ there exist $\eta_q > 0$ and $L_q > 0$ such that $\|z_q(t;q) - z_q(t;\tilde{q})\|_{\mathcal{L}(\mathbb{R}^4,Z)} \le L_q\,|q - \tilde{q}|_{\mathbb{R}^4}$ whenever $|\tilde{q} - q| < \eta_q$. Hence

$$\begin{aligned}
\|M(t;q) - M(t;\tilde{q})\| &= \sup_{h \in \mathbb{R}^4,\, |h|=1} \|[M(t;q) - M(t;\tilde{q})]\, h\|_Y \\
&= \sup_{h \in \mathbb{R}^4,\, |h|=1} \|\mathcal{C} z_q(t;q)h - \mathcal{C} z_q(t;\tilde{q})h\|_Y \\
&\le \|\mathcal{C}\|_{\mathcal{L}(Z,Y)} \sup_{h \in \mathbb{R}^4,\, |h|=1} \left\{ \|z_q(t;q) - z_q(t;\tilde{q})\|_{\mathcal{L}(\mathbb{R}^4,Z)}\, |h| \right\} \\
&\le \|\mathcal{C}\|_{\mathcal{L}(Z,Y)}\, L_q\, |q - \tilde{q}|_{\mathbb{R}^4} \\
&\doteq \mathcal{L}_q\, |q - \tilde{q}|_{\mathbb{R}^4}
\end{aligned}$$

for every $\tilde{q}$ such that $|q - \tilde{q}| < \eta_q$. ■

The following lemma is an immediate consequence of Lemma 4 and its proof is therefore omitted.

Lemma 5. *Let $D(q)$, $q \in \mathcal{Q}$, be as defined above. Then the mapping $q \to D(q)$ is locally Lipschitz continuous from $\mathcal{Q} \to \mathcal{L}(\mathbb{R}^4, \mathbb{R}^4)$.*

Before stating the main results concerning the convergence of the quasilinearization algorithm (QA), it is necessary to introduce the concept of *point of attraction*. Its definition is given below as well as a sufficient condition for an iteration mapping on a Banach space to have a point of attraction.

Definition 6. *Let U be an open subset of a Banach space X and let E be a mapping from U into X. Then, $x^* \in U$ is said to be a* point of attraction *of the iteration $x^{k+1} = E(x^k)$ if there exists an open neighborhood S of x^* such that $S \subset U$ and for any $x^0 \in S$, the iterates $x^k \in U$, for all $k \ge 1$ and $x^k \to x^*$ as $k \to \infty$.*

Lemma 7 (Contraction mapping theorem). *Let U be an open subset of a Banach space X, $E : U \to X$, $x^* \in U$ and suppose there is an open ball $B = B(x^*, \eta) \subset U$ and $\alpha \in (0,1)$ such that*

$$\|E(x) - x^*\| \le \alpha \,\|x - x^*\|\,, \qquad \forall x \in B.$$

Then x^* is a point of attraction of the iteration $x^{k+1} = E(x^k)$.

Proof. Whenever $x^0 \in B$, $\|x^1 - x^*\| = \|E(x^0) - x^*\| \le \alpha\,\|x^0 - x^*\|$, from which $x^1 \in B$. By induction,

$$\left\|x^{k+1} - x^*\right\| = \left\|E(x^k) - x^*\right\| \le \alpha \left\|x^k - x^*\right\| \le \alpha^{k+1} \left\|x^0 - x^*\right\|$$

and $\alpha^{k+1}\,\|x^0 - x^*\| \to 0$ as $k \to \infty$. ■

Theorem 8 (Local convergence of the QA under exact fit-to-data assumption). *Let $M(t;q)$, $D(q)$, and $E(q)$ be as defined in the previous section. Assume that there exist* *an open set $U \subset \mathcal{Q}$ and $q^* \in U$ such that $[D(q^*)]^{-1}$ exists and $J(q^*) = 0$ (exact fit-to-data assumption). Let E be the mapping defined at the end of the previous section. Then, for every $\epsilon > 0$, there exists a constant $\delta > 0$ such that $|q - q^*| < \delta$ implies*

$$|E(q) - q^*| \leq K|q - q^*|^2 + \epsilon|q - q^*|$$

where K is a constant depending only on q^* (not on ϵ). In particular, q^* is a point of attraction of the iteration $q^{k+1} = E(q^k)$.

