WA17 11:40

Proceedings of the 34th
Conference on Decision & Control
New Orleans, LA - December 1995

Adaptive Poleplacement: The Division By Zero Problem

Krzysztof Arent[1] , Iven Mareels[2], Jan Willem Polderman[3]

Abstract— We re-examine the division by zero problem which occurs in certainty equivalence based indirect adaptive control algorithms applied to linear systems. By exploiting a parametrization for linear systems induced by the continued fraction description of its transfer function, the division by zero problem obtains a very simple geometric representation that can be used to virtually eliminate the problem in the adaptive algorithm.

Keywords: Adaptive pole placement control, pole-zero cancellation, continued fraction representation

Mathematics subject classification: 93C40

I. Introduction

This paper introduces a novel approach to the pole-zero cancellation problem encountered in adaptive indirect pole placement control applied to linear finite dimensional discrete time systems [5]. This work exploits ideas of [1] and [11].

Essentially we propose to exploit a reparametrization of the coefficients of the polynomials that specify the linear system to be controlled based on the continued fraction expansion of the system's transfer function. The important property of this system representation, from the point of view of adaptive pole placement control, is that verifying controllability, expressed in the new model parameters, is equivalent to checking whether some parameters are zero or not. This simple structure of uncontrollable models in the new parameter space is useful to derive a systematic solution of the stabilizability problem