IDENTIFICATION OF LINEAR ERROR-MODELS WITH PROJECTED DYNAMICAL SYSTEMS

K. Kuhnen[+] and P. Krejci*

[+]Laboratory for Process Automation, Saarland University, Germany

*Mathematical Institute of the Academy of Sciences, Prague, Czech Republic

Abstract. Linear error models are an integral part of several parameter identification methods for feedforward and feedback control systems and lead in connection with the L_2-norm to a convex distance measure which has to be minimised for identification purposes. The parameters are hereby often subject to specific restrictions whose intersections span a convex solution set with non-differentiability points on its boundary. For solving these well conditioned problems on-line the paper formulates the solution of the bounded convex minimisation problem as a stable equilibrium set of a proper system of differential equations. The vector field of the corresponding system of differential equations is based on a projection of the negative gradient of the distance measure. A general drawback of this approach is the discontinuous right-hand side of the differential equation caused by the projection transformation. The consequence are difficulties for the verification of the existence, uniqueness and stability of a solution trajectory. Therefore the first subject of this paper is the derivation of an alternative formulation of the projected dynamical system, which exhibits, in contrast to the original formulation, a continuous right-hand side and thus is accessible to conventional analysis methods. For this purpose the multi-dimensional stop operator is used and the existence, uniqueness and stability properties of the solution trajectories are established. The second part of this paper deals with the numerical integration of the projected dynamical system which is used for an implementation of the identification method on a digital signal processor for example. To demonstrate the performance the application of this on-line identification method to the hysteretic filter synthesis with the modified Prandtl-Ishlinskii approach is presented in the last part of this paper.

1. Introduction

In general the design of control systems bases on a more or less precise model of the control plant. Frequently a model structure for the characteristic of the plant is given with fixed but unknown model parameters. A central step for the design of feedforward or feedback controllers consists in the identification of the unknown parameters of the plant based on measurements of the characteristic [4,7]. The starting point is the definition of an error signal e which describes the deviation of the model characteristic from the measured characteristic in dependence on the unknown model parameters $\boldsymbol{w}$ for every time t. The next step consists in the derivation of a measure V for the distance between the modeled and the measured data from the error signal e by using a well suited cost functional. Proper error model parameters are then determined by the minimisation of this distance measure V. An important task which consists in obtaining a practically well-posed identification problem is to find an error model of the form

$$e(t) = \zeta(t) + \boldsymbol{\Psi}^T(t) \cdot \boldsymbol{w} \tag{1}$$

with linear dependence on the parameters. Here, ζ and $\boldsymbol{\Psi}$ are a scalar and a vector time function which contain the unparameterised characteristic of the plant. Starting from this error model the square of the L_2-norm leads to

$$\begin{aligned} F(\boldsymbol{w}) &= \frac{1}{2}\boldsymbol{w}^T \cdot \int_0^{t_E} \boldsymbol{\Psi}(t)\boldsymbol{\Psi}(t)^T dt \cdot \boldsymbol{w} + \int_0^{t_E} \zeta(t)\boldsymbol{\Psi}(t)^T dt \cdot \boldsymbol{w} + \frac{1}{2}\int_0^{t_E} \zeta(t)^2 dt \\ &= \frac{1}{2}\boldsymbol{w}^T \cdot \boldsymbol{H} \cdot \boldsymbol{w} + \boldsymbol{g}^T \cdot \boldsymbol{w} + f \end{aligned} \tag{2}$$

which is a quadratic distance measure for the identification of the error model parameters $\boldsymbol{w}$ on the measurement time interval $[0,t_E]$. By construction, the matrix $\boldsymbol{H}$ given by

$$\boldsymbol{w}^T \cdot \boldsymbol{H} \cdot \boldsymbol{w} = \int_0^{t_E} (\boldsymbol{\Psi}^T(t) \cdot \boldsymbol{w})^2 dt \geq 0 \quad \forall \boldsymbol{w} \in \Re^n , \tag{3}$$

is always symmetric and positive-semidefinite. From this follows the convexity of the distance measure $F(\boldsymbol{w})$. In many cases the parameter set consists only of a convex subset Z of $\Re^n$ which can for instance be given in terms of a concave function u as

$$Z = \left\{ \boldsymbol{w} \in \Re^n \;\middle|\; u(\boldsymbol{w}) \geq 0 \right\}. \tag{4}$$

A typical example for such an identification problem is the compensator design for memoryless and hysteretic nonlinearities with the so-called Prandtl-Ishlinskii approach. In this approach the conditions for the invertibility of the corresponding nonlinearities are then formulated as linear inequality constraints for the error model parameters [7], and the convex subset Z is therefore a convex polyhedron which results from the intersection of convex halfspaces defined by linear inequality contraints. In this case the optimisation problem to solve is given by

$$\min_{\boldsymbol{w} \in Z} \{F(\boldsymbol{w})\}. \tag{5}$$

Because of the convexity of the distance measure V and the parameter set Z the bounded optimisation problem (5) is convex itself and thus the set of global minima is convex, too [9]. For solving these well conditioned problems before putting the control system into operation the optimisation theory provides multiple powerful algorithms. But these algorithms are not suitable for an optimisation during the operation of the control system, because they require too much computing power. A possibility to avoid these difficulties is to formulate the solution of the bounded quadratic optimisation problem as a stable equilibrium point of a proper system of differential equations. This dynamical system can then be solved through numerical integration very efficiently from time step to time step during the operation of the control system. In the unconstrained case, for instance, the right-hand side of the differential equation is given by the negative gradient of the quadratic target function. The theory of the projected dynamical systems [8] offers a starting point for the formulation of a respective differential equation if the convex constraint Z is present. Then the vector field of the differential equation is also based on the negative gradient of the quadratic target function. However, the right-hand side of the differential equation is obtained in this case from a projection of the negative gradient which ensures that the trajectory of the system under no circumstances leaves the admissible solution set. In contrast to parameter projection methods normally used the in the field of adaptive systems, the applied projection transformation considers also the non-

differentiability points of the boundaries of the convex solution set. As in the case of a convex polyhedron these non-differentiability points often result from the intersection of smooth convex sets. The projection transformation leads to a discontinuous right-hand side of the differential equation. The consequence are difficulties for the verification of the existence, uniqueness and stability of a solution trajectory. The main subject of this paper is the derivation of an alternative formulation of the projected dynamical system, which exhibits, in contrast to the original formulation, a continuous right-hand side and thus is accessible to conventional analysis methods. For this purpose the multi-dimensional stop operator, well-known from the hysteretic systems theory, is used [5,6]. Finally the existence and uniqueness of the solution trajectories are verified and the stability properties of the projected dynamical system are investigated.

2. Geometry of convex sets

We consider the vector space $X = \Re^n$ for some $n \in \aleph$, endowed with a scalar product $\langle \cdot, \cdot \rangle$ and with norm

$$\|\boldsymbol{x}\| = \langle \boldsymbol{x}, \boldsymbol{x} \rangle^{1/2} \tag{6}$$

for $\boldsymbol{x} \in X$, and a fixed closed convex set $Z \subset X$. For $\boldsymbol{x} \in Z$ we define the outward normal cone $N(\boldsymbol{x})$ and the tangential cone $T(\boldsymbol{x})$ by the formula

$$N(\boldsymbol{x}) = \{\boldsymbol{y} \in X \mid \langle \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{z} \rangle \geq 0 \quad \forall \boldsymbol{z} \in Z\} \tag{7}$$

and

$$T(\boldsymbol{x}) = \{\boldsymbol{v} \in X \mid \langle \boldsymbol{y}, \boldsymbol{v} \rangle \leq 0 \quad \forall \boldsymbol{y} \in N(\boldsymbol{x})\}. \tag{8}$$

We will see in the sequel that it is appropriate here to consider a general abstract scalar product in X rather than just the canonical one

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle_{\boldsymbol{I}} = \boldsymbol{x}^T \cdot \boldsymbol{y} = \sum_{i=1}^{n} x_i y_i \,. \tag{9}$$

A suitable choice of the scalar product will enable us in section 5 and 6 to simplify considerably the numerical computations. A geometrical interpretation of the normal cone and the tangential cone is given in Fig. 1. Note that we have

$$\boldsymbol{z} - \boldsymbol{x} \in T(\boldsymbol{x}) \quad \forall \boldsymbol{x}, \boldsymbol{z} \in Z \,. \tag{10}$$

We denote by $\boldsymbol{Q}$ the orthogonal projection of X onto Z, that is,

$$\boldsymbol{Q}(\boldsymbol{r}) = \arg\min_{\boldsymbol{z} \in Z} \{\|\boldsymbol{r} - \boldsymbol{z}\|\} \tag{11}$$

for $\boldsymbol{r} \in X$. We also make systematic use of the dual mapping $\boldsymbol{P}(\boldsymbol{r}) = \boldsymbol{r} - \boldsymbol{Q}(\boldsymbol{r})$. In particular (11) can be equivalently written as a variational inequality

$$\boldsymbol{Q}(\boldsymbol{r}) \in Z, \quad \langle \boldsymbol{P}(\boldsymbol{r}), \boldsymbol{Q}(\boldsymbol{r}) - \boldsymbol{z} \rangle \geq 0 \quad \forall \boldsymbol{r} \in X \quad \forall \boldsymbol{z} \in Z \,. \tag{12}$$

Indeed, this follows from the elementary identity

$$\|\boldsymbol{r} - ((1-\alpha)\boldsymbol{Q}(\boldsymbol{r}) + \alpha \boldsymbol{z})\|^2 = \|\boldsymbol{r} - \boldsymbol{Q}(\boldsymbol{r})\|^2 + \alpha^2 \|\boldsymbol{Q}(\boldsymbol{r}) - \boldsymbol{z}\|^2 + 2\alpha \langle \boldsymbol{P}(\boldsymbol{r}), \boldsymbol{Q}(\boldsymbol{r}) - \boldsymbol{z} \rangle \tag{13}$$

for all $\alpha \in [0,1]$, $\boldsymbol{r} \in X$, and $\boldsymbol{z} \in Z$. Assuming that (12) holds, we obtain from (13) for $\alpha = 1$ that $\|\boldsymbol{r} - \boldsymbol{z}\|^2 \geq \|\boldsymbol{r} - \boldsymbol{Q}(\boldsymbol{r})\|^2$ for all $\boldsymbol{z} \in Z$, hence (11) is verified. Conversely, if (11) holds, then $\|\boldsymbol{r} - ((1-\alpha)\boldsymbol{Q}(\boldsymbol{r}) + \alpha\boldsymbol{z})\| \geq \|\boldsymbol{r} - \boldsymbol{Q}(\boldsymbol{r})\|$ and from (13) it follows that $\alpha\|\boldsymbol{Q}(\boldsymbol{r}) - \boldsymbol{z}\|^2 + 2\langle \boldsymbol{P}(\boldsymbol{r}), \boldsymbol{Q}(\boldsymbol{r}) - \boldsymbol{z}\rangle \geq 0$ for all $\alpha \in]0,1]$. Letting α tend to 0+ we obtain (12).

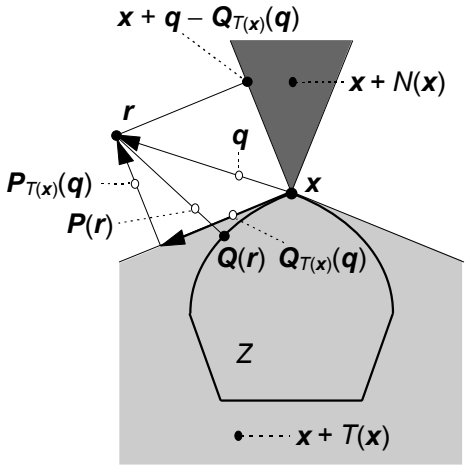


Figure 1: Schematic representation of the geometry of convex sets with arbitrary points $\boldsymbol{x} \in Z$, $\boldsymbol{r} \in Z$, $\boldsymbol{q} = \boldsymbol{r} - \boldsymbol{x}$.

The tangential cone $T(\boldsymbol{x})$ is obviously a convex closed set, we thus can define in the same way the projection $\boldsymbol{Q}_{T(\boldsymbol{x})}$ of X onto $T(\boldsymbol{x})$ for $\boldsymbol{x} \in Z$ analogously to (11) as

$$\boldsymbol{Q}_{T(\boldsymbol{x})}(\boldsymbol{q}) = \arg\min_{\boldsymbol{y} \in T(\boldsymbol{x})} \{\|\boldsymbol{q} - \boldsymbol{y}\|\} \tag{14}$$

for $\boldsymbol{q} = (\boldsymbol{r} - \boldsymbol{x}) \in X$. The relationship between $\boldsymbol{Q}$ and $\boldsymbol{Q}_{T(\boldsymbol{x})}$ will be characterized in the following two Lemmas.

Lemma 2.1 *Let $\boldsymbol{x} \in Z$ and $\boldsymbol{q}, \boldsymbol{v} \in X$ be given. Then*

$$\boldsymbol{v} = \boldsymbol{Q}_{T(\boldsymbol{x})}(\boldsymbol{q}) \Leftrightarrow \boldsymbol{v} \in T(\boldsymbol{x}) \ , \ \langle \boldsymbol{q} - \boldsymbol{v}, \boldsymbol{v} \rangle = 0 \ , \ \boldsymbol{x} = \boldsymbol{Q}(\boldsymbol{x} + \boldsymbol{q} - \boldsymbol{v}) \, . \tag{15}$$

Proof. Assume first that $\boldsymbol{v} = \boldsymbol{Q}_{T(\boldsymbol{x})}(\boldsymbol{q})$. The variational formulation analogous to (12) reads

$$\langle \boldsymbol{q} - \boldsymbol{v}, \boldsymbol{v} - \boldsymbol{s} \rangle \geq 0 \quad \forall \boldsymbol{s} \in T(\boldsymbol{x}) \, . \tag{16}$$

As $T(\boldsymbol{x})$ is a cone, we can choose consecutively $\boldsymbol{s} = \boldsymbol{o}$ and $\boldsymbol{s} = 2\boldsymbol{v} \in T(\boldsymbol{x})$ and obtain $\langle \boldsymbol{q} - \boldsymbol{v}, \boldsymbol{v} \rangle = 0$ and

$$\langle \boldsymbol{q} - \boldsymbol{v}, -\boldsymbol{s} \rangle \geq 0 \quad \forall \boldsymbol{s} \in T(\boldsymbol{x}) \, . \tag{17}$$

By (10) we can put $\boldsymbol{s} = \boldsymbol{z} - \boldsymbol{x}$ for an arbitrary $\boldsymbol{z} \in Z$, hence

$$\langle \boldsymbol{q} - \boldsymbol{v}, \boldsymbol{x} - \boldsymbol{z} \rangle \geq 0 \quad \forall \boldsymbol{z} \in Z \tag{18}$$

which is equivalent to $\boldsymbol{x} = \boldsymbol{Q}(\boldsymbol{x} + \boldsymbol{q} - \boldsymbol{v})$ according to (12). Conversely, if the right-hand side of (15) is satisfied and if $\boldsymbol{s} \in T(\boldsymbol{x})$ is arbitrary, then

$$\langle \boldsymbol{q} - \boldsymbol{v}, \boldsymbol{x} - \boldsymbol{Q}(\boldsymbol{x} + \delta\boldsymbol{s}) \rangle \geq 0 \quad \forall \delta > 0 \, , \tag{19}$$

or equivalently

$$\langle \boldsymbol{q}-\boldsymbol{v}, \frac{1}{\delta}\boldsymbol{P}(\boldsymbol{x}+\delta\boldsymbol{s})-\boldsymbol{s}\rangle \geq 0 \quad \forall \delta > 0\,. \tag{20}$$

The assertion will immediately follow provided we check that

$$\lim_{\delta\to 0+}\frac{1}{\delta}\boldsymbol{P}(\boldsymbol{x}+\delta\boldsymbol{s}) = 0 \quad \forall \boldsymbol{x}\in Z \quad \forall \boldsymbol{s}\in T(\boldsymbol{x})\,. \tag{21}$$

To prove this conjecture, we use (12) for $\boldsymbol{q} = \boldsymbol{x} + \delta\boldsymbol{s}$ and obtain

$$\left\langle \boldsymbol{P}(\boldsymbol{x}+\delta\boldsymbol{s}), \boldsymbol{Q}(\boldsymbol{x}+\delta\boldsymbol{s})-\boldsymbol{z}\right\rangle \geq 0 \quad \forall \boldsymbol{z}\in Z \quad \forall \delta > 0\,. \tag{22}$$