Proof. By definition

$$E(q) = q - [D(q)]^{-1} \left\{ \sum_{i=1}^{m} M(t_i; q)^* \left(\mathcal{C} z(t_i; q) - \hat{z}_i \right) \right\}$$

whenever $[D(q)]^{-1}$ exists. Hence

$$\begin{aligned}
E(q) - q^* &= q - [D(q)]^{-1} \left\{ \sum_{i=1}^{m} M(t_i; q)^* \left(\mathcal{C} z(t_i; q) - \hat{z}_i \right) \right\} - q^* \\
&= [D(q)]^{-1} \left\{ D(q)\,(q - q^*) - \sum_{i=1}^{m} M(t_i; q)^* \left(\mathcal{C} z(t_i; q) - \hat{z}_i \right) \right\} \\
&= [D(q)]^{-1} \left\{ \sum_{i=1}^{m} M(t_i; q)^* \left[M(t_i; q) - M(t_i; q^*) \right] (q - q^*) \right\} \\
&\quad - [D(q)]^{-1} \left\{ \sum_{i=1}^{m} M(t_i; q)^* \left[\mathcal{C} z(t_i; q) - \mathcal{C} z(t_i; q^*) - M(t_i; q^*)\,(q - q^*) \right] \right\} \\
&\quad - [D(q)]^{-1} \left\{ \sum_{i=1}^{m} M(t_i; q)^* \left[\mathcal{C} z(t_i; q^*) - \hat{z}_i \right] \right\}.
\end{aligned}$$

Since $J(q^*) = 0$, the third term on the right hand side equals zero. Also, since $[D(q^*)]^{-1}$ exists, by continuity there exist positive constants δ_1 and D so that if $|q - q^*| < \delta_1$, then $\left\| [D(q)]^{-1} \right\| \leq D$. From Lemma 4 there exists M such that $\|M(t_i; q)^*\| \leq M$ for $i = 1, 2, \ldots, m$, whenever $|q - q^*| < \delta_1$.

Consequently,

$$\begin{aligned}
|E(q) - q^*| &\leq DM \sum_{i=1}^{m} \left\| \left[M\,(t_i; q) - M\,(t_i; q^*) \right] (q - q^*) \right\| \\
&\quad + DM \sum_{i=1}^{m} \left\| \mathcal{C} z(t_i; q) - \mathcal{C} z(t_i; q^*) - M(t_i; q^*)\,(q - q^*) \right\| \\
&\doteq A + B.
\end{aligned}$$

By Lemma 4, if $|q - q^*| < \eta_{q^*}$, then $A \leq DMm\mathcal{L}_{q^*} \, |q - q^*|^2$. Also, since

$$M(t_i; q^*)\,(q - q^*) = \mathcal{C} z_q(t_i; q^*)\,(q - q^*)\,,$$

from the definition of the Fréchet derivative $z_q(t; q)$, for every $\epsilon > 0$, there exists $\delta_2 = \delta_2(\epsilon, q^*) > 0$ such that $|q - q^*| < \delta_2$ implies

$$\left\| \mathcal{C} z(t_i; q) - \mathcal{C} z(t_i; q^*) - M(t_i; q^*)\,(q - q^*) \right\| \leq \epsilon \, |q - q^*|\,,$$

$i = 1, 2, \ldots, m$.

Summarizing,

$$|E(q) - q^*| \leq DMm \left[\mathcal{L}_{q^*} |q - q^*|^2 + \epsilon |q - q^*|\right]$$

for any q such that $|q - q^*| < \delta^* \doteq \min\{\delta_1, \delta_2, \eta_{q^*}\}$. By Lemma 7, q^* is a point of attraction of the iteration $q^{k+1} = E(q^k)$. ■

It is important to note that in Theorem 8 an exact fit-to-data at the minimizer q^* was assumed. In practice, when working with real parameter identification problems, this is not a realistic assumption due to possible observation, measuring and modeling errors. In the next theorem this exact fit-to-data assumption is weakened.

Theorem 9 (Local convergence of the QA with noisy data). *Let $M(t; q)$, $D(q)$, and $E(q)$ be as before. Assume that there exist an open set $U \subset \mathcal{Q}$ and $q^* \in U$ such that $D(q^*)$ is nonsingular and $q^* = E(q^*)$ (fixed point). Let $D \doteq \sup\left\{\|D(q)\|^{-1} : |q - q^*| \leq \delta_1\right\}$ with δ_1 as in Theorem 8 and $\mathcal{L}$ the smallest constant satisfying*

$$\|M(t_i; q)^* - M(t_i; q^*)^*\| \leq \mathcal{L} |q - q^*|, \qquad \forall\, |q - q^*| < \delta_1, \qquad i = 1, 2, \ldots, m,$$

and suppose

$$\sum_{i=1}^{m} \|\mathcal{C}z(t_i; q^*) - \hat{z}_i\| < \frac{1}{D\mathcal{L}}.$$

Then q^* is a point of attraction of the iteration $q^{k+1} = E\left(q^k\right)$.

Proof. Following the same steps as in the proof of Theorem 8, it follows that

$$|E(q) - q^*| \leq DMm \left[\mathcal{L} |q - q^*|^2 + \epsilon |q - q^*|\right] + \left|[D(q)]^{-1} \sum_{i=1}^{m} M(t_i; q)^* \left[\mathcal{C}z(t_i; q^*) - \hat{z}_i\right]\right|. \tag{9}$$

But,

$$\sum_{i=1}^{m} M(t_i; q^*)^* \left[\mathcal{C}z(t_i; q^*) - \hat{z}_i\right] = 0, \tag{10}$$

since, by assumption, $q^* = E(q^*)$. Combining (9) and (10)

$$\begin{aligned}
|E(q) - q^*| &\leq DMm \left[\mathcal{L} |q - q^*|^2 + \epsilon |q - q^*|\right] \\
&\quad + D \left\| \sum_{i=1}^{m} \left[M(t_i; q)^* - M(t_i; q^*)^*\right] \left[\mathcal{C}z(t_i; q^*) - \hat{z}_i\right] \right\| \\
&\leq DMm \left[\mathcal{L} |q - q^*|^2 + \epsilon |q - q^*|\right] \\
&\quad + D\mathcal{L} |q - q^*| \sum_{i=1}^{m} \|\mathcal{C}z(t_i; q^*) - \hat{z}_i\| \\
&= DMm \left[\mathcal{L} |q - q^*|^2 + \epsilon |q - q^*|\right] + \gamma |q - q^*|
\end{aligned}$$

where $\gamma < 1$ by hypothesis. This concludes the proof. ■

5 Numerical Results

Here, the algorithm proposed in the previous sections is applied to the identification of the non-physical parameters appearing in (1)–(5). Measurements are taken at several points in the space-time domain and are specified in the examples.

In the following examples, the parameter values reported by F. Falk in [15] for the alloy $Au_{23}Cu_{30}Zn_{47}$ are used. These values are: $\alpha_2 = 24$ J cm^{-3} K^{-1}, $\alpha_4 = 1.5 \times 10^5$ J cm^{-3}, $\alpha_6 = 7.5 \times 10^6$ J cm^{-3} K^{-1}, $\theta_1 = 208$ K, $C_v = 2.9$ J cm^{-3} K^{-1}, $k = 1.9$ w cm^{-1} K^{-1}, $\rho = 11.1$ g cm^{-3}. Also the value of γ was chosen to be $\gamma = 10^{-12}$ J cm^{-1} as reported in [17], and β was chosen to be $\beta = 1$ (this choice has no particular physical meaning). Figure 1 shows the stress-strain curves obtained from the potential (6) for these values of the parameters. The doted lines indicate the unstable parts of the curves, while the horizontal lines indicate possible hysteresis loops.