For $\boldsymbol{z} = \boldsymbol{x}$ this yields in particular

$$\left\langle \boldsymbol{P}(\boldsymbol{x}+\delta\boldsymbol{s}), \delta\boldsymbol{s}-\boldsymbol{P}(\boldsymbol{x}+\delta\boldsymbol{s})\right\rangle \geq 0 \quad \forall \delta > 0\,, \tag{23}$$

hence $\|\boldsymbol{P}(\boldsymbol{x}+\delta\boldsymbol{s})\| \leq \delta\,\|\boldsymbol{s}\|$. The system $\{\boldsymbol{P}(\boldsymbol{x}+\delta\boldsymbol{s})/\delta\,;\ \delta > 0\}$ is bounded, we can therefore extract a sequence $\delta_n \to 0+$ and an element $\boldsymbol{y} \in X$ such that

$$\lim_{n\to\infty}\frac{1}{\delta_n}\boldsymbol{P}(\boldsymbol{x}+\delta_n\boldsymbol{s}) = \boldsymbol{y}\,. \tag{24}$$

Dividing (22) for $\delta = \delta_n$ by δ_n and letting n tend to ∞ we obtain

$$\left\langle \boldsymbol{y}, \boldsymbol{x}-\boldsymbol{z}\right\rangle \geq 0 \quad \forall \boldsymbol{z}\in Z\,, \tag{25}$$

hence $\boldsymbol{y}$ belongs to the normal cone $N(\boldsymbol{x})$. We further divide (23) for $\delta = \delta_n$ by δ_n^2 and pass to the limit. This yields

$$\left\langle \boldsymbol{y}, \boldsymbol{s}-\boldsymbol{y}\right\rangle \geq 0\,. \tag{26}$$

By definition of $N(\boldsymbol{x})$ and $T(\boldsymbol{x})$ we have $\langle \boldsymbol{y},\boldsymbol{s}\rangle \leq 0$, hence $\boldsymbol{y} = \boldsymbol{o}$ independently of the sequence $\{\delta_n\}$, and the proof is complete. ■

Lemma 2.2 *For every $\boldsymbol{q} \in X$ and $\boldsymbol{x} \in Z$ we have*

$$\boldsymbol{Q}_{T(x)}(\boldsymbol{q}) = \lim_{\delta\to 0+}\frac{1}{\delta}(\boldsymbol{Q}(\boldsymbol{x}+\delta\boldsymbol{q})-\boldsymbol{x})\,.$$

Proof. Let $\boldsymbol{q} \in X$ *and* $\boldsymbol{x} \in Z$ be given. Similarly as in (22) we have

$$\left\langle \boldsymbol{P}(\boldsymbol{x}+\delta\boldsymbol{q}), \boldsymbol{Q}(\boldsymbol{x}+\delta\boldsymbol{q})-\boldsymbol{x}\right\rangle = \left\langle \boldsymbol{x}+\delta\boldsymbol{q}-\boldsymbol{Q}(\boldsymbol{x}+\delta\boldsymbol{q}), \boldsymbol{Q}(\boldsymbol{x}+\delta\boldsymbol{q})-\boldsymbol{x}\right\rangle \geq 0 \quad \forall \delta > 0\,, \tag{27}$$

hence $\|\boldsymbol{Q}(\boldsymbol{x}+\delta\boldsymbol{q}) - \boldsymbol{x}\| \leq \delta\,\|\boldsymbol{q}\|$ for every $\delta > 0$. There exists again a sequence $\delta_n \to 0$ and an element $\boldsymbol{v} \in X$ such that

$$\lim_{n\to\infty}\frac{1}{\delta_n}(\boldsymbol{Q}(\boldsymbol{x}+\delta_n\boldsymbol{q})-\boldsymbol{x}) = \boldsymbol{v}\,. \tag{28}$$

For every $\boldsymbol{y} \in N(\boldsymbol{x})$ and $\delta > 0$ we have by definition of $N(\boldsymbol{x})$ that $\langle \boldsymbol{y},\boldsymbol{x} - \boldsymbol{Q}(\boldsymbol{x}+\delta\boldsymbol{q})\rangle \geq 0$. Dividing this inequality by $\delta = \delta_n$ by δ_n and passing to the limit we obtain that $\langle \boldsymbol{y},\boldsymbol{v}\rangle \leq 0$ for every $\boldsymbol{y} \in N(\boldsymbol{x})$, hence $\boldsymbol{v} \in T(\boldsymbol{x})$. To check that $\boldsymbol{v} = \boldsymbol{Q}_{T(x)}(\boldsymbol{q})$, we choose an arbitrary $\boldsymbol{s} \in T(\boldsymbol{x})$ and compute

$$\left\langle \boldsymbol{q}-\frac{1}{\delta_n}(\boldsymbol{Q}(\boldsymbol{x}+\delta_n\boldsymbol{q})-\boldsymbol{x}),\frac{1}{\delta_n}(\boldsymbol{Q}(\boldsymbol{x}+\delta_n\boldsymbol{q})-\boldsymbol{x})-\boldsymbol{s}\right\rangle$$
$$=\frac{1}{\delta_n^{\,2}}\left\langle \boldsymbol{P}(\boldsymbol{x}+\delta_n\boldsymbol{q}),\boldsymbol{Q}(\boldsymbol{x}+\delta_n\boldsymbol{q})-(\boldsymbol{x}+\delta_n\boldsymbol{s})\right\rangle =: A_n - B_n, \tag{29}$$

where

$$A_n=\frac{1}{\delta_n^{\,2}}\left\langle \boldsymbol{P}(\boldsymbol{x}+\delta_n\boldsymbol{q}),\boldsymbol{Q}(\boldsymbol{x}+\delta_n\boldsymbol{q})-\boldsymbol{Q}(\boldsymbol{x}+\delta_n\boldsymbol{s})\right\rangle \geq 0\,,$$

$$B_n=\frac{1}{\delta_n^{\,2}}\left\langle \boldsymbol{P}(\boldsymbol{x}+\delta_n\boldsymbol{q}),\boldsymbol{P}(\boldsymbol{x}+\delta_n\boldsymbol{s})\right\rangle$$

with $\lim_{n\to\infty}B_n = 0$ by virtue of (21) and (28). Passing to the limit in (29) we obtain (16), hence $\boldsymbol{v} = \boldsymbol{Q}_{T(\boldsymbol{x})}(\boldsymbol{q})$. The independence of the limit in (27) from the choice of δ_n completes the proof.■

Let us consider now a convex continuously differentiable function $F: X \to \Re$, and let $\boldsymbol{F}': X \to X$ denote its gradient, that is,

$$\left\langle \boldsymbol{F}'(\boldsymbol{x}),\boldsymbol{q}\right\rangle = \lim_{\delta\to 0+}\frac{1}{\delta}(F(\boldsymbol{x}+\delta\boldsymbol{q})-F(\boldsymbol{x})) \quad \text{for } \boldsymbol{x},\, \boldsymbol{q}\in X. \tag{30}$$

Let

$$K=\{\boldsymbol{v}\in Z \mid F(\boldsymbol{v})\leq F(\boldsymbol{z}) \quad \forall \boldsymbol{z}\in Z\} \tag{31}$$

denote the set (possibly empty if Z is unbounded and F is unbounded from below) where F attains its minimum on the set Z. For $\boldsymbol{v}_0,\boldsymbol{v}_1 \in K$ we have indeed $F(\boldsymbol{v}_0) = F(\boldsymbol{v}_1) =: F_{\min}$, and for $\alpha \in [0,1]$ it follows that

$$F_{\min}\leq F(\alpha\boldsymbol{v}_1+(1-\alpha)\boldsymbol{v}_0)\leq \alpha F(\boldsymbol{v}_1)+(1-\alpha)F(\boldsymbol{v}_0)=F_{\min}\,, \tag{32}$$

hence K is convex. We conclude this section with the following easy and classical result.

Lemma 2.3 *Let F, K be as above, and set*

$$\hat{K}=\{\boldsymbol{v}\in Z \mid \left\langle \boldsymbol{F}'(\boldsymbol{v}),\boldsymbol{z}-\boldsymbol{v}\right\rangle \geq 0 \quad \forall \boldsymbol{z}\in Z\}\,,$$

$$K_\infty=\{\boldsymbol{v}\in Z \mid \boldsymbol{o}=\boldsymbol{Q}_{T(\boldsymbol{v})}(-\boldsymbol{F}'(\boldsymbol{v}))\}\,.$$

Then $K_\infty = \hat{K} = K$.

Proof. For $\boldsymbol{v}\in K$ and $\boldsymbol{z}\in Z$ we have $(F(\boldsymbol{v}+\delta(\boldsymbol{z}-\boldsymbol{v})) - F(\boldsymbol{v}))/\delta \geq 0$, and letting δ tend to 0+ we obtain $\langle \boldsymbol{F}'(\boldsymbol{v}),\boldsymbol{z}-\boldsymbol{v}\rangle \geq 0$, hence $K\subset\hat{K}$. To prove the inclusion $\hat{K}\subset K$, we first notice that the convexity of F yields

$$F(\boldsymbol{v}+\delta(\boldsymbol{z}-\boldsymbol{v}))-F(\boldsymbol{v})\leq\delta(F(\boldsymbol{z})-F(\boldsymbol{v})) \quad \forall \boldsymbol{z},\boldsymbol{v}\in Z \quad \forall \delta\in[0,1]\,. \tag{33}$$

For $\delta\to 0+$ this implies that

$$\left\langle \boldsymbol{F}'(\boldsymbol{v}),\boldsymbol{z}-\boldsymbol{v}\right\rangle \leq F(\boldsymbol{z})-F(\boldsymbol{v}) \quad \forall \boldsymbol{z},\boldsymbol{v}\in Z\,. \tag{34}$$

For all $\boldsymbol{v} \in \hat{K}$ we thus have $F(\boldsymbol{z}) - F(\boldsymbol{v}) \geq 0$ for all $\boldsymbol{z} \in Z$, hence $\boldsymbol{v} \in K$. The identity $K_\infty = \hat{K}$ is straightforward. By Lemma 2.1 we have

$$\boldsymbol{o} = \boldsymbol{Q}_{T(\boldsymbol{v})}(-\boldsymbol{F}'(\boldsymbol{v})) \Leftrightarrow \boldsymbol{v} = \boldsymbol{Q}(\boldsymbol{v} - \boldsymbol{F}'(\boldsymbol{v}))$$

which is by virtue of (12) in turn equivalent to the variational inequality

$$\langle -\boldsymbol{F}'(\boldsymbol{v}), \boldsymbol{v} - \boldsymbol{z} \rangle \geq 0 \quad \forall \boldsymbol{z} \in Z$$

and Lemma 2.3 is proved. ■

3. Dynamical systems and the stop operator

The main part of this section is devoted to the projected dynamical system

$$\dot{\boldsymbol{w}}(t) = \boldsymbol{Q}_{T(\boldsymbol{w}(t))}(-\boldsymbol{F}'(\boldsymbol{w}(t))) \quad , \quad \boldsymbol{w}(0) = \boldsymbol{w}_0 \in Z \, , \tag{35}$$

where F is a convex function as at the end of the previous section and where the dot denotes the derivatives with respect to t. The equilibrium point or stationary point of the projected dynamical system (35) is defined as the vector $\boldsymbol{w}_\infty$ which fullfils the equation

$$\boldsymbol{o} = \boldsymbol{Q}_{T(\boldsymbol{w}_\infty)}(-\boldsymbol{F}'(\boldsymbol{w}_\infty)) \, . \tag{36}$$

The equivalence between the set K_∞ of the equilibrium points of the projected dynamical system (35) and the set of set of global minima K of the convex function F is established in Lemma 2.3. This result permits the minimization of $F(\boldsymbol{w})$ with $\boldsymbol{w} \in Z$ by solving the projected dynamical system (35) provided that a solution trajectory of the projected dynamical system (35) exists, is unique and converges asymptotically to the convex set K_∞ of equilibrium points. But the projection transformation $\boldsymbol{Q}_{T(\boldsymbol{w})}$ introduces a discontinuity into the right-hand side of the differential equation (35) and the problem becomes difficult. Instead of trying to extend to our case the complicated methods which are described in [8], we transform the equation (35) into an operator-differential equation with a Lipschitz-continuous right-hand side. For this purpose we work with the space $AC(\Re_+;X)$ of absolutely continuous functions $\boldsymbol{q} : \Re_+ \to X$, where $\Re_+$ denotes the interval $[0,\infty[$. Keeping still fixed the convex closed set $Z \subset X$ from the previous section, we refer to [3,5,6] to recall that for every $\boldsymbol{q} \in AC(\Re_+;X)$ and every initial condition $\boldsymbol{w}_0 \in Z$ there exists a unique solution $\boldsymbol{w} \in AC(\Re_+;X)$ to the variational inequality

$$\boldsymbol{w}(t) \in Z \quad \forall t \geq 0 \ , \ \boldsymbol{w}(0) = \boldsymbol{w}_0 \ , \ \langle \dot{\boldsymbol{q}}(t) - \dot{\boldsymbol{w}}(t), \boldsymbol{w}(t) - \boldsymbol{z} \rangle \geq 0 \ \text{ a.e. } \ \forall \boldsymbol{z} \in Z \, . \tag{37}$$

The solution operator

$$\boldsymbol{S} : Z \times AC(\Re_+;X) \to AC(\Re_+;X) : (\boldsymbol{w}_0, \boldsymbol{q}) \to \boldsymbol{w} \tag{38}$$

is called the *stop* and its analytical properties have been systematically studied e.g. in Chapter 2 of [3]. We will need the following result which in principle goes back to [5].

Lemma 3.1 *Let $\boldsymbol{q},\boldsymbol{w} \in AC(\Re_+;X)$ be given, $\boldsymbol{w}(t) \in Z$ for all $t \geq 0$, and let $\boldsymbol{S} : Z \times AC(\Re_+;X) \to AC(\Re_+;X)$ be the stop. The following two conditions are equivalent.*

(*i*) $\dot{\boldsymbol{w}}(t) = \boldsymbol{Q}_{T(\boldsymbol{w}(t))}(\dot{\boldsymbol{q}}(t))$ a.e.

(*ii*) $\boldsymbol{w} = \boldsymbol{S}[\boldsymbol{w}(0), \boldsymbol{q}]$.

Proof. We use the equivalence statement in Lemma 1.1. If (*i*) holds for some $t > 0$, then (18) reads

$$\langle \dot{\boldsymbol{q}}(t) - \dot{\boldsymbol{w}}(t), \boldsymbol{w}(t) - \boldsymbol{z} \rangle \geq 0 \quad \forall \boldsymbol{z} \in Z\,, \tag{39}$$

hence (*ii*) is fulfilled. Conversely, if (39) holds for some $t > 0$, then putting alternatively $\boldsymbol{z} = \boldsymbol{w}(t+h)$, $\boldsymbol{z} = \boldsymbol{w}(t-h)$ for $h \in]0, t[$, dividing (39) by h, and letting h tend to 0+ we obtain

$$\langle \dot{\boldsymbol{q}}(t) - \dot{\boldsymbol{w}}(t), \dot{\boldsymbol{w}}(t) \rangle = 0\,. \tag{40}$$

Moreover, for $\boldsymbol{y} \in N(\boldsymbol{w}(t))$ we have

$$\langle \boldsymbol{y}, \boldsymbol{w}(t) - \boldsymbol{z} \rangle \geq 0 \quad \forall \boldsymbol{z} \in Z\,, \tag{41}$$

and arguing similarly as above we obtain

$$\langle \boldsymbol{y}, \dot{\boldsymbol{w}}(t) \rangle = 0\,, \tag{42}$$

hence $\dot{\boldsymbol{w}}(t) \in T(\boldsymbol{w}(t))$. From Lemma 2.1 and (39) - (42) we conclude that (*i*) holds. ■

Assume that (35) admits a solution $\boldsymbol{w} \in AC(\Re_+;X)$ such that $\boldsymbol{w}(t) \in Z$ for all $t \geq 0$. For $t \geq 0$ we define an auxiliary fuction $\boldsymbol{q} \in AC(\Re_+;X)$ by the formula

$$\boldsymbol{q}(t) = -\int_0^t \boldsymbol{F}'(\boldsymbol{w}(s))\mathrm{d}s\,. \tag{43}$$

From Lemma 3.1 it follows that $\boldsymbol{w} = \boldsymbol{S}[\boldsymbol{w}_0, \boldsymbol{q}]$, and differentiating (43) we obtain the system

$$\begin{aligned} &(i) \quad \dot{\boldsymbol{q}}(t) + \boldsymbol{F}'(\boldsymbol{w}(t)) = \boldsymbol{o}, \\ &(ii) \quad \boldsymbol{w} = \boldsymbol{S}[\boldsymbol{w}_0, \boldsymbol{q}], \\ &(iii) \quad \boldsymbol{q}(0) = \boldsymbol{o}. \end{aligned} \tag{44}$$

By virtue of Lemma 3.1, problems (35) and (44) are equivalent. We will see that independently of the choice of Z, the stop $\boldsymbol{S}[\boldsymbol{w}_0, \cdot]$ is for every $T > 0$ a Lipschitz continuous mapping from the restriction $AC([0,T];X)$ of $AC(\Re_+;X)$ onto the interval $[0,T]$ into the space $C([0,T];X)$ of continuous functions on $[0,T]$. In order to simplify the presentation, we moreover assume that $\boldsymbol{F}' : X \to X$ is Lipschitz continuous. This enables us to treat the *discontinuous* problem (35) by the standard technique of *continuous* dynamical systems and construct the solution in a standard way by successive approximations.