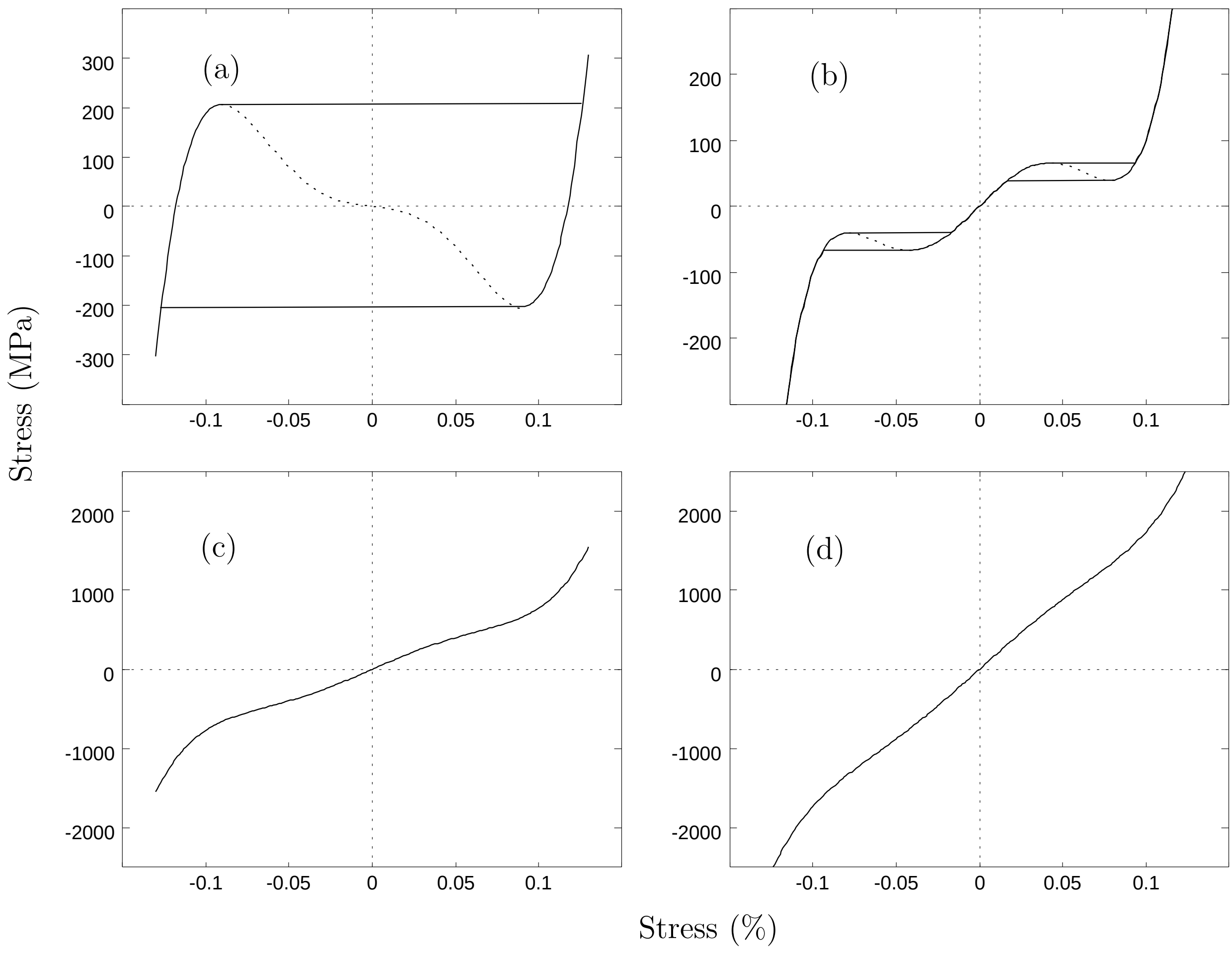


Figure 1: Stress-Strain curves obtained with Ψ as in (6) with α_2, α_4, α_6 and θ_1 as in [15], for different temperatures: (a) $\theta = 200^0$K; (b) $\theta = 260^0$K; (c) $\theta = 400^0$K; (d) $\theta = 600^0$K. Dotted lines indicate unstable parts of the curves. Horizontal lines indicate possible hysteresis loops.

Under certain general conditions, the identifiability of these parameters can be proved (see [24]). The choice of the examples below was done in order to fulfill these conditions.

The parameter to be estimated is $q^* = (\alpha_2,\, \alpha_4,\, \alpha_6,\, \theta_1) = (24,\, 1.5 \times 10^5,\, 7.5 \times 10^6,\, 208)$.
<u>Example 1</u>: <u>*Exact data*</u>.

For this example the initial data is chosen to be $u_0 \equiv 0$, $v_0 \equiv 0$, $\theta_0 \equiv 200$ K, $g(x,t) \equiv 0$,

$$f(x,t) = \begin{cases} 1 \times 10^5, & \text{if } 0.4 \le x \le 0.6, \\ 0, & \text{otherwise} \end{cases}$$

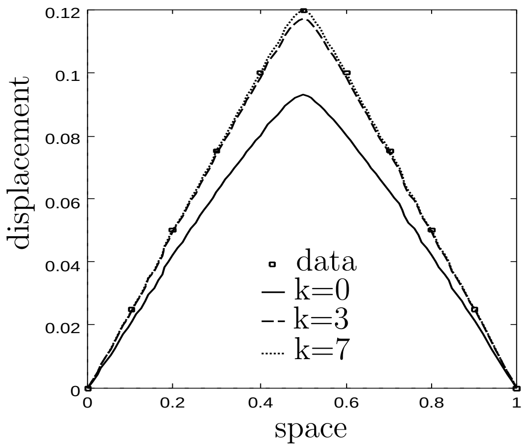
and $T = 0.01$. First, $u(t,x,q^*)$ and $\theta(t,x,q^*)$ are obtained by numerically solving the problem. For this purpose, the spectral method proposed in [19] is used. The observations are then taken to be $\hat{z}_i = \left\{ \begin{pmatrix} u(x_j,t_i;q^*) \\ \theta(x_j,t_i;q^*) \end{pmatrix} \right\}_{j=1}^{9}$, where $t_i = 0.001\,i$, $i = 1,2,\dots,10$, and $x_j = 0.1\,j$, $j = 1,2,\dots,9$. We start with an initial estimate $q^0 = (50,\ 3 \times 10^5,\ 15 \times 10^6,\ 420)$, approximately equal to twice q^*. The results of the iterations produced by the quasilinearization algorithm are shown in Table 1. Figure 2a shows a comparison between $u(x,T;q^*)$ and $u(x,T;q^k)$ while in Figure 2b, $\theta(x,T;q^*)$ and $\theta(x,T;q^k)$ are drawn for different values of k.

Table 1: Values of the parameters and of the error criterion at different iteration steps in Example 1.

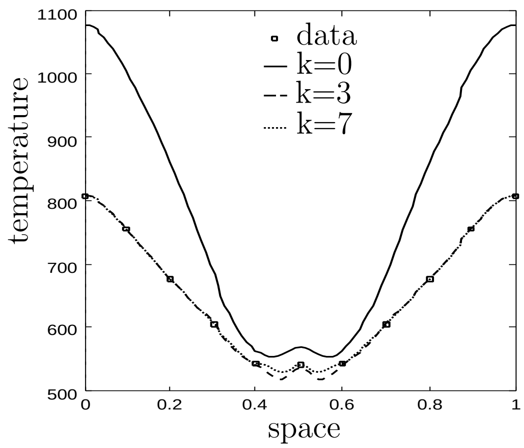
k	α_2	α_4	α_6	θ_1	$J(q^k)$
0	50.0000	300000	1.50000e+07	420.000	1994.6900
1	16.1807	228111	1.40769e+07	459.904	611.1950
2	26.1790	222964	8.71784e+06	33.096	280.8220
3	25.3531	246241	8.83171e+06	126.468	15.3156
4	24.2770	178223	7.87660e+06	181.091	7.1313
5	24.0166	151184	7.51451e+06	206.550	0.6210
6	24.0012	150073	7.50096e+06	207.927	0.0122
7	24.0001	150006	7.50008e+06	207.994	0.0030
8	24.0001	150002	7.50003e+06	207.998	0.0029
9	24.0000	150002	7.50002e+06	207.998	0.0029
10	24.0000	150002	7.50002e+06	207.998	0.0029
11	24.0000	150002	7.50002e+06	207.998	0.0029
12	24.0000	150002	7.50002e+06	207.998	0.0029

Example 2: *Noisy data*.