Proposition 3.2 *Let $Z \subset X$ be an arbitrary non-empty convex closed set and let $F : U \to \Re$ be a convex continuously differentiable function defined in a convex open set $U \subset X$ such that $Z \subset U$. Assume furthermore that $\boldsymbol{F}'$: $X \to X$ is Lipschitz continuous. Then the system (44) admits a unique solution $(\boldsymbol{q},\boldsymbol{w}) \in AC(\Re_+;X) \times AC(\Re_+;X)$ such that $d\boldsymbol{q}/dt$ is continuous and $d\boldsymbol{w}/dt \in L^\infty_{loc}(\Re_+; X)$*

Proof. We fix some final time $T > 0$, and for $t \in [0,T]$ and $n \in \aleph$ define recursively the sequences

$$\boldsymbol{q}^{(0)}(t) = \boldsymbol{o}, \quad \boldsymbol{w}^{(n-1)}(t) = \boldsymbol{S}[\boldsymbol{w}_0, \boldsymbol{q}^{(n-1)}](t), \quad \boldsymbol{q}^{(n)}(t) = \int_0^t -\boldsymbol{F}'(\boldsymbol{w}^{(n-1)}(\tau))\mathrm{d}\tau. \tag{45}$$

From (37) it follows that

$$\left\langle \dot{\boldsymbol{q}}^{(n)}(t)-\dot{\boldsymbol{w}}^{(n)}(t)-\dot{\boldsymbol{q}}^{(m)}(t)+\dot{\boldsymbol{w}}^{(m)}(t),\boldsymbol{w}^{(n)}(t)-\boldsymbol{w}^{(m)}(t)\right\rangle \geq 0 \ \text{ a.e. } \forall n,m\in \aleph\cup\{0\}, \tag{46}$$

hence $\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left\|\boldsymbol{w}^{(n)}(t)-\boldsymbol{w}^{(m)}(t)\right\|^2 \leq \left\|\dot{\boldsymbol{q}}^{(n)}(t)-\dot{\boldsymbol{q}}^{(m)}(t)\right\|\left\|\boldsymbol{w}^{(n)}(t)-\boldsymbol{w}^{(m)}(t)\right\|$ a.e. $\forall n,m\in \aleph\cup\{0\}$ or simply

$$\frac{\mathrm{d}}{\mathrm{d}t}\left\|\boldsymbol{w}^{(n)}(t)-\boldsymbol{w}^{(m)}(t)\right\| \leq \left\|\dot{\boldsymbol{q}}^{(n)}(t)-\dot{\boldsymbol{q}}^{(m)}(t)\right\| \ \text{ a.e. } \forall n,m\in \aleph\cup\{0\} \tag{47}$$

Integrating (47) yields

$$\left\|\boldsymbol{w}^{(n)}(t)-\boldsymbol{w}^{(m)}(t)\right\| \leq \int_0^t \left\|\dot{\boldsymbol{q}}^{(n)}(\tau)-\dot{\boldsymbol{q}}^{(m)}(\tau)\right\|\mathrm{d}\tau \quad \forall t\in[0,T] \quad \forall n,m\in \aleph\cup\{0\}. \tag{48}$$

Let $L > 0$ be the Lipschitz constant of $\boldsymbol{F}'$, that is,

$$\left\|\boldsymbol{F}'(\boldsymbol{x})-\boldsymbol{F}'(\boldsymbol{y})\right\| \leq L\left\|\boldsymbol{x}-\boldsymbol{y}\right\| \quad \forall \boldsymbol{x},\boldsymbol{y}\in U$$

By (45) and (48) we have for every $n \in \aleph$ and $t \in [0,T]$ that

$$\left\|\dot{\boldsymbol{q}}^{(n+1)}(t)-\dot{\boldsymbol{q}}^{(n)}(t)\right\| \leq L\left\|\boldsymbol{w}^{(n)}(t)-\boldsymbol{w}^{(n-1)}(t)\right\| \leq L\int_0^t \left\|\dot{\boldsymbol{q}}^{(n)}(\tau)-\dot{\boldsymbol{q}}^{(n-1)}(\tau)\right\|\mathrm{d}\tau. \tag{49}$$

By induction over n we obtain

$$\int_0^t \left\|\dot{\boldsymbol{q}}^{(n)}(\tau)-\dot{\boldsymbol{q}}^{(n-1)}(\tau)\right\|\mathrm{d}\tau \leq \left\|\boldsymbol{F}'(\boldsymbol{w}_0)\right\|\frac{L^{n-1}t^n}{n!}. \tag{50}$$

The series $\sum_{n=1}^{\infty}\frac{L^{n-1}t^n}{n!}$ is convergent, hence $\{\boldsymbol{q}^{(n)}\}$ is a fundamental sequence in $AC([0,T];X)$. Let $\boldsymbol{q} \in AC([0,T];X)$ denote its limit. Put $\boldsymbol{w} = \boldsymbol{S}[\boldsymbol{w}_0,\boldsymbol{q}]$. Repeating the argument of (48) we obtain

$$\left\|\boldsymbol{w}^{(n)}(t)-\boldsymbol{w}(t)\right\| \leq \int_0^t \left\|\dot{\boldsymbol{q}}^{(n)}(\tau)-\dot{\boldsymbol{q}}(\tau)\right\|\mathrm{d}\tau \quad \forall t\in[0,T] \quad \forall n\in \aleph, \tag{51}$$

hence $\boldsymbol{w}^{(n)}$ converge uniformly to $\boldsymbol{w}$. Passing to the limit in (45) as $n \to \infty$ we check that $(\boldsymbol{q},\boldsymbol{w}) \in AC([0,T];X) \times AC([0,T];X)$ is a solution to (44), hence in particular $\mathrm{d}\boldsymbol{q}/\mathrm{d}t$ is continuous. From (40) it follows that $\|\mathrm{d}\boldsymbol{w}/\mathrm{d}t\| \leq \|\mathrm{d}\boldsymbol{q}/\mathrm{d}t\|$ a.e., hence $\mathrm{d}\boldsymbol{w}/\mathrm{d}t \in L^{\infty}([0,T];X)$. The uniqueness is easy: for two solutions $(\boldsymbol{q}_i,\boldsymbol{w}_i)$, $i = 1,2$, we have

$$\left\langle \dot{\boldsymbol{q}}_1-\dot{\boldsymbol{q}}_2,\boldsymbol{w}_1-\boldsymbol{w}_2\right\rangle+\left\langle \boldsymbol{F}'(\boldsymbol{w}_1)-\boldsymbol{F}'(\boldsymbol{w}_2),\boldsymbol{w}_1-\boldsymbol{w}_2\right\rangle = 0 \ \text{ a.e.},$$

where $\left\langle \dot{\boldsymbol{q}}_1-\dot{\boldsymbol{q}}_2,\boldsymbol{w}_1-\boldsymbol{w}_2\right\rangle \geq \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left\|\boldsymbol{w}_1-\boldsymbol{w}_2\right\|^2$ a.e., $\left\langle \boldsymbol{F}'(\boldsymbol{w}_1)-\boldsymbol{F}'(\boldsymbol{w}_2),\boldsymbol{w}_1-\boldsymbol{w}_2\right\rangle \geq 0$, hence $\boldsymbol{w}_1 = \boldsymbol{w}_2$, $\boldsymbol{q}_1 = \boldsymbol{q}_2$. We thus proved the existence of a unique solution to (44) on each interval $[0,T]$. The solution thus can be extended to $\Re_+$ and the proof is complete. ■

The main result of this section reads as follows.

Proposition 3.3 *Assume that the set K is non-empty, and let* $\boldsymbol{q}$, $\boldsymbol{w} \in AC(\Re_+;X)$ *satisfy (44). Then there exists* $\boldsymbol{w}_\infty \in K$ *such that*

(*i*) $\lim_{t\to\infty} \boldsymbol{w}(t) = \boldsymbol{w}_\infty$.

(*ii*) *There exists a nonincreasing function* $p : \Re_+ \to \Re_+$ *such that* $\|\mathrm{d}\boldsymbol{w}(t)/\mathrm{d}t\| = p(t)$ a.e. and $\lim_{t\to\infty} \sqrt{t}\, p(t) = 0$.

(*iii*) $F(\boldsymbol{w}(t)) \le F(\boldsymbol{w}(s))$ *for all* $t > s \ge 0$, *and* $\int_0^\infty \|\dot{\boldsymbol{w}}(t)\|^2 \,\mathrm{d}t = F(\boldsymbol{w}_0) - F(\boldsymbol{w}_\infty)$.

Proof. For $t > 0$ and $h > 0$ the solution satisfies the identity

$$\dot{\boldsymbol{q}}(t+h) - \dot{\boldsymbol{q}}(t) + \boldsymbol{F}'(\boldsymbol{w}(t+h)) - \boldsymbol{F}'(\boldsymbol{w}(t)) = \boldsymbol{o}. \tag{52}$$

We take the scalar product of its left-hand side with $\boldsymbol{w}(t+h) - \boldsymbol{w}(t)$. From (37) it follows that

$$\langle \dot{\boldsymbol{q}}(t+h) - \dot{\boldsymbol{w}}(t+h), \boldsymbol{w}(t+h) - \boldsymbol{w}(t) \rangle \ge 0 \quad \text{a.e.},$$

$$\langle \dot{\boldsymbol{q}}(t) - \dot{\boldsymbol{w}}(t), \boldsymbol{w}(t) - \boldsymbol{w}(t+h) \rangle \ge 0 \quad \text{a.e.},$$

hence

$$\langle \dot{\boldsymbol{q}}(t+h) - \dot{\boldsymbol{q}}(t), \boldsymbol{w}(t+h) - \boldsymbol{w}(t) \rangle \ge \langle \dot{\boldsymbol{w}}(t+h) - \dot{\boldsymbol{w}}(t), \boldsymbol{w}(t+h) - \boldsymbol{w}(t) \rangle.$$

The convexity of F implies that $\boldsymbol{F}'$ is monotone, that is,

$$\langle \boldsymbol{F}'(\boldsymbol{w}(t+h)) - \boldsymbol{F}'(\boldsymbol{w}(t)), \boldsymbol{w}(t+h) - \boldsymbol{w}(t) \rangle \ge 0.$$

The above considerations lead to the inequality

$$\frac{\mathrm{d}}{\mathrm{d}t} \|\boldsymbol{w}(t+h) - \boldsymbol{w}(t)\|^2 \le 0 \quad \textit{a.e.}\ \forall h > 0, \tag{53}$$

hence

$$\|\boldsymbol{w}(t+h) - \boldsymbol{w}(t)\| \le \|\boldsymbol{w}(s+h) - \boldsymbol{w}(s)\| \quad \forall t > s \ge 0 \quad \forall h > 0. \tag{54}$$

Letting h tend to 0+ we obtain

$$\|\dot{\boldsymbol{w}}(t)\| \le \|\dot{\boldsymbol{w}}(s)\| \quad \text{for all Lebesgue points } t > s > 0 \text{ of } \dot{\boldsymbol{w}}. \tag{55}$$

We now put $p(t) = \operatorname{sup\,ess}\{\|\mathrm{d}\boldsymbol{w}(r)/\mathrm{d}r\|;\ r \in\ ]t,\infty[\}$. Then p is nonincreasing, and the complement of the set A of all $t > 0$ which are Lebesgue points of $\mathrm{d}\boldsymbol{w}/\mathrm{d}t$ and continuity points of p is of measure zero. For all r, $t \in A$, $r > t > 0$, we have $\|\mathrm{d}\boldsymbol{w}(r)/\mathrm{d}r\| \le \|\mathrm{d}\boldsymbol{w}(t)/\mathrm{d}t\|$, hence $p(t) \le \|\mathrm{d}\boldsymbol{w}(t)/\mathrm{d}t\|$. On the other hand, for $t \in A$ and $\varepsilon > 0$ there exists $\delta > 0$ such that $\|\mathrm{d}\boldsymbol{w}(t)/\mathrm{d}t\| \le p(s) \le p(t) + \varepsilon$ for $s \in\ ]t - \delta, t[$. Letting ε tend to 0+ we obtain $\|\mathrm{d}\boldsymbol{w}(t)/\mathrm{d}t\| = p(t)$ for all $t \in A$. We further take the scalar product of (44) (*i*) with $\mathrm{d}\boldsymbol{w}(t)/\mathrm{d}t$. Then (40) yields

$$\|\dot{\boldsymbol{w}}(t)\|^2 + \frac{\mathrm{d}}{\mathrm{d}t} F(\boldsymbol{w}(t)) = 0 \quad \textit{a.e.} \tag{56}$$

hence

$$\int_0^{t^*} \|\dot{\boldsymbol{w}}(t)\|^2 \,\mathrm{d}t + F(\boldsymbol{w}(t^*)) - F(\boldsymbol{w}_0) = 0 \quad \forall t^* > 0, \tag{57}$$

which for $t^* \to \infty$ yields

$$\int_0^\infty p(t)^2 \,\mathrm{d}t \le F(\boldsymbol{w}_0) - F_{\min} . \tag{58}$$

Assume that there exists $c > 0$ and a sequence $0 < t_1 < t_2 < \ldots$ such that $\lim_{n\to\infty} t_n = \infty$ and $p(t_n) \ge c t_n^{-1/2}$ for all $n \in \aleph$. Then

$$\int_0^\infty p(t)^2 \,\mathrm{d}t \ge \sum_{n=1}^{\infty} \int_{t_n}^{t_{n+1}} p(t)^2 \,\mathrm{d}t \ge c^2 \sum_{n=1}^{\infty} (1 - \frac{t_n}{t_{n+1}}) .$$

Set $a_n = 1 - t_n/t_{n+1}$ for $n \in \aleph$. Then $\log(t_{n+1}) = \log(t_n) - \log(1 - a_n)$, hence

$$+\infty = -\sum_{n=1}^{\infty} \log(1-a_n) \le \max\{\frac{\log(1-a_n)}{-a_n}; n \in \aleph\} \sum_{n=1}^{\infty} a_n < +\infty \tag{59}$$

which is a contradiction, hence (*ii*) is verified. The next estimate consists in choosing an arbitrary $\boldsymbol{v} \in K$, and taking the scalar product of (44)(*i*) with $\boldsymbol{w}(t) - \boldsymbol{v}$. Then

$$\left\langle \dot{\boldsymbol{q}}(t), \boldsymbol{w}(t) - \boldsymbol{v} \right\rangle + \left\langle \boldsymbol{F}'(\boldsymbol{w}(t)), \boldsymbol{w}(t) - \boldsymbol{v} \right\rangle = 0 . \tag{60}$$

We have $\left\langle \dot{\boldsymbol{q}}(t), \boldsymbol{w}(t) - \boldsymbol{v} \right\rangle \ge \left\langle \dot{\boldsymbol{w}}(t), \boldsymbol{w}(t) - \boldsymbol{v} \right\rangle$ by (37) and $\left\langle \boldsymbol{F}'(\boldsymbol{w}(t)), \boldsymbol{w}(t) - \boldsymbol{v} \right\rangle \ge F(\boldsymbol{w}(t)) - F(\boldsymbol{v})$ by (34), and from (60) it follows that

$$\left\langle \dot{\boldsymbol{w}}(t), \boldsymbol{w}(t) - \boldsymbol{v} \right\rangle + F(\boldsymbol{w}(t)) - F(\boldsymbol{v}) \le 0 \tag{61}$$

that is

$$\frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \left\| \boldsymbol{w}(t) - \boldsymbol{v} \right\|^2 + F(\boldsymbol{w}(t)) - F(\boldsymbol{v}) \le 0 . \tag{62}$$