This example is analogous to Example 1, except that random noise is added to the observation data in order to simulate measuring errors. More precisely, the observations are taken to be $\hat{z}_i = \left\{ \begin{pmatrix} u(x_j,t_i;q^*) + r_{i,j} \\ \theta(x_j,t_i;q^*) + \tilde{r}_{i,j} \end{pmatrix} \right\}_{j=1}^{9}$, where $r_{i,j}$ and $\tilde{r}_{i,j}$ are random numbers uniformly distributed in $(-0.05\overline{u}, 0.05\overline{u})$ and $\left(-0.05\,\overline{\theta}, 0.05\,\overline{\theta}\right)$, respectively, with $\overline{u} = \dfrac{1}{90}\sum_{i=1}^{10}\sum_{j=1}^{9} |u(x_i,t_i;q^*)|$ and $\overline{\theta} = \dfrac{1}{90}\sum_{i=1}^{10}\sum_{j=1}^{9} |\theta(x_i,t_i;q^*)|$. The initial estimate is again $q^0 = (50,\ 3 \times 10^5,\ 15 \times 10^6,\ 420)$. The results of the iterations are shown in Table 2. Figure 3a shows a comparison between $u(x,T;q^*)$ and $u(x,T;q^k)$ while in Figure 3b, $\theta(x,T;q^*)$ and $\theta(x,T;q^k)$ are drawn for different values of k.



(a)



(b)

Figure 2: Displacement (a) and Temperature (b) at $T = 0.01$ for $q = q^k$, $k = 0, 3, 7$.

Table 2: Values of the parameters and of the error criterion at different iteration steps.

k	α_2	α_4	α_6	θ_1	$J(q^k)$
0	50.0000	300000	1.50000e+07	420.000	1987.240
1	16.5263	251533	1.43413e+07	450.975	604.570
2	26.7351	173032	7.92651e+06	77.3584	261.591
3	25.1282	223785	8.54573e+06	148.386	111.619
4	24.2875	176007	7.84280e+06	189.479	111.030
5	24.4436	183683	7.95702e+06	183.663	110.985
6	24.4070	180771	7.91592e+06	186.193	110.977
7	24.4184	181677	7.92857e+06	185.411	110.979
8	24.4151	181408	7.92483e+06	185.645	110.978
9	24.4161	181487	7.92593e+06	185.576	110.979
10	24.4158	181464	7.92560e+06	185.596	110.978
11	24.4159	181471	7.92570e+06	185.590	110.978
12	24.4159	181469	7.92567e+06	185.592	110.978
13	24.4159	181469	7.92568e+06	185.592	110.978

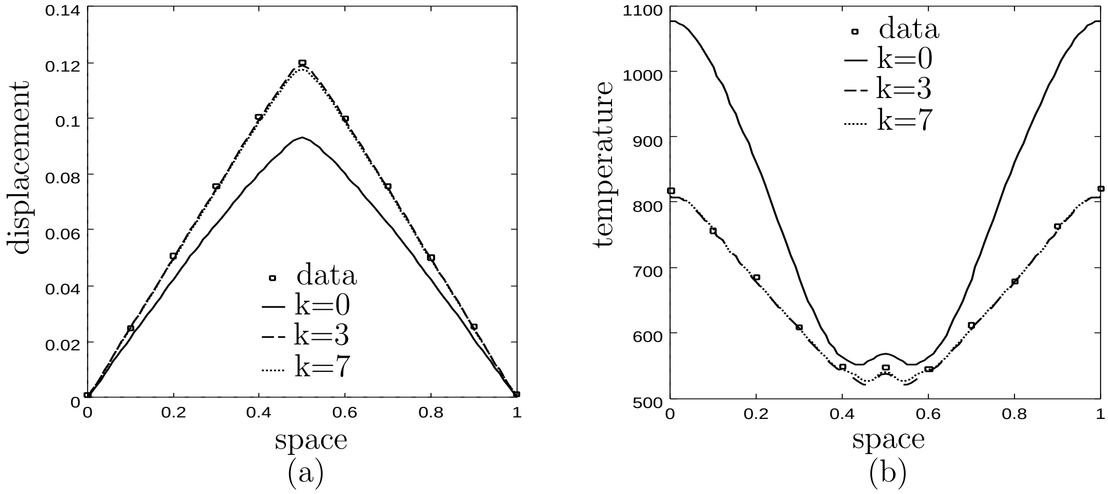


Figure 3: Displacement (a) and Temperature (b) at $T = 0.01$ for $q = q^k$, $k = 0, 3, 7$.