We have in particular that

$$\left\| \boldsymbol{w}(t) - \boldsymbol{v} \right\| \le \left\| \boldsymbol{w}(s) - \boldsymbol{v} \right\| \quad \forall t > s \ge 0 . \tag{63}$$

We find a sequence $t_n \to \infty$ and an element $\boldsymbol{w}_\infty \in Z$ such that

$$\dot{\boldsymbol{w}}(t_n) \to 0, \quad \boldsymbol{w}(t_n) \to \boldsymbol{w}_\infty \quad \text{as} \quad n \to \infty . \tag{64}$$

Inequality (61) for $t = t_n$ and $n \to \infty$ yields that $F(\boldsymbol{w}_\infty) \le F(\boldsymbol{v}) = F_{\min}$, hence $\boldsymbol{w}_\infty \in K$. This enables us to finish the proof. Choosing $\boldsymbol{v} = \boldsymbol{w}_\infty$ in (63) we see that the whole trajectory of $\boldsymbol{w}$ converges to $\boldsymbol{w}_\infty$, hence (*i*) holds,while (*iii*) follows from (56) and (57) for $t^* \to \infty$. ■

In the following we restrict ourselves to the quadratic distance measure (2) with a symmetric matrix $\boldsymbol{H}$ which results from a linear error-model (1) and which fulfils the condition

$$\boldsymbol{w}^T \cdot \boldsymbol{H} \cdot \boldsymbol{w} = \int_0^{t_E} (\boldsymbol{\Psi}^T(t) \cdot \boldsymbol{w})^2 \,\mathrm{d}t > 0 \quad \forall \boldsymbol{w} \in \Re^n \setminus \{\boldsymbol{o}\} . \tag{65}$$

In this case $\boldsymbol{H}$ is positive-definite and the distance measure F according to (2) is strictly convex. It is a well-known result from the optimization theory that a strictly convex function subject to convex contraints has only one global minimum in Z [9]. Therefore according to Lemma 2.3 the set K_∞ consists of one single point $\boldsymbol{w}_\infty$, which is according to Proposition 3.3 a globally asymtotically stable equilibrium point of the corresponding projected dynamical sy-

stem (44). If we choose the canonical scalar product according to (9) for the general scalar product an expression for $\boldsymbol{F}'(\boldsymbol{x})$ is given by (30), that is,

$$\left\langle \boldsymbol{F}'(\boldsymbol{x}), \boldsymbol{q} \right\rangle_{\boldsymbol{I}} = \left\langle \boldsymbol{H} \cdot \boldsymbol{x} + \boldsymbol{g}, \boldsymbol{q} \right\rangle_{\boldsymbol{I}}, \tag{66}$$

hence

$$\boldsymbol{F}'(\boldsymbol{x}) = \boldsymbol{H} \cdot \boldsymbol{x} + \boldsymbol{g} \quad \forall \boldsymbol{x} \in \boldsymbol{X}. \tag{67}$$

In this case system (44) can therefore be written in the form

$$\begin{aligned} &(i) \quad \dot{\boldsymbol{q}}(t) = -\boldsymbol{H} \cdot \boldsymbol{w}(t) - \boldsymbol{g}, \\ &(ii) \quad \boldsymbol{w} = \boldsymbol{S}[\boldsymbol{w}_0, \boldsymbol{q}], \\ &(iii) \quad \boldsymbol{q}(0) = \boldsymbol{o}. \end{aligned} \tag{68}$$

Due to the special structure of the gradient in (67) the stability results (*i*) and (*ii*) from Proposition 3.3 can be formulated in a stronger way by

Proposition 3.4. *Assume that the set K is non-empty, and let* $\boldsymbol{q}$, $\boldsymbol{w} \in AC(\Re_+;X)$ *satisfy (44) with a quadratic distance measure F according to (2) which fulfils the condition (65). Then*

$$\left\| \boldsymbol{w}(t) - \boldsymbol{w}_\infty \right\|_2 \leq \mathrm{e}^{-\lambda_{\min}(\boldsymbol{H})t} \left\| \boldsymbol{w}_0 - \boldsymbol{w}_\infty \right\|_2$$

holds with $\lambda_{\min}(\boldsymbol{H}) > 0$ *as the smallest eigenvalue of* $\boldsymbol{H}$.

Proof. The starting point of our considerations is the defining variational inequality (37) of the stop. According to (37) it holds for any $\boldsymbol{w}_1(t) = \boldsymbol{S}[\boldsymbol{w}_{10}, \boldsymbol{q}_1](t) \in Z$ und $\boldsymbol{w}_2(t) = \boldsymbol{S}[\boldsymbol{w}_{20}, \boldsymbol{q}_2](t) \in Z$

$$0 \leq (\dot{\boldsymbol{q}}_1(t) - \dot{\boldsymbol{w}}_1(t))^T \cdot (\boldsymbol{w}_1(t) - \boldsymbol{w}_2(t)) + (\dot{\boldsymbol{q}}_2(t) - \dot{\boldsymbol{w}}_2(t))^T \cdot (\boldsymbol{w}_2(t) - \boldsymbol{w}_1(t)). \tag{69}$$

A reorganization of the right-hand side yields to

$$0 \leq (\dot{\boldsymbol{q}}_1(t) - \dot{\boldsymbol{q}}_2(t) - (\dot{\boldsymbol{w}}_1(t) - \dot{\boldsymbol{w}}_2(t)))^T \cdot (\boldsymbol{w}_1(t) - \boldsymbol{w}_2(t)) \tag{70}$$

and from this follows the inequality

$$(\dot{\boldsymbol{w}}_1(t) - \dot{\boldsymbol{w}}_2(t))^T \cdot (\boldsymbol{w}_1(t) - \boldsymbol{w}_2(t)) \leq (\dot{\boldsymbol{q}}_1(t) - \dot{\boldsymbol{q}}_2(t))^T \cdot (\boldsymbol{w}_1(t) - \boldsymbol{w}_2(t)). \tag{71}$$

Due to

$$\frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \left\| \boldsymbol{w}_1(t) - \boldsymbol{w}_2(t) \right\|^2 = (\dot{\boldsymbol{w}}_1(t) - \dot{\boldsymbol{w}}_2(t))^T \cdot (\boldsymbol{w}_1(t) - \boldsymbol{w}_2(t)) \tag{72}$$

(71) leads to

$$\frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \left\| \boldsymbol{w}_1(t) - \boldsymbol{w}_2(t) \right\|^2 \leq (\dot{\boldsymbol{q}}_1(t) - \dot{\boldsymbol{q}}_2(t))^T \cdot (\boldsymbol{w}_1(t) - \boldsymbol{w}_2(t)). \tag{73}$$

Assume now that both $(\boldsymbol{w}_1, \boldsymbol{q}_1)$ and $(\boldsymbol{w}_2, \boldsymbol{q}_2)$ are solutions of (68) with respective initial states $\boldsymbol{w}_{10}$, $\boldsymbol{w}_{20}$. From (73) and (68)(*i*) it follows that

$$\frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \left\| \boldsymbol{w}_1(t) - \boldsymbol{w}_2(t) \right\|^2 \leq -(\boldsymbol{w}_1(t) - \boldsymbol{w}_2(t))^T \cdot \boldsymbol{H} \cdot (\boldsymbol{w}_1(t) - \boldsymbol{w}_2(t)). \tag{74}$$

Due to

$$(\boldsymbol{w}_1(t)-\boldsymbol{w}_2(t))^T\cdot\boldsymbol{H}\cdot(\boldsymbol{w}_1(t)-\boldsymbol{w}_2(t))\geq\lambda_{\min}(\boldsymbol{H})(\boldsymbol{w}_1(t)-\boldsymbol{w}_2(t))^T\cdot(\boldsymbol{w}_1(t)-\boldsymbol{w}_2(t))$$
$$=\lambda_{\min}(\boldsymbol{H})\left\|\boldsymbol{w}_1(t)-\boldsymbol{w}_2(t)\right\|^2 \tag{75}$$

the right-hand side of (74) can be estimated by

$$\frac{\mathrm{d}}{\mathrm{d}t}\left\|\boldsymbol{w}_1(t)-\boldsymbol{w}_2(t)\right\|^2\leq-2\lambda_{\min}(\boldsymbol{H})\left\|\boldsymbol{w}_1(t)-\boldsymbol{w}_2(t)\right\|^2. \tag{76}$$

Let now $(\boldsymbol{w},\boldsymbol{q})$ be an arbitrary solution of (68) and let $\boldsymbol{w}_\infty$ be the equilibrium point. From (37) and Lemma 2.3 it follows that $\boldsymbol{w}_2(t)=\boldsymbol{w}_\infty$, $\boldsymbol{q}_2(t)=-t(\boldsymbol{H}\cdot\boldsymbol{w}_\infty+\boldsymbol{g})$ satisfies (68) with $\boldsymbol{w}_{20}=\boldsymbol{w}_\infty$. Using (76) with $(\boldsymbol{w}_1,\boldsymbol{q}_1)=(\boldsymbol{w},\boldsymbol{q})$ we obtain the inequality

$$\frac{\mathrm{d}}{\mathrm{d}t}\left\|\boldsymbol{w}(t)-\boldsymbol{w}_\infty\right\|^2\leq-2\lambda_{\min}(\boldsymbol{H})\left\|\boldsymbol{w}(t)-\boldsymbol{w}_\infty\right\|^2 \tag{77}$$

and the time evolution of the Euclidean distance between the equilibrium point $\boldsymbol{w}_\infty$ and $\boldsymbol{w}(t)$ can be estimated by

$$\left\|\boldsymbol{w}(t)-\boldsymbol{w}_\infty\right\|\leq\mathrm{e}^{-\lambda_{\min}(\boldsymbol{H})t}\left\|\boldsymbol{w}_0-\boldsymbol{w}_\infty\right\| \tag{78}$$

and the proof is complete. ■

Remark

In practical applications (and we will see a typical situation in Section 5), orthogonal projections in usual sense applied to an arbitrary convex constraint may lead to technical difficulties. Our method, however, works independently of the orthogonality concept induced by a concrete scalar product. It may therefore be useful to choose the scalar product so as to fit with the geometry of the convex set Z. To be more specific, consider a non-singular matrix $\boldsymbol{U}\in\Re^{n\times n}$ (that is $\det\boldsymbol{U}\neq0$), and the scalar product

$$\left\langle\boldsymbol{x},\boldsymbol{y}\right\rangle_U=\left\langle\boldsymbol{U}\cdot\boldsymbol{x},\boldsymbol{U}\cdot\boldsymbol{y}\right\rangle_I=\boldsymbol{x}^T\cdot\boldsymbol{U}^T\cdot\boldsymbol{U}\cdot\boldsymbol{y}, \tag{79}$$

and let us derive an explicit form of (44) in this situation. An expression for $\boldsymbol{F}'(\boldsymbol{x})$ is given by (30), that is,

$$\left\langle\boldsymbol{F}'(\boldsymbol{x}),\boldsymbol{q}\right\rangle_U=\left\langle\boldsymbol{H}\cdot\boldsymbol{x}+\boldsymbol{g},\boldsymbol{q}\right\rangle_I, \tag{80}$$

hence

$$\boldsymbol{F}'(\boldsymbol{x})=\boldsymbol{U}^{-1}\cdot\boldsymbol{U}^{-1T}\cdot(\boldsymbol{H}\cdot\boldsymbol{x}+\boldsymbol{g})\quad\forall\boldsymbol{x}\in\boldsymbol{X}. \tag{81}$$

The right-hand side of (80) follows directly from the definition of the scalar function F in (2) and the limit process in (30), whereas the left-hand side of (80) depends on the concrete scalar product which is chosen in this case according to (79). System (44) can therefore be written in the form

$$\begin{aligned}&(i)\quad\boldsymbol{U}\cdot\dot{\boldsymbol{q}}(t)+\boldsymbol{U}^{-1T}\cdot(\boldsymbol{H}\cdot\boldsymbol{w}(t)+\boldsymbol{g})=\boldsymbol{o},\\&(ii)\quad\left\langle\boldsymbol{U}\cdot\dot{\boldsymbol{q}}(t)-\boldsymbol{U}\cdot\dot{\boldsymbol{w}}(t),\boldsymbol{U}\cdot\boldsymbol{w}(t)-\boldsymbol{U}\cdot\boldsymbol{z}\right\rangle_I\geq0\quad\text{a.e.}\quad\forall\boldsymbol{z}\in Z,\\&(iii)\quad\boldsymbol{q}(0)=\boldsymbol{o},\qquad\boldsymbol{w}(0)=\boldsymbol{w}_0,\end{aligned} \tag{82}$$

or, introducting new quantities $\boldsymbol{w}^*=\boldsymbol{U}\cdot\boldsymbol{w}$, $\boldsymbol{q}^*=\boldsymbol{U}\cdot\boldsymbol{q}$, $Z^*=\boldsymbol{U}(Z)$, $\boldsymbol{g}^*=\boldsymbol{U}^{-1T}\cdot\boldsymbol{g}$, $\boldsymbol{H}^*=\boldsymbol{U}^{-1T}\cdot\boldsymbol{H}\cdot\boldsymbol{U}^{-1}$, we obtain the system

$$
\begin{aligned}
&(i)\quad \dot{\boldsymbol{q}}^*(t) = -\boldsymbol{H}^* \cdot \boldsymbol{w}^*(t) - \boldsymbol{g}^*, \\
&(ii)\quad (\dot{\boldsymbol{q}}^*(t) - \dot{\boldsymbol{w}}^*(t))^T \cdot (\boldsymbol{w}^*(t) - \boldsymbol{z}^*) \geq 0 \quad \text{a.e.} \quad \forall \boldsymbol{z}^* \in Z^*, \\
&(iii)\quad \boldsymbol{q}^*(0) = \boldsymbol{o}, \qquad \boldsymbol{w}^*(0) = \boldsymbol{U} \cdot \boldsymbol{w}_0,
\end{aligned}
\tag{83}
$$

which corresponds to the form of (68). Therefore Proposition 3.4 and the results of Section 4 for the system (68) hold also for the transformed system (83). This and the time-discrete system (86) below in Section 4 suggest a hint how to choose the matrix $\boldsymbol{U}$ in (79). It corresponds to an optimal deformation $Z \rightarrow Z^* = \boldsymbol{U}(Z)$ of the convex set Z which ensures the simplest possible numerical evaluation of the orthogonal projection $\boldsymbol{Q}^*$ onto Z^*.

4. Numerical Integration

The starting point for the formulation of a well suited numerical integration procedure for the projected dynamical system (68) on a sampling data system with a constant sampling time T_s is the alternative formulation (35). Applying Lemma 2.2. to the right-hand side of (35) and replacing the left-hand side of (35) by the definition of d$\boldsymbol{w}(t)$/dt leads to

$$
\lim_{\delta \rightarrow +0} \frac{\boldsymbol{w}(t+\delta) - \boldsymbol{w}(t)}{\delta} = \lim_{\delta \rightarrow +0} \frac{1}{\delta} (\boldsymbol{Q}(\boldsymbol{w}(t) - \delta(\boldsymbol{H} \cdot \boldsymbol{w}(t) + \boldsymbol{g})) - \boldsymbol{w}(t)) . \tag{84}
$$

From this follows for a finite length of the stepsize δ the expression

$$
\boldsymbol{w}(t+\delta) = \boldsymbol{Q}(\boldsymbol{w}(t) - \delta(\boldsymbol{H} \cdot \boldsymbol{w}(t) + \boldsymbol{g})) \quad , \quad \boldsymbol{w}(t_0) = \boldsymbol{w}_0 \in Z . \tag{85}
$$

If we assume a constant integration stepsize δ which corresponds to the sampling time T_s the continuous time t and $t + \delta$ in (85) can be replaced by the discrete time kT_s and $(k + 1)T_s$. Neglecting the sampling time T_s in the arguments of $\boldsymbol{w}$ leads to the vector difference equation

$$
\boldsymbol{w}(k+1) = \boldsymbol{Q}(\boldsymbol{w}(k) - T_s(\boldsymbol{H} \cdot \boldsymbol{w}(k) + \boldsymbol{g})) \quad , \quad \boldsymbol{w}(0) = \boldsymbol{w}_0 \in Z \tag{86}
$$

as a time-discrete approximation of (68). From the practical point of view we are searching for the maximum value of T_s for which the discrete-time series $\boldsymbol{w}(k)$, $k \geq 0$ has a unique fixed point $\boldsymbol{w}_*$ which corresponds to the unique equilibrium point $\boldsymbol{w}_\infty$ of (68).