6 Conclusions

In this article, an algorithm for the identification of the parameters α_2, α_4, α_6 and θ_1 in the initial-boundary value problem (1)–(5) has been proposed. Its convergence has been proved for both exact and noisy data. Numerical results have been presented. Although the numerical experiments have been conducted using artificial data, efforts are currently underway to obtain measurements which will also allow to test the algorithm against experimental data.

References

[1] M. Achenbach and I. Muller (1982), A Model for Shape Memory, *J. Physique C4* **43**, No. 12, Suppl., pp. 163–167.

[2] M. Achenbach and I. Muller (1983), Creep and Yield in Martensitic Transformations, *Ingenieur-Archiv.* **53**, pp. 73-83.

[3] M. Achenbach and I. Muller, Shape Memory as a Thermally Activated Process, *in* "Plasticity Today: Modeling Methods and Applications", pp. 515-531, Sawczuk and Bianchi.

[4] G. F. Andreasen, and R. E. Morrow, Laboratory and Clinical Analyses of NiTinol Wire, *Am. J. Orthod.* **73**, No. 2, pp. 142-151.

[5] H. T. Banks, and G. M. Groome, Jr. (1973), Convergence Theorems for Parameter Estimation by Quasilinearization, *J. Math. Anal. Appl.* **42**, pp. 91-109.

[6] R. E. Bellman, and R. E. Kalaba (1965), "Quasilinearization and Nonlinear Boundary Value Problems", American Elsevier, New York.

[7] A. Bensoussan, G. Da Prato, M. Delfour and S. Mitter (1992), "Representation and Control of Infitnite Dimensional Systems", Vol. I y II, Birkäuser.

[8] D. W. Brewer, J. A. Burns, and E. M. Cliff (1993), Parameter Identification for an Abstract Cauchy Problem, *Quart. Math. Anal.* **51**, pp. 1–22.

[9] M. Brokate and J. Sprekels (1996), "Hysteresis and Phase Transitions", Applied Mathematical Sciences, Vol. 121, Springer, New York.

[10] J. Buell, and R. E. Kalaba (1969), Quasilinearization and the Fitting of Nonlinear Models of Drug Metabolism to Experimental Knetic Data, *Math. Biosciences* **5**, pp. 121–132.

[11] J. A. Burns, and E. M. Cliff (1980), An Abstract Quasilinearization Algorithm for Estimating Parameters in Hereditary Systems, *IEEE Trans. Automat. Contr.* **25**, pp.126–129.

[12] J. A. Burns, P. Morin, and R. D. Spies (1997), Parameter Differentiability of the Solution of a Nonlinear Abstract Cauchy Problem, ICAM Report 97-06-01, Interdisciplinary Center for Applied Mathematics, Virginia Polytechnic Institute and State University, Blacksburg, VA.

[13] L. S. Castelman, S. M. Motzkin (1976), F. P. Alicandri and V. L. Bonawit, Biocompatibility of Nitinol Alloy as an Implant Material, *J. Biomedical Materials Res.* **10**, pp. 614–627.

[14] H. Funakubo (Ed.) (1987), "Shape Memory Alloys", Precision Machinery and Robotics, Vol. 1, Translated from the Japanese by J. B. Kennedy, Gordon and Breach Science Publishers.

[15] F. Falk (1980), Model Free Energy, Mechanics and Thermodynamics of Shape Memory Alloys, *Acta Metallurgica* **28**, pp. 1773–1780.

[16] F. Falk (1983), One Dimensional Model of Shape Memory Alloys, *Arch. Mech.* **35**, pp. 63–84.

[17] A. Friedman, and J. Sprekels (1990), Steady-States of Austenitic-Martensitic Domains in the Ginzburg-Landau Theory of Shape Memory Alloys, *Continuum Mech. Thermodin.*, Vol. 2, pp. 199–213.

[18] D. Henry (1989), "Geometric Theory of Semilinear Parabolic Equations", Lecture Notes in Mathematics No. 840, Springer-Verlag.

[19] T. L. Herdman, P. Morin, and R. D. Spies Convergent Spectral Approximation for the Thermomechanical Processes in Shape Memory Alloys, *J. Nonlinear Anal.: Theory, Methods, and Applications*, to appear.

[20] K. H. Hoffmann and S. Zheng (1986), Uniquenes for Nonlinear Coupled Equations Arising from Alloy Mechanism, Preprint 118, Institut Fr Mathematik, Ausburg.