Proposition 4.1 *Let Z, $\boldsymbol{Q}$, $\boldsymbol{H}$, and $\boldsymbol{w}_\infty$ be as above, and let $T_s > 0$ be chosen such that*

$$
T_s < \frac{2}{\lambda_{\max}(\boldsymbol{H})} , \tag{87}
$$

where $\lambda_{\max}(\boldsymbol{H}) > 0$ is the largest eigenvalue of $\boldsymbol{H}$. Let $\boldsymbol{w}(k) \in Z$ be the sequence defined by (86). Then $\boldsymbol{w}(k)$ converges to $\boldsymbol{w}_\infty$ as $k \rightarrow \infty$, and $\boldsymbol{w}_\infty$ is the unique fixed point of the difference equation (86). Moreover, assuming

$$
T_s \leq \frac{2}{\lambda_{\max}(\boldsymbol{H}) + \lambda_{\min}(\boldsymbol{H})} , \tag{88}
$$

we estimate the number N of necessary integration steps for any given accuracy bound $0 < \varepsilon_w < 1$ by the implication

$$
N \geq -\frac{\ln(\varepsilon_w)}{T_s \lambda_{\min}(\boldsymbol{H})} \quad \Rightarrow \quad \frac{\left\| \boldsymbol{w}(N) - \boldsymbol{w}_\infty \right\|}{\left\| \boldsymbol{w}_0 - \boldsymbol{w}_\infty \right\|} < \varepsilon_w . \tag{89}
$$

Proof. We first show that the mapping defined by the right-hand side of (86) is a contraction and then use the Banach Contraction Principle to conclude that it admits a unique fixed point $\boldsymbol{w}_* \in Z$, see [2], that is we look for a constant $L_{\boldsymbol{Q}}(T_s) < 1$ such that

$$\left\|\boldsymbol{Q}(\boldsymbol{w}_2 - T_s(\boldsymbol{H}\cdot\boldsymbol{w}_2 + \boldsymbol{g})) - \boldsymbol{Q}(\boldsymbol{w}_1 - T_s(\boldsymbol{H}\cdot\boldsymbol{w}_1 + \boldsymbol{g}))\right\| \le L_{\boldsymbol{Q}}(T_s)\left\|\boldsymbol{w}_2 - \boldsymbol{w}_1\right\|_2 . \tag{90}$$

For the calculation of the Lipschitz constant $L_{\boldsymbol{Q}}(T_s)$ in (90) the expression

$$\begin{aligned}\left\|\boldsymbol{r}_2 - \boldsymbol{r}_1\right\|^2 &= \left\|\boldsymbol{Q}(\boldsymbol{r}_2) + \boldsymbol{P}(\boldsymbol{r}_2) - (\boldsymbol{Q}(\boldsymbol{r}_1) + \boldsymbol{P}(\boldsymbol{r}_1))\right\|^2 \\ &= \left\|\boldsymbol{Q}(\boldsymbol{r}_2) - \boldsymbol{Q}(\boldsymbol{r}_1)\right\|^2 + \left\|\boldsymbol{P}(\boldsymbol{r}_2) - \boldsymbol{P}(\boldsymbol{r}_1)\right\|^2 + 2(\boldsymbol{P}(\boldsymbol{r}_2) - \boldsymbol{P}(\boldsymbol{r}_1))^T \cdot (\boldsymbol{Q}(\boldsymbol{r}_2) - \boldsymbol{Q}(\boldsymbol{r}_1))\end{aligned} \tag{91}$$

can be used. This leads to

$$\begin{aligned}\left\|\boldsymbol{r}_2 - \boldsymbol{r}_1\right\|^2 &= \left\|\boldsymbol{Q}(\boldsymbol{r}_2) - \boldsymbol{Q}(\boldsymbol{r}_1)\right\|^2 + \left\|\boldsymbol{P}(\boldsymbol{r}_2) - \boldsymbol{P}(\boldsymbol{r}_1)\right\|^2 \\ &\quad + 2\boldsymbol{P}(\boldsymbol{r}_2)^T \cdot (\boldsymbol{Q}(\boldsymbol{r}_2) - \boldsymbol{Q}(\boldsymbol{r}_1)) + 2\boldsymbol{P}(\boldsymbol{r}_1)^T \cdot (\boldsymbol{Q}(\boldsymbol{r}_1) - \boldsymbol{Q}(\boldsymbol{r}_2)).\end{aligned} \tag{92}$$

Due to $\|\boldsymbol{P}(\boldsymbol{r}_2) - \boldsymbol{P}(\boldsymbol{r}_1)\|^2 \ge 0$, $\boldsymbol{Q}(\boldsymbol{r}_1) \in Z$, $\boldsymbol{Q}(\boldsymbol{r}_2) \in Z$, $\boldsymbol{P}(\boldsymbol{r}_2)^T \cdot (\boldsymbol{Q}(\boldsymbol{r}_2) - \boldsymbol{Q}(\boldsymbol{r}_1)) \ge 0$ and $\boldsymbol{P}(\boldsymbol{r}_1)^T \cdot (\boldsymbol{Q}(\boldsymbol{r}_1) - \boldsymbol{Q}(\boldsymbol{r}_2)) \ge 0$ according to (12), (92) permits the estimation

$$\left\|\boldsymbol{r}_2 - \boldsymbol{r}_1\right\| \ge \left\|\boldsymbol{Q}(\boldsymbol{r}_2) - \boldsymbol{Q}(\boldsymbol{r}_1)\right\| . \tag{93}$$

This leads with (90) to

$$\begin{aligned}\left\|\boldsymbol{Q}(\boldsymbol{w}_2 - T_s(\boldsymbol{H}\cdot\boldsymbol{w}_2 + \boldsymbol{g})) - \boldsymbol{Q}(\boldsymbol{w}_1 - T_s(\boldsymbol{H}\cdot\boldsymbol{w}_1 + \boldsymbol{g}))\right\| &\le \left\|(\boldsymbol{I} - T_s\boldsymbol{H})\cdot(\boldsymbol{w}_2 - \boldsymbol{w}_1)\right\| \\ &\le \left\|\boldsymbol{I} - T_s\boldsymbol{H}\right\|\left\|\boldsymbol{w}_2 - \boldsymbol{w}_1\right\| \\ &\le \left|\lambda_{\max\,\mathrm{abs}}(\boldsymbol{I} - T_s\boldsymbol{H})\right|\left\|\boldsymbol{w}_2 - \boldsymbol{w}_1\right\|\end{aligned} \tag{94}$$

where $\lambda_{\mathrm{maxabs}}(\boldsymbol{I} - T_s\boldsymbol{H})$ is the eigenvalue of $\boldsymbol{I} - T_s\boldsymbol{H}$ with maximal absolute value. We can therefore put $L_Q(T_s) = |\lambda_{\mathrm{maxabs}}(\boldsymbol{I} - T_s\boldsymbol{H})|$ provided we ensure that

$$\left|\lambda_{\max\,\mathrm{abs}}(\boldsymbol{I} - T_s\boldsymbol{H})\right| < 1 . \tag{95}$$

A one-to-one correspondence between the eigenvalues of $\boldsymbol{I} - T_s\boldsymbol{H}$ and those of $\boldsymbol{H}$ follow from the formula

$$(\boldsymbol{I} - T_s\boldsymbol{H})\cdot\boldsymbol{x} = \lambda\boldsymbol{x} \quad \Leftrightarrow \quad \boldsymbol{H}\cdot\boldsymbol{x} = \frac{1-\lambda}{T_s}\boldsymbol{x}$$

for $\boldsymbol{x} \in \Re^n$. By virtue of (87), each eigenvalue λ of $\boldsymbol{I} - T_s\boldsymbol{H}$ thus satisfies the inequality

$$-1 < 1 - T_s\lambda_{\max}(\boldsymbol{H}) \le \lambda(\boldsymbol{I} - T_s\boldsymbol{H}) \le 1 - T_s\lambda_{\min}(\boldsymbol{H}) < 1 , \tag{96}$$

and (95) follows. We conclude that there exists a unique $\boldsymbol{w}_* \in Z$ such that

$$\boldsymbol{w}_* = \boldsymbol{Q}(\boldsymbol{w}_* - T_s(\boldsymbol{H}\cdot\boldsymbol{w}_* + \boldsymbol{g})) , \tag{97}$$

and that $\boldsymbol{w}(k)$ converges to $\boldsymbol{w}_*$ as $k \to \infty$. We have to check that $\boldsymbol{w}_* = \boldsymbol{w}_\infty$. Indeed, by virtue of Lemma 2.1, the equilibrium condition (36) is fullfiled if and only if $-\boldsymbol{F}'(\boldsymbol{w}_\infty) = -(\boldsymbol{H}\cdot\boldsymbol{w}_\infty + \boldsymbol{g}) \in N(\boldsymbol{w}_\infty)$, which is in turn equivalent to $-T_s(\boldsymbol{H}\cdot\boldsymbol{w}_\infty + \boldsymbol{g}) \in N(\boldsymbol{w}_\infty)$ for every $T_s > 0$, that is

$$\boldsymbol{o} = \boldsymbol{Q}_{T(\boldsymbol{w}_\infty)}(-T_s(\boldsymbol{H}\cdot\boldsymbol{w}_\infty + \boldsymbol{g})) . \tag{98}$$

Using now Lemma 2.1 with $\boldsymbol{v} = \boldsymbol{o}$, $\boldsymbol{w} = \boldsymbol{w}_\infty$ and $\boldsymbol{q} = -T_s(\boldsymbol{H}\cdot\boldsymbol{w}_\infty + \boldsymbol{g})$ we infer that equations (97) and (98) are equivalent, hence $\boldsymbol{w}_*$ and $\boldsymbol{w}_\infty$ coincide.

It remains to establish the estimate (89) for N under condition (88). It follows from (96) that $L_Q(T_s) = 1 - T_s\lambda_{\min}(\boldsymbol{H})$, so that for every $k \in \aleph$ we have

$$\left\|\boldsymbol{w}(k+1) - \boldsymbol{w}_\infty\right\| \le L_{\boldsymbol{Q}}(T_s)\left\|\boldsymbol{w}(k) - \boldsymbol{w}_\infty\right\|.$$

By induction we obtain

$$\left\|\boldsymbol{w}(N) - \boldsymbol{w}_\infty\right\| \le (L_{\boldsymbol{Q}}(T_s))^N \left\|\boldsymbol{w}_0 - \boldsymbol{w}_\infty\right\|. \tag{99}$$

Assume now that

$$N \ge -\frac{\ln(\varepsilon_{\boldsymbol{w}})}{T_s\lambda_{\min}(\boldsymbol{H})}.$$

Then

$$(L_{\boldsymbol{Q}}(T_s))^N \le \varepsilon_{\boldsymbol{w}}^{\kappa}, \text{ where } \kappa = -\frac{\ln(1 - T_s\lambda_{\min}(\boldsymbol{H}))}{T_s\lambda_{\min}(\boldsymbol{H})}.$$

Taking into account the fact that $-\ln(1 - x) > x$ for every $x < 1$, we obtain $\kappa > 1$, and combining the last inequality with (99) we complete the proof. ■

5. Projection mapping Q for polyhedral constraints with regular matrix U

The calculation of the projection mapping $\boldsymbol{Q}$ in (86) requires the solution of the minimum norm problem (11). The solution of this problem can be equivalently formulated as the solution of the special quadratic program

$$\boldsymbol{Q}(\boldsymbol{r}) = \arg\min_{z\in Z}\{\frac{1}{2}\boldsymbol{z}^T\cdot\boldsymbol{z} - \boldsymbol{r}^T\cdot\boldsymbol{z}\} \tag{100}$$

with $\boldsymbol{r} = \boldsymbol{w}(k) - T_s(\boldsymbol{H}\cdot\boldsymbol{w}(k) + \boldsymbol{g})$ and in general have to be solved numerically for every time step k. In principle this is also the case if we restrict ourselves to the practically important case of a feasible solution set

$$Z = \left\{\boldsymbol{w}\in\Re^n \;\middle|\; \boldsymbol{U}\cdot\boldsymbol{w} - \boldsymbol{u} \ge \boldsymbol{o}\right\} \tag{101}$$

which is defined by linear inequality constraints with a quadratic matrix $\boldsymbol{U} \in \Re^{n\times n}$ and a vector $\boldsymbol{u} \in \Re^n$. Every row $U_{ij}\, w_j \ge u_i$, $i = 1 \ldots n$ in the inequality constraints of (101) defines a half space in $\Re^n$. The convex solution set Z is then given as the intersection of this n half spaces. The convex subset Z formed by the linear inequality contraints in (101) is a polyhedral cone with vertex at $\boldsymbol{u}$ and thus a special case of the more general convex subset Z formed by the concave function according to (4).

If we assume that the convex subset Z is an orthant which means a convex polyhedron defined by

$$Z = \left\{\boldsymbol{w}\in\Re^n \;\middle|\; \boldsymbol{I}\cdot\boldsymbol{w} - \boldsymbol{u} \ge \boldsymbol{o}\right\} \tag{102}$$

with the identity matrix $\boldsymbol{I} \in \Re^{n \times n}$, the solution of the quadratic program (100) and thus the projection mapping $\boldsymbol{Q}$ can be formulated analytically. Fig. 2 shows the geometrical interpretation of the projection mapping $\boldsymbol{Q}$ in the plane.

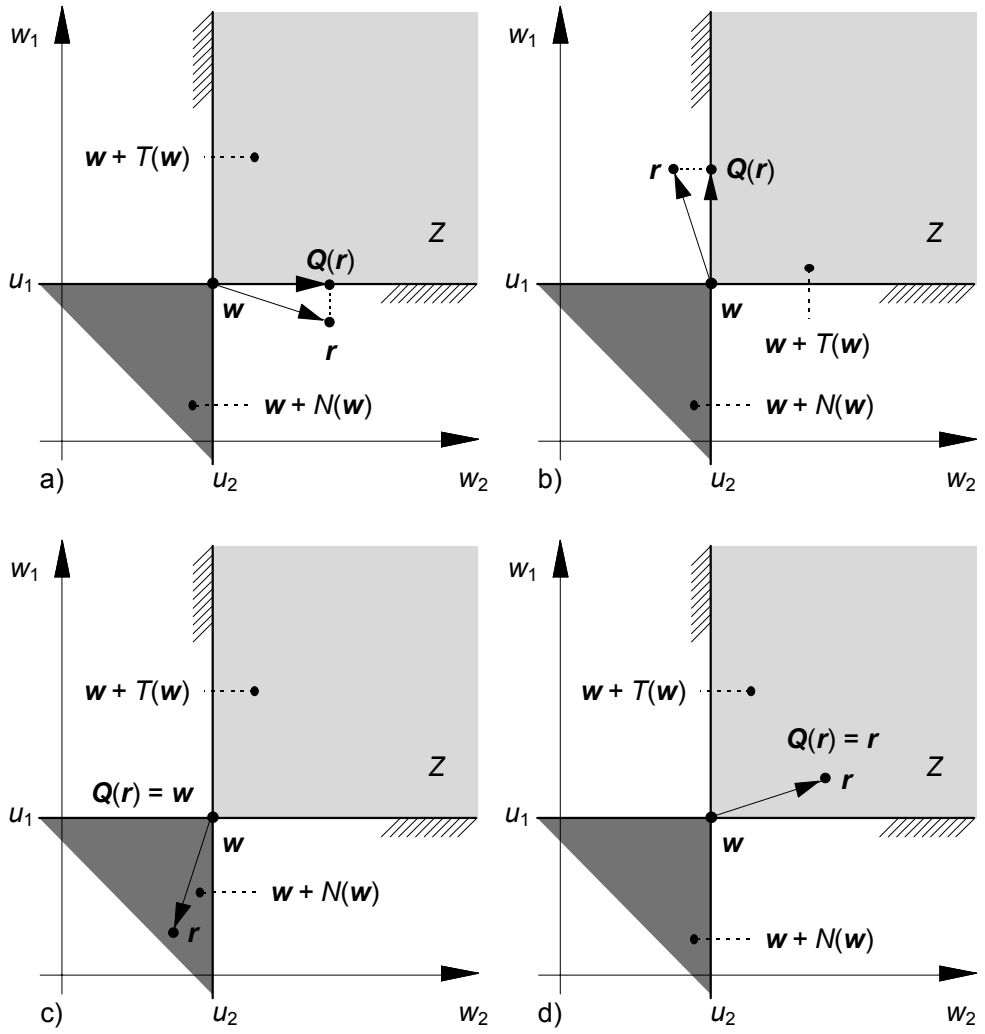


Figure 2: Geometrical interpretation of the projection mapping $\boldsymbol{Q}$ for an orthant (102) with $n = 2$ in a corner $\boldsymbol{w}$: a,b) $\boldsymbol{r} \notin \boldsymbol{w} + N(\boldsymbol{w}) \cup \boldsymbol{w} + T(\boldsymbol{w})$ c) $\boldsymbol{r} \in \boldsymbol{w} + N(\boldsymbol{w})$ d) $\boldsymbol{r} \in \boldsymbol{w} + T(\boldsymbol{w})$

Obviously in this special case the solution of the quadratic program (100) can be expressed explicitly by

$$\boldsymbol{Q}(\boldsymbol{r})_i = \begin{cases} r_i & , \quad r_i \geq u_i \\ u_i & , \quad r_i < u_i \end{cases} , \quad i = 1 \,..\, n \,. \tag{103}$$

In this case the solution of the vector difference equation (86) decouples into the solution of n scalar difference equations which can be calculated independently after the determination of the gradient $\boldsymbol{F}'(\boldsymbol{w}(k)) = \boldsymbol{H} \cdot \boldsymbol{w}(k) + \boldsymbol{g}$ of the unbounded problem.

According to the remark at the end of section 3 a deformation of the convex set Z defined by (101) to an orthant $Z^* \rightarrow \boldsymbol{U}(Z)$ defined by (102) is always possible by choosing the concrete scalar product (79) instead of the canonical one defined by (9). In this case the corresponding

projection mapping $\boldsymbol{Q}^*$ in the numerical formulation of (83) can be calculated according to (103) instead of (100) and this justifies the choice of the scalar product in (81)-(82).

6. Application example: Hysteresis filter synthesis with the modified Prandtl-Ishlinskii approach

The modified Prandtl-Ishlinskii hysteresis operator has been developed recently for the modeling and compensation of asymmetrically complex hysteretic nonlinearities [7]. It is defined as the concatenation of a Prandtl-Ishlinskii hysteresis operator H and a Prandtl-Ishlinskii superposition operator S and in vector notation is given by

$$\Gamma[x](t) = S[H[x]](t) = \boldsymbol{w}_S^{\,T} \cdot \boldsymbol{S}_{\boldsymbol{r}_S}[\boldsymbol{w}_H^{\,T} \cdot \boldsymbol{H}_{\boldsymbol{r}_H}[x,\boldsymbol{z}_{H0}]](t)\,. \tag{104}$$

The operator H consists of a weighted linear superposition of d+1 elementary play operators H_{rH}, which are included in (104) in the d+1-dimensional vector $\boldsymbol{H}_{\boldsymbol{r}H}$. The rate-independent characteristic of the play is characterised by the threshold-dependent x-y-trajectory, see Fig. 3a. The weights w_{Hi}, the thresholds r_{Hi} and the initial values z_{H0i}, $i = 0\,..\,d$ of the play operators are considered in the vector notation (104) by the vector of weights $\boldsymbol{w}_H^{\,T} = (w_{H0}\ w_{H1}\ ..\ w_{Hd})$, the vector of thresholds $\boldsymbol{r}_H^{\,T} = (r_{H0}\ r_{H1}\ ..\ r_{Hd})$ with $r_{H0} = 0$ and the vector of the initial values $\boldsymbol{z}_{H0}^{\,T} = (z_{H00}\ z_{H01}\ ..\ z_{H0d})$. The outputs of the elementary operators $z_{Hi} = H_{rH}[x,z_{H0i}]$, $i = 0\,..\,d$ represent the inner system state or the memory of the discrete-threshold Prandtl-Ishlinskii hysteresis operator.

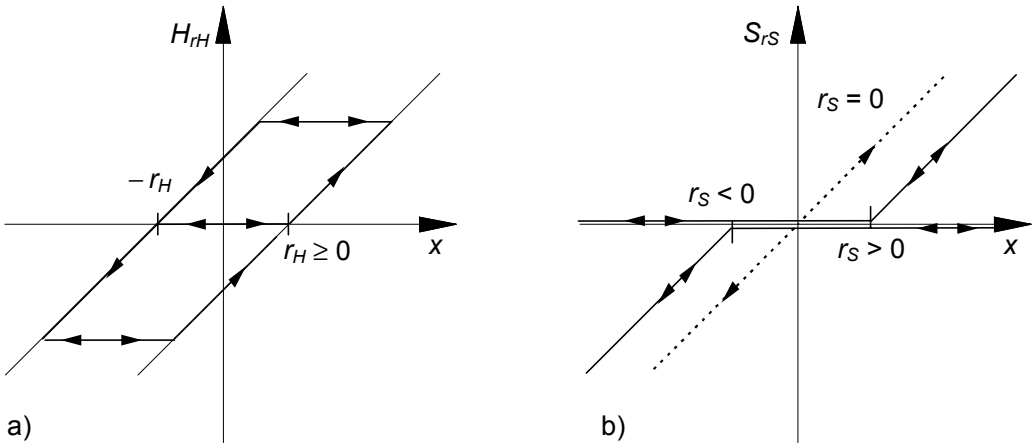


Figure 3: x-y-trajectory of the play (a) and the one-sided dead-zone operator (b)

The memoryless superposition operator S describes the deviation of the real characteristic from the odd symmetry property of the operator H [7]. It consists of the weighted linear superposition of $2l$+1 one-sided dead-zone operators S_{rS}, which are included in (104) in the $2l$+1-dimensional vector $\boldsymbol{S}_{\boldsymbol{r}S}$. The rate-independent transfer characteristic is characterised by the threshold-dependent x-y-trajectory shown in Fig. 3b. The weights w_{Si} and the thresholds r_{Si}, $i = -l\,..\,+l$ of the one-sided dead-zone operators are considered in the vector notation (104) by the vector of weights $\boldsymbol{w}_S^{\,T} = (w_{S-l}\,..\,w_{S0}\,..\,w_{Sl})$ and the vector of thresholds $\boldsymbol{r}_S^{\,T} = (r_{S-l}\,..\,r_{S0}\,..\,r_{Sl})$ with $r_{S0} = 0$. The corresponding compensator

$$\Gamma^{-1}[y](t) = H^{-1}[S^{-1}[y]](t) = \boldsymbol{w}'^{\,T}_H \cdot \boldsymbol{H}_{\boldsymbol{r}'_H}[\boldsymbol{w}'^{\,T}_S \cdot \boldsymbol{S}_{\boldsymbol{r}'_S}[y],\boldsymbol{z}'_{H0}](t) \tag{105}$$

exists uniquely in the convex polyhedron

$$\Omega = \left\{ \begin{pmatrix} \boldsymbol{w}_H \\ \boldsymbol{w}_S \end{pmatrix} \in \begin{pmatrix} \Re^{d+1} \\ \Re^{2l+1} \end{pmatrix} \;\middle|\; \begin{pmatrix} \boldsymbol{U}_H & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{U}_S \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{w}_H \\ \boldsymbol{w}_S \end{pmatrix} - \begin{pmatrix} \boldsymbol{u}_H \\ \boldsymbol{u}_S \end{pmatrix} \geq \begin{pmatrix} \boldsymbol{o} \\ \boldsymbol{o} \end{pmatrix} \right\} \tag{106}$$

and follows from the inversion of H and S and the concatenation of S^{-1} and H^{-1} [7]. The matrices and vectors $\boldsymbol{U}_H \in \Re^{d+1\times d+1}$, $\boldsymbol{u}_H \in \Re^{d+1}$, $\boldsymbol{U}_S \in \Re^{2l+1\times 2l+1}$ and $\boldsymbol{u}_S \in \Re^{2l+1}$ in the inequality constraints in (106) are given by

$$\boldsymbol{U}_H = \begin{pmatrix} 1 & 0 & . & . & 0 \\ 1 & 1 & . & . & 0 \\ . & . & . & & . \\ . & . & & . & . \\ 1 & 1 & . & . & 1 \end{pmatrix}, \qquad \boldsymbol{u}_H = \begin{pmatrix} \varepsilon \\ \varepsilon \\ . \\ . \\ \varepsilon \end{pmatrix},$$

$$\boldsymbol{U}_S = \begin{pmatrix} 1 & \cdots & 1 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \iddots & \vdots \\ 0 & \cdots & 1 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 & 1 & \cdots & 0 \\ \vdots & \iddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 1 & \cdots & 1 \end{pmatrix} \text{ and } \boldsymbol{u}_S = \begin{pmatrix} \varepsilon \\ \vdots \\ \varepsilon \\ \varepsilon \\ \varepsilon \\ \vdots \\ \varepsilon \end{pmatrix}.$$

$\varepsilon > 0$ is a lower bound and a design parameter which permits the change of strict inequality constraints by the inequality constraints in (106).

According to (105) the model structures of the Prandtl-Ishlinskii operators H and S are obviously invariant with respect to the inversion operation. For the calculation of the inverse filter from the model and vice versa only the thresholds $\boldsymbol{r}_H$ and $\boldsymbol{r}_H$', the weights $\boldsymbol{w}_H$ and $\boldsymbol{w}_H$'and the initial values $\boldsymbol{z}_{H0}$ and $\boldsymbol{z}_{H0}$' have to be calculated by the corresponding mappings

$$\boldsymbol{w}_H' = \boldsymbol{\Phi}_H(\boldsymbol{w}_H) \quad \text{or} \quad \boldsymbol{w}_H = \boldsymbol{\Phi}_H(\boldsymbol{w}_H'), \tag{107}$$

$$\boldsymbol{r}_H' = \boldsymbol{\Psi}_H(\boldsymbol{r}_H, \boldsymbol{w}_H) \quad \text{or} \quad \boldsymbol{r}_H = \boldsymbol{\Psi}_H(\boldsymbol{r}_H', \boldsymbol{w}_H') \tag{108}$$

$$\boldsymbol{z}_{H0}' = \boldsymbol{\Theta}_H(\boldsymbol{z}_{H0}, \boldsymbol{w}_H) \quad \text{or} \quad \boldsymbol{z}_{H0} = \boldsymbol{\Theta}_H(\boldsymbol{z}_{H0}', \boldsymbol{w}_H') \tag{109}$$

$$\boldsymbol{w}_S' = \boldsymbol{\Phi}_S(\boldsymbol{w}_S) \quad \text{or} \quad \boldsymbol{w}_S = \boldsymbol{\Phi}_S(\boldsymbol{w}_S') \tag{110}$$

and

$$\boldsymbol{r}_S' = \boldsymbol{\Psi}_S(\boldsymbol{r}_S, \boldsymbol{w}_S) \quad \text{or} \quad \boldsymbol{r}_S = \boldsymbol{\Psi}_S(\boldsymbol{r}_S', \boldsymbol{w}_S') \,. \tag{111}$$

The derivation of these mappings is not the aim of this article. For this purpose we refer to the original papers [7].

6.1 Generalised error model

The invariance of the Prandtl-Ishlinskii operators H and S with respect to the inversion operation makes it possible to generate a generalised error model

$$e(t) = H[x](t) - S^{-1}[y](t) = \boldsymbol{w}_H^{\,T} \cdot \boldsymbol{H}_{\boldsymbol{r}_H}[x, \boldsymbol{z}_{H0}](t) - \boldsymbol{w}_S'^{\,T} \cdot \boldsymbol{S}_{\boldsymbol{r}_S'}[y](t) \tag{112}$$

which depends linearly on the weights $\boldsymbol{w}_H$ and $\boldsymbol{w}_S$'. This generalised error model is the starting point for the synthesis of the modified Prandtl-Ishlinskii hysteresis operator Γ and its compensator Γ^{-1} starting from the measured input signal $x(t)$ and output signal $y(t)$. For the real values $\boldsymbol{r}_H$, $\boldsymbol{w}_H$, $\boldsymbol{z}_{H0}$, $\boldsymbol{r}_S$', $\boldsymbol{w}_S$' of the generalised error model follows

$$e(t) = w_{H0}\left(\frac{\boldsymbol{w}_H^{\,T}}{w_{H0}} \cdot \boldsymbol{H}_{\boldsymbol{r}_H}[x,\boldsymbol{z}_{H0}](t) - \frac{\boldsymbol{w}_S'^{\,T}}{w_{H0}} \cdot \boldsymbol{S}_{\boldsymbol{r}_S'}[y](t)\right) = 0$$

for all t and $w_{H0} > 0$. Therefore the expression between the parentheses has to be zero for all times. Thus the error model is overdetermined with one degree of freedom and consequently $w_{H0} = 1$ can be given as a real value. With $x = H_{rH0}[x,z_{H00}]$ follows the alternative representation (1)

$$e(t) = \zeta(t) + \boldsymbol{\Psi}^T(t) \cdot \boldsymbol{w}$$

with

$$\zeta(t) = x(t)\,,$$

$$\boldsymbol{w}^T = \begin{pmatrix} w_{H1} & \cdots & w_{Hd} & \boldsymbol{w}_S'^{\,T} \end{pmatrix}$$

and

$$\boldsymbol{\Psi}^T(t) = \begin{pmatrix} H_{r_{H1}}[x,\boldsymbol{z}_{H01}](t) & \cdots & H_{r_{Hd}}[x,\boldsymbol{z}_{H0d}](t) & -\boldsymbol{S}_{\boldsymbol{r}_S'}(y(t)) \end{pmatrix}.$$

The synthesis of the modified Prandtl-Ishlinskii hysteresis operator Γ and its compensator Γ^{-1} is realised in four steps.

6.2 Model and compensator synthesis procedure

In the first synthesis step the thresholds and the initial hysteretic state values are determined by the formulas

$$r_{Hi} = \frac{i}{d+1} \max_{0 \le t \le t_E} \{|x(t)|\} \quad ; \quad i = 1\,..\,d\,, \tag{113}$$

$$\begin{aligned} r_{Si}' &= \frac{i}{l+1} \min_{0 \le t \le t_E} \{y(t)\} \quad ; \quad i = -l\,..\,-1 \\ r_{Si}' &= \frac{i}{l+1} \max_{0 \le t \le t_E} \{y(t)\} \quad ; \quad i = +0\,..\,+l \end{aligned} \tag{114}$$

and

$$z_{H0i} = 0 \quad ; \quad i = 1\,..\,d\,. \tag{115}$$

In addition to the model orders d and l, the maximum of the absolute value of the measured input signal and the maximum and minimum value of the output signal must be given. Moreover, during identification (115) assumes the evolution of the hysteretic state from the so-called „virgin" or „demagnetised" state.

The determination of the weights is the aim of the second step and follows from the constrained least-square minimisation of the error model (1), which means by the quadratic minimisation of (2) with the polyhedral constraints (101). The matrix $\boldsymbol{U}$ in (101) follows from

the matrix in (106) by canceling the first row and column. The vector $\boldsymbol{u}^T$ in (101) results from the vector $(\boldsymbol{u}_H{}^T\ \boldsymbol{u}_S{}^T) - (\boldsymbol{i}^T\ \boldsymbol{o}^T)$ with $\boldsymbol{i}^T = (1\ 1\ .\ .\ 1)$ followed by canceling the first column.

This procedure guarantee the existence of the operators H^{-1} and S starting from the operators H and S^{-1} and vice versa and thus the applicability of the transformation mappings (107-111). According to the last sections the quadratic minimisation of (2) with the polyhedral constraints (101) can be realized by the numerical time-integration of the corresponding projected dynamical system (68). For this purpose the initial values of (68) must lie within the feasible region. A meaningful assumption for the determination of the initial values is to have no information about the hysteretic nonlinearity. In this case it is meaningful to choose the initial values for the error model in such a way that the model Γ and its compensator Γ^{-1} exhibit the behaviour of the identity operator I. From this idea follows the initial values

$$\boldsymbol{w}_0{}^T = (\underbrace{0 \quad \cdots \quad 0}_{d}\ \underbrace{0 \quad \cdots \quad 0 \quad 1 \quad 0 \quad \cdots \quad 0}_{2l+1}) \tag{116}$$

with $w_{S00} = 1$.

According to section 3 the stability properties of the projected dynamical system (68) depend decisively on the definiteness properties of the matrix $\boldsymbol{H}$ according to (65). Condition (65) demands the nonexistence of any vector $\boldsymbol{w}$ apart from the zero vector which is perpendicular to the signal vector $\boldsymbol{\Psi}(t)$ generated by the elementary operators for all $0 \le t \le t_E$. This condition is fulfilled if the components of the signal vector are all different from zero and moreover linearly independent in the interval $0 \le t \le t_E$. The latter condition is given by the fact that because of (113-114) the elementary operators in the signal vector have different thresholds. The former condition requires the crossing of every threshold r_{Hi}, $i = 1\ .\ .\ d$ and $r_S{}'_i$, $i = -l\ .\ .\ l$ by the amplitudes of the input signal $x(t)$ and the output signal $y(t)$. Because of (113-114) this property of the input and output signal is given a-priori as well.