[21] K. H. Hoffmann and S. Zheng (1988), Uniquenes for Structural Phase Transitions in Shape Memory Alloys, *Math. Methods Appl. Sci.* **10**, pp. 145-151.

[22] R. Lohman and I. Muller, A Model for the Qualitative Description of Martensitic Transformations in Memory Alloys, *in* "Phase Transformations" (E. Aifantis and J. Gittus, Eds.), pp. 55–75.

[23] E. C. Malloy (22 June 1990), Nitinol Provides Shape Memory Capabilities, *On the Surf. Mag.*, pp. 1–3.

[24] P. Morin and R. D. Spies (1997), Identifiability of the Landau–Ginzburg Potential in a Mathematical Model of Shape Memory Alloys *J. Math. Anal. Appl.* **212**, No. 1, pp. 292-315.

[25] I. Muller (1979), A Model for a Body with Shape Memory, *Arch. Rational Mechanics Anal.*, **70**, pp. 61–67.

[26] I. Muller and P. Villaggio (1977), A Model for an Elastic-Plastic Body, *Arch. Rational Mechanics Anal.* **65**, No. 1, pp. 25–46.

[27] I. Muller and K. Wilmanski (1980), A Model for Phase Transitions in Pseudoelastic Bodies, *Il Nuovo Cimento* **57b**, No. 2, pp. 283–318.

[28] M. Niezgodka and J. Sprekels (1988), Existence of Solutions for a Mathematical Model of Structural Phase Transitions in Shape Memory Alloys, *Math. Methods Appl. Sci.* **10**, pp. 197–223.

[29] M. Niezgodka and J. Sprekels (1991), Convergent Numerical Approximations of the Thermomechanical Phase Transitions in Shape Memory Alloys, *Numer. Math.* **58**, pp. 759–778.

[30] M. Niezgodka and J. Sprekels (1988), Global Solution to a Model of Structural Phase Transitions in Shape Memory Alloys, *J. Math. Anal. Appl.* **130**, pp. 39–54.

[31] A. Pazy (1983), "Semigroups of Linear Operators and Applications to Partial Differential Equations", corrected 2nd printing, Springer-Verlag, New York/Berlin.

[32] C. A. Rogers, C. Liang and J. Jia (April 1989), Behavior of Shape Memory Alloy Reinforced Composites, Part I: Model Formulations and Control Problems", *in* 30^{th} AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference, Mobile, AL.

[33] M. A. Schmerling, M. A. Wilkov, A. E. Sanders and J. E. Woosley (1975), A Proposed Medical Application of the Shape Memory Effect: A NiTi Harrington Rod for the Treatment of Scoliosis, *in* "Shape Memory Effects in Alloys", Jeff Perkins (ed.), Plenum Press, pp.563–574.

[34] R. D. Spies (1995), A State-Space Approach to a One-Dimensional Mathematical Model for the Dynamics of Phase Transitions in Pseudoelastic Materials, *J. Math. Anal. Appl.* **190**, pp. 58–100.

[35] R. D. Spies (1994), Results on a Mathematical Model of Thermomechanical Phase Transitions in Thape Memory Materials, *J. Smart Materials Structures* **3**, pp. 459–469.

[36] J. Sprekels (1989), Automatic Control of One-Dimensional Thermomechanical Phase Transitions, *in* "Mathematical Models for Phase Change Problems", International Series of Numerical Mathematics, Vol. 88, pp 89-98, Birkhäuser, Basel.

[37] J. Sprekels (1989), Global Existence for Thermomechanical Processes with Nonconvex Free Energies of Ginzburg-Landau Form, *J. Math. Anal. Appl.* **141**, pp. 333–348.

[38] J. Sprekels and S. Zheng (1989), Global Solutions to the Equations of a Ginzburg-Landau Theory for Structural Phase Transitions in Shape Memory Alloys, *Physica D* **39**, pp. 59–76.

[39] K. Wilmansky (1983), Propagation of the Interface in the Stress-Induced Austenite-Martensite Transformation, *Ingenieur-Archiv.* **53**, pp. 291–301.

[40] S. Zheng (Jan. 1989), Global Solutions to the Thermomechanical Equations with Non-convex Landau-Ginzburg Free Energy, *J. Appl. Math. Phys. (ZAMP)*, pp. 111–127.