In the third step the weights of the identified Prandtl-Ishlinskii hysteresis operator H have to be completed by

$$\boldsymbol{w}_H{}^T = (1 \quad \boldsymbol{w}_1 \quad \cdots \quad \boldsymbol{w}_d)\,. \tag{117}$$

In the fourth step the corresponding model Γ and the corresponding compensator Γ^{-1} are generated by the transformation mappings (110-111) and (107)-(109), respectively.

6.3 Identification results

To test the filter design method presented above, a micropositioning stage based on a piezoelectric stack actuator as a driving element was realized. Fig. 4 illustrates the test and measurement equipment of the experimental part. It consists of a digital signal processor (DSP) which generates the digital driving signal for the piezoelectric actuator. This signal is converted by a 12-bit digital-to-analog converter and amplified by an high-voltage power amplifier to an analog high-voltage signal of up to 1 kV. The output of the high-voltage power amplifier is measured by an additional voltage measurement circuit and converted to a digital signal by a 12-bit analog-to-digital converter. The displacement of the actuator is directly digitally measured by a high-precision laser-interferometer with a resolution of 5 nm. The signals from the voltage measurement circuit and the interferometer are fed back to the DSP for characterization, identification and control purposes.

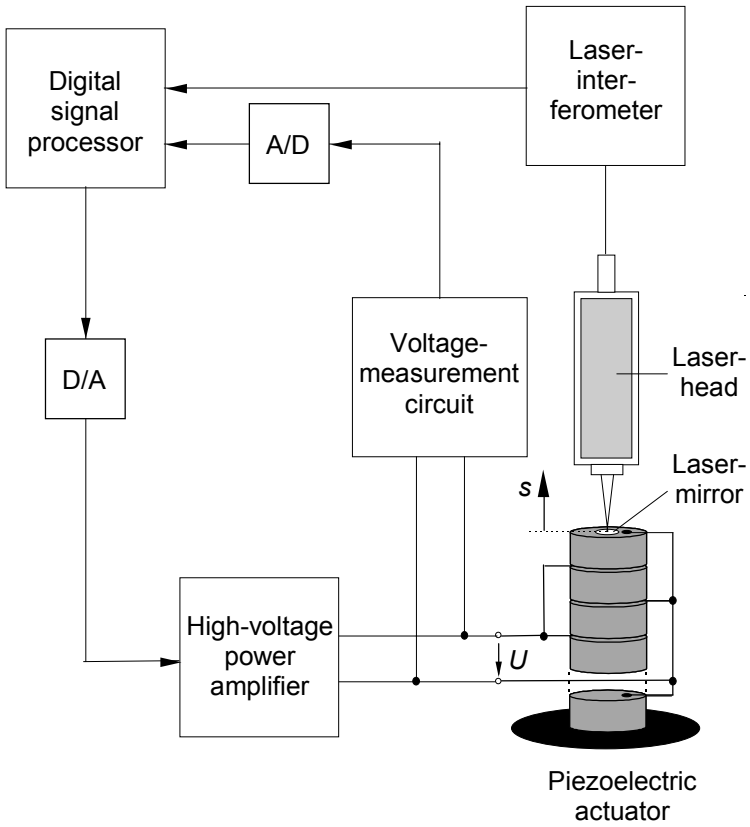


Figure 4: Test and measurement equipment of the experimental part

The measured hysteretic relation W between the driving voltage signal U and the displacement signal s of the actuator in the large signal range is shown in Fig. 5a. The light grey curve in Fig. 5b depicts the relationship between the voltage U as the input quantity x and the displacement s as the output quantity y modeled by a modified Prandtl-Ishlinskii hysteresis operator Γ with the model orders $d = 5$ and $l = 2$ and identified with the procedure described in section 6.2. The relative modeling error defined by

$$e_{\Gamma} = \frac{\left\| \Gamma[U] - s \right\|_{\infty}}{\left\| \Gamma[U] \right\|_{\infty}} \tag{118}$$

amounts 17.3% for the best linear approximation through the origin and 1.5% in the case of the operator-based approximation. The black curve in Fig 5b shows the relationship between the given displacement s_c and the voltage U of the corresponding inverse modified Prandtl-Ishlinskii hysteresis operator Γ^{-1}. The dark grey curve in Fig. 5b displays the compensation effect of the concatenation $W[\Gamma^{-1}]$ of the inverse hysteretic filter and the real hysteretic nonlinearity of the piezoelectric actuator. The relative feedforward control error, definded by

$$e_{W[\Gamma^{-1}]} = \frac{\left\| W[\Gamma^{-1}[s_s]] - s_s \right\|_{\infty}}{\left\| W[\Gamma^{-1}[s_s]] \right\|_{\infty}} \tag{119}$$

amounts in this case to 2.2%. The thresholds and weights of the generalized error model (112) as well as the thresholds and weights of modified Prandtl-Ishlinskii hysteresis operator (104) and ist inverse (105) are shown in table 1.

l	r_H	w_H	r_H'	w_H'
0	$+0.0\cdot10^{+0}$	$+1.0\cdot10^{+0}$	$+0.0\cdot10^{+0}$	$+1.0\cdot10^{+0}$
1	$+8.3\cdot10^{+1}$	$+3.0\cdot10^{-1}$	$+8.3\cdot10^{+1}$	$-2.3\cdot10^{-1}$
2	$+1.6\cdot10^{+2}$	$+2.0\cdot10^{-1}$	$+1.9\cdot10^{+2}$	$-1.0\cdot10^{-1}$
3	$+2.5\cdot10^{+2}$	$+1.5\cdot10^{-1}$	$+3.2\cdot10^{+2}$	$-6.0\cdot10^{-2}$
4	$+3.3\cdot10^{+2}$	$+1.4\cdot10^{-1}$	$+4.5\cdot10^{+2}$	$-4.9\cdot10^{-2}$
5	$+4.2\cdot10^{+2}$	$+1.6\cdot10^{-1}$	$+6.0\cdot10^{+2}$	$-4.6\cdot10^{-2}$

j	r_S	w_S	r_S'	w_S'
–2	$-5.3\cdot10^{+2}$	$+2.6\cdot10^{-3}$	$-1.6\cdot10^{+1}$	$-2.4\cdot10^{-2}$
–1	$-2.7\cdot10^{+2}$	$+2.5\cdot10^{-3}$	$-8.0\cdot10^{+0}$	$-2.7\cdot10^{-2}$
0	$+0.0\cdot10^{+0}$	$+2.8\cdot10^{-2}$	$+0.0\cdot10^{+0}$	$+3.5\cdot10^{-1}$
+1	$+2.4\cdot10^{+2}$	$-1.6\cdot10^{-3}$	$+6.5\cdot10^{+0}$	$+2.3\cdot10^{-2}$
+2	$+4.9\cdot10^{+2}$	$-2.1\cdot10^{-3}$	$+1.3\cdot10^{+1}$	$+3.5\cdot10^{-2}$

Table 1: Thresholds and weights of the modified Prandtl-Ishlinskii hysteresis operator Γ and the inverse Γ^{-1}

The thresholds $\boldsymbol{r}_H$ and $\boldsymbol{r}_S$' of the generalized error model was determined according to (113), (114) with $\max\{|U(t)|\} = +500\text{V}$, $\min\{s(t)\} = -24{,}08\mu\text{m}$ and $\max\{s(t)\} = +19{,}38\mu\text{m}$ for $0 \le t \le t_E$. The weights $\boldsymbol{w}_H$ and $\boldsymbol{w}_S$' of the generalized error model (112) was determined by the time integration of the projected dynamical system (68) which corresponds to the alternative formulation (1) of the generalized error model (112) and by the completion of the weights according to (117). This procedure is described in detail below. The remaining thresholds $\boldsymbol{r}_S$ and weights $\boldsymbol{w}_S$ of the modified Prandtl-Ishlinskii hysteresis operator as well as the remaining thresholds $\boldsymbol{r}_H$' and weights $\boldsymbol{w}_H$' of the inverse modified Prandtl-Ishlinskii hysteresis operator results from the transformation mappings (110-111) and (107)-(109), respectively.

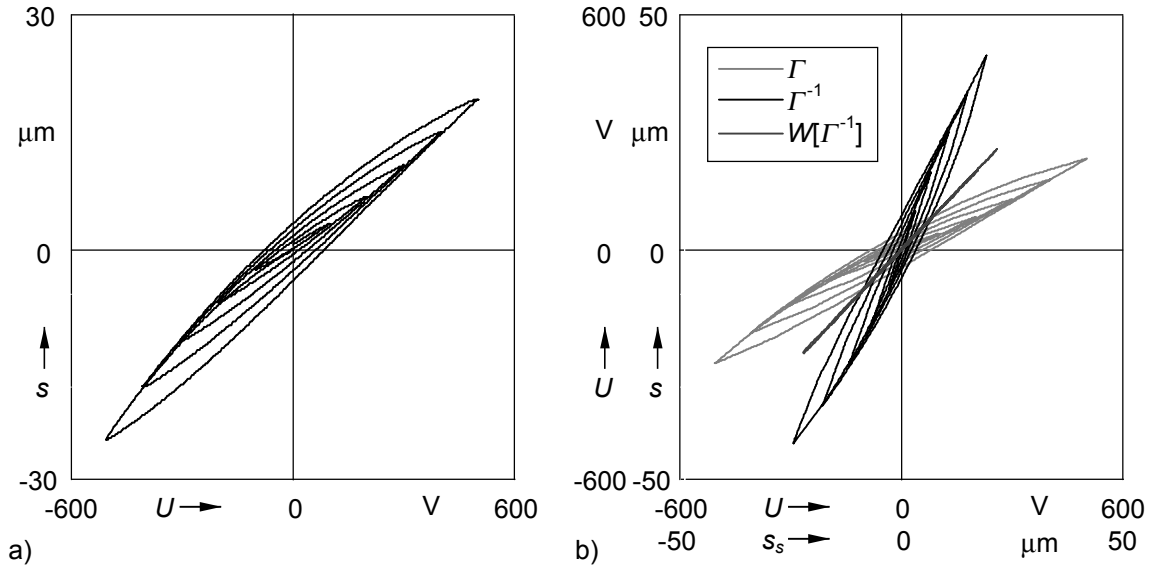


Figure 5: a) Complex hysteretic nonlinearity W of a piezoelectric stack actuator b) Modified Prandtl-Ishlinskii operator Γ, the inverse Γ^{-1} and the concatenation $W[\Gamma^{-1}]$

The determination of the weights $\boldsymbol{w}$ of the error model (1) with projected dynamical systems was the main subject of this article. The application of this method should now be demonstrated. With the input and output measurement data shown in Fig. 5a we obtain a Hessian matrix

$$\boldsymbol{H} = \begin{pmatrix} +1.3\cdot10^{+8} & +1.1\cdot10^{+8} & +7.0\cdot10^{+7} & +3.5\cdot10^{+7} & +1.1\cdot10^{+7} & -3.1\cdot10^{+5} & -1.4\cdot10^{+6} & -6.0\cdot10^{+6} & -1.3\cdot10^{+6} & -2.9\cdot10^{+5} \\ +1.1\cdot10^{+8} & +9.7\cdot10^{+7} & +7.0\cdot10^{+7} & +3.8\cdot10^{+7} & +1.3\cdot10^{+7} & -2.5\cdot10^{+5} & -1.1\cdot10^{+6} & -4.7\cdot10^{+6} & -1.0\cdot10^{+6} & -2.3\cdot10^{+5} \\ +7.0\cdot10^{+7} & +7.0\cdot10^{+7} & +5.8\cdot10^{+7} & +3.6\cdot10^{+7} & +1.4\cdot10^{+7} & -1.8\cdot10^{+5} & -7.7\cdot10^{+5} & -3.1\cdot10^{+6} & -6.8\cdot10^{+5} & -1.7\cdot10^{+5} \\ +3.5\cdot10^{+7} & +3.8\cdot10^{+7} & +3.6\cdot10^{+7} & +2.6\cdot10^{+7} & +1.1\cdot10^{+7} & -1.2\cdot10^{+5} & -4.4\cdot10^{+5} & -1.6\cdot10^{+6} & -3.8\cdot10^{+5} & -1.0\cdot10^{+5} \\ +1.1\cdot10^{+7} & +1.3\cdot10^{+7} & +1.4\cdot10^{+7} & +1.1\cdot10^{+7} & +6.5\cdot10^{+6} & -5.3\cdot10^{+4} & -1.4\cdot10^{+5} & -4.9\cdot10^{+5} & -1.4\cdot10^{+5} & -4.6\cdot10^{+5} \\ -3.1\cdot10^{+5} & -2.5\cdot10^{+5} & -1.8\cdot10^{+5} & -1.2\cdot10^{+5} & -5.3\cdot10^{+4} & +4.2\cdot10^{+3} & +1.0\cdot10^{+4} & +1.7\cdot10^{+4} & +0.0\cdot10^{+0} & +0.0\cdot10^{+0} \\ -1.4\cdot10^{+6} & -1.1\cdot10^{+6} & -7.7\cdot10^{+5} & -4.4\cdot10^{+5} & -1.4\cdot10^{+5} & +1.0\cdot10^{+4} & +3.8\cdot10^{+4} & +7.1\cdot10^{+4} & +0.0\cdot10^{+0} & +0.0\cdot10^{+0} \\ -6.0\cdot10^{+6} & -4.7\cdot10^{+6} & -3.1\cdot10^{+6} & -1.6\cdot10^{+6} & -4.9\cdot10^{+5} & +1.7\cdot10^{+4} & +7.1\cdot10^{+4} & +2.8\cdot10^{+5} & +5.5\cdot10^{+4} & +1.3\cdot10^{+4} \\ -1.3\cdot10^{+6} & -1.0\cdot10^{+6} & -6.8\cdot10^{+5} & -3.8\cdot10^{+5} & -1.4\cdot10^{+5} & +0.0\cdot10^{+0} & +0.0\cdot10^{+0} & +5.5\cdot10^{+4} & +3.0\cdot10^{+4} & +7.9\cdot10^{+3} \\ -2.9\cdot10^{+5} & -2.3\cdot10^{+5} & -1.7\cdot10^{+5} & -1.0\cdot10^{+5} & -4.6\cdot10^{+5} & +0.0\cdot10^{+0} & +0.0\cdot10^{+0} & +1.3\cdot10^{+4} & +7.9\cdot10^{+3} & +3.1\cdot10^{+3} \end{pmatrix}$$

a vector

$$\boldsymbol{g}^T = \begin{pmatrix} +1.3\cdot10^{+8} & +9.7\cdot10^{+7} & +5.8\cdot10^{+7} & +2.6\cdot10^{+7} & +6.9\cdot10^{+6} & -3.4\cdot10^{+5} & -1.5\cdot10^{+6} & -6.4\cdot10^{+6} & -1.3\cdot10^{+6} & -3.2\cdot10^{+5} \end{pmatrix}$$

and a scalar $f = +7.6\cdot10^{+7}$ as well as a matrix of inequality constraints

$$\boldsymbol{U} = \begin{pmatrix} +1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ +1 & +1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ +1 & +1 & +1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ +1 & +1 & +1 & +1 & 0 & 0 & 0 & 0 & 0 & 0 \\ +1 & +1 & +1 & +1 & +1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & +1 & +1 & +1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & +1 & +1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & +1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & +1 & +1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & +1 & +1 & +1 \end{pmatrix}$$

and a vector of inequality constraints

$$\boldsymbol{u}^T = \begin{pmatrix} +\varepsilon-1 & +\varepsilon-1 & +\varepsilon-1 & +\varepsilon-1 & +\varepsilon-1 & +\varepsilon & +\varepsilon & +\varepsilon & +\varepsilon & +\varepsilon \end{pmatrix}$$

with $\varepsilon = 1\cdot10^{-6}$. In this case the inverse of the matrix $\boldsymbol{U}$ is give by

$$\boldsymbol{U}^{-1} = \begin{pmatrix} +1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & +1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & +1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & +1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & +1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & +1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & +1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & +1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & +1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & +1 \end{pmatrix}.$$

The determination of the largest and smallest eigenvalue of the matrix $\boldsymbol{U}^{-1T}\boldsymbol{H}\boldsymbol{U}^{-1}$ yields the parameter $\lambda_{\min}(\boldsymbol{U}^{-1T}\boldsymbol{H}\boldsymbol{U}^{-1}) = 9.25\cdot10^{+2}$ 1/s and $\lambda_{\max}(\boldsymbol{U}^{-1T}\boldsymbol{H}\boldsymbol{U}^{-1}) = 3.16\cdot10^{+7}$ 1/s. The numerical integration is carried out with an error bound of $\varepsilon_{\boldsymbol{w}}^* = 1\cdot10^{-6}$ and an integration step size of $T_s = (\lambda_{\max}(\boldsymbol{U}^{-1T}\boldsymbol{H}\boldsymbol{U}^{-1}) + \lambda_{\min}(\boldsymbol{U}^{-1T}\boldsymbol{H}\boldsymbol{U}^{-1}))^{-1} = 31.64\cdot10^{-9}$ s. From this follows a necessary number of integration steps of $N = 4.73\cdot10^{+5}$.

Table 2 shows the initial values $\boldsymbol{w}_0$ and the equilibrium values $\boldsymbol{w}_\infty(\varepsilon_{\boldsymbol{w}}^*)$ after the numerical integration. Fig. 6 shows the time evolution of the distance measure F which decreases monotonically from the initial value $F(\boldsymbol{w}_0) = 6.9\cdot10^{+7}$ to the finite value $F(\boldsymbol{w}_\infty(\varepsilon_{\boldsymbol{w}}^*)) = 2.9\cdot10^{+4}$ in the equilibrium point.

i	$\boldsymbol{w}_{E0}$	$\boldsymbol{w}_{E\infty}(\varepsilon_{\boldsymbol{w}}^*)$
1	$+0.0\cdot10^{+0}$	$+3.0\cdot10^{-1}$
2	$+0.0\cdot10^{+0}$	$+2.0\cdot10^{-1}$
3	$+0.0\cdot10^{+0}$	$+1.5\cdot10^{-1}$
4	$+0.0\cdot10^{+0}$	$+1.4\cdot10^{-1}$
5	$+0.0\cdot10^{+0}$	$+1.6\cdot10^{-1}$
6	$+0.0\cdot10^{+0}$	$-2.4\cdot10^{-2}$
7	$+0.0\cdot10^{+0}$	$-2.7\cdot10^{-2}$
8	$+1.0\cdot10^{+0}$	$+3.5\cdot10^{-1}$
9	$+0.0\cdot10^{+0}$	$+2.3\cdot10^{-2}$
10	$+0.0\cdot10^{+0}$	$+3.5\cdot10^{-2}$

Table2: Initial values $\boldsymbol{w}_0$ and equilibrium values $\boldsymbol{w}_\infty(\varepsilon_{\boldsymbol{w}}^*)$

The dashed curve in Fig. 6 displays the time evolution of F in the unconstrained case $Z = \Re^{10}$ and the solid curve depicts the time evolution of F in the constrained case $Z = \{\boldsymbol{w} \in \Re^{10} \mid \boldsymbol{U}\cdot\boldsymbol{w} - \boldsymbol{u} \geq \boldsymbol{o}\}$. Obviously the influence of the projection mechanism to the time evolution of F is only weak.

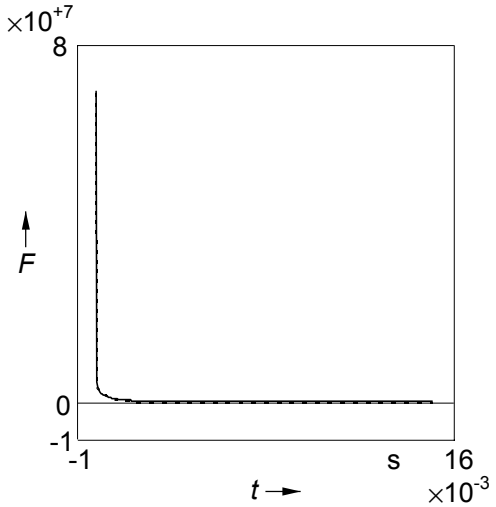
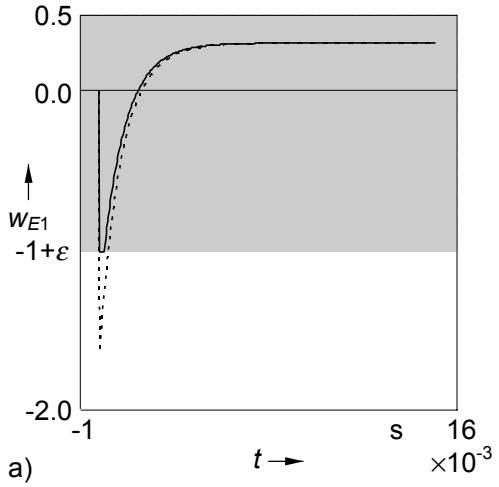


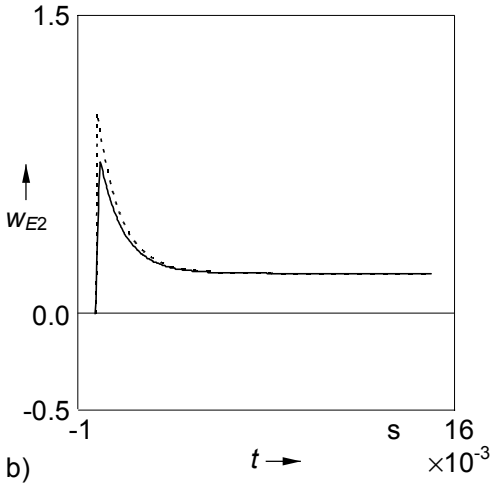
Figure 6: Time evolution of the distance measure F

The Figs. 7a-j shows the time evolution of the single components w_i, $i = 1 \ldots 10$ of the vector $\boldsymbol{w}$. Also in these cases the dashed curves display the time evolution of w_i, $i = 1 \ldots 10$ in the unconstraind case $Z = \Re^{10}$ and the solid curves depicts the time evolution of of w_i, $i = 1 \ldots 10$ in the constrained case $Z = \{\boldsymbol{w} \in \Re^{10} \mid \boldsymbol{U}\cdot\boldsymbol{w} - \boldsymbol{u} \geq \boldsymbol{o}\}$. Due to the special structure of the matrix $\boldsymbol{U}$ and $\boldsymbol{U}^{-1}$ the components w_1 and w_8 are not influenced by the transformation $Z^* \rightarrow \boldsymbol{U}(Z)$. In these two coordinate directions the boundary of the feasible set Z depends not on the other components of $\boldsymbol{w}$. Therefore the boundary ∂Z is parallel to the w_1-coordinate axis and the w_8-coordinate axis. In both cases the feasible set is marked as a grey area and starts in w_1-direction at $-1+\varepsilon$ and in w_8-direction at $+\varepsilon$, see Fig. 7a and Fig. 7h. The effect of the projection mechanism can be recognized very clearly by the time evolution of the component w_1 depicted in Fig. 7a. In the unconstrained case the component w_1 leaves the feasible set Z for a certain time period whereas in the constrained case due to the projection mechanism the component w_1 slides along the boundary before it returns into the feasible set. Both in the unconstrained and in the constrained case the equilibrium points $\boldsymbol{w}_\infty(\varepsilon_{\boldsymbol{w}}^*)$ are equivalent. In contrast to the

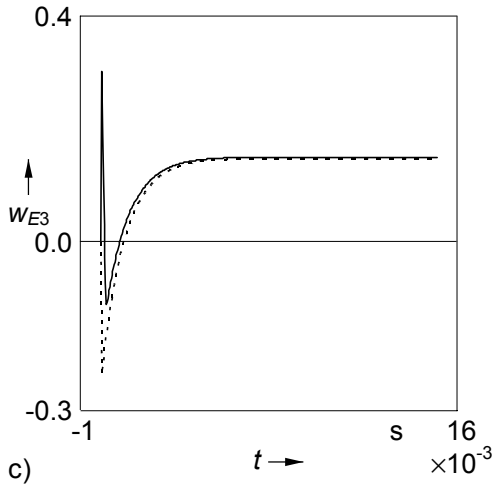
constrained case there exists time periods in the unconstrained case where the compensator does not exists.



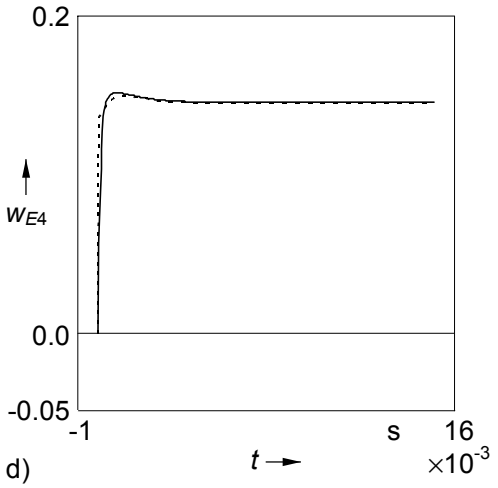
a)



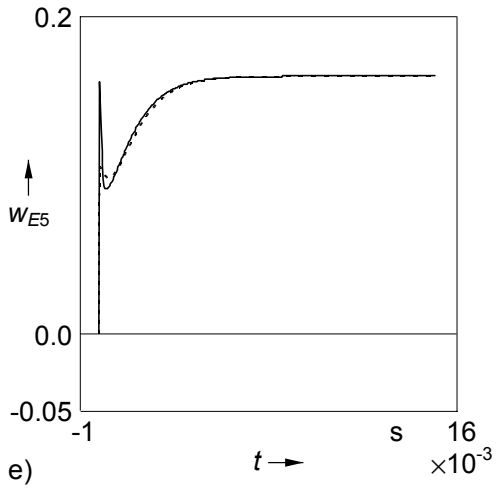
b)



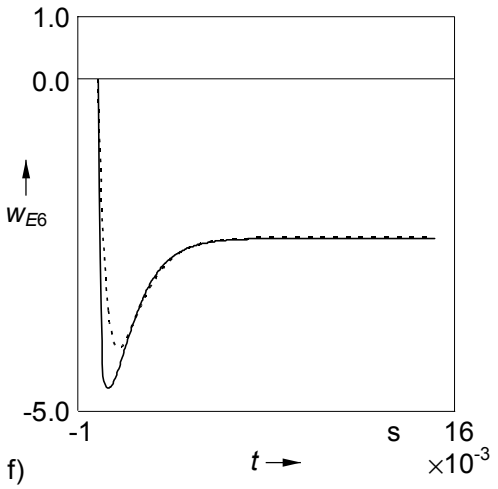
c)



d)



e)



f)

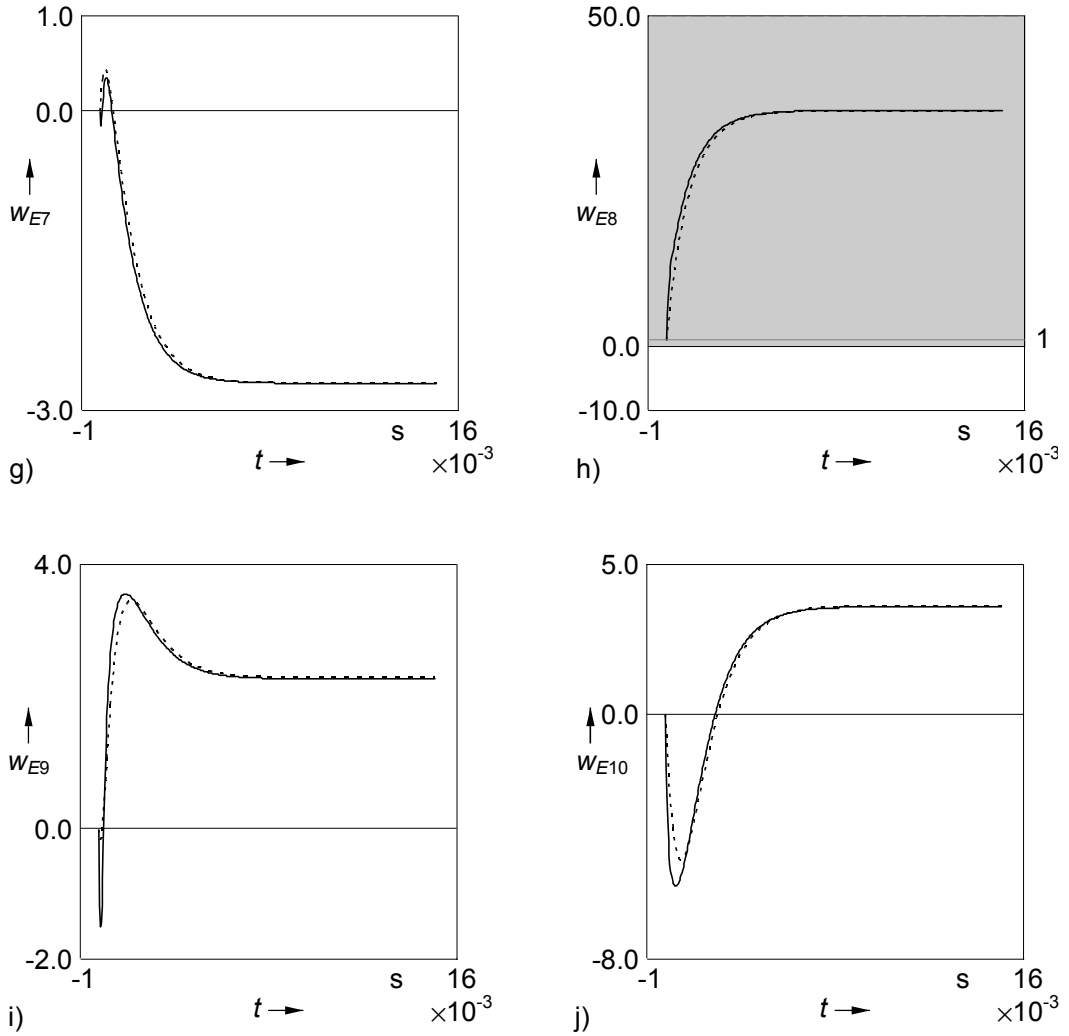


Figure 7: Time evolution of the vector $\boldsymbol{w}$

But the existence and uniqueness at every time point of the identification is a basic prediction for every iterative self-learning or adaptive inverse feedforward control scheme. It is exactly this property which was guaranteed by the projection mechanism in (68).

7. Conclusion

In control theory many identification problems with parameters which are subject to specific restrictions can be stated as convex programming problems. These problems can be solved on-line by the time-integration of a special time-invariant projected dynamical system with a discontinuous right-hand side. The paper develops an alternative formulation of this projected dynamical system based on a multidimensional stop operator. In contrast to the original formulation the new right-hand side is continuous and the problem is thus accessible to conventional analysis methods which easily give results on existence, uniqueness and stability properties of the corresponding solution trajectories. In future works the presented on-line identification method will be used as a part of an iterative compensation scheme for memoryless and complex hysteretic nonlinearities.

Acknowledgement

The authors thank Prof. Dr.-Ing. habil. Hartmut Janocha from the Laboratory for Process Automation (LPA) at Saarland University for the support of this project.

References

1. Brokate, M. and Sprekels, J., Hysteresis and phase transitions. Appl. Math Sci. 121, Springer, New York, 1996.
2. Burg, K.; Haf, H.; Wille, F.: *Höhere Mathematik für Ingenieure Band I: Analysis*. Teubner-Verlag, Stuttgart, 1997.
3. Drabek, P., Krejci, P. and Takac, P., Nonlinear differential equations. Research Notes in Mathmatics, Vol. 404, Chapman & Hall/CRC, London, 1999.
4. Ioannou, P.A. and Sun, J., Robust adaptive control. PTR Prentice-Hall, Upper Saddle River, 1996.
5. Krasnosel'skii, M.A. and Pokrovskii, A.V., Systems with hysteresis. Springer, Berlin, 1989.
6. Krejci, P., Hysteresis, convexity and dissipation in hyperbolic equations. Gakuto Int. Ser. Math. Sci. Appl., Vol. 8, Gakkotosho, Tokyo, 1996.
7. Kuhnen, K., Inverse Steuerung piezoelektrischer Aktoren mit Hysterese-, Kriech- und Superpositionsoperatoren. Doctoral Thesis, Saarland University, Shaker, Aachen, 2001.
8. Nagurney, A and Zhang, D., Projected dynamical systems and variational inequalities with applications. Kluwer, Boston/Dortrecht/London, 1996.
9. Papageorgiou, M., Optimierung. Oldenbourg, München, 1991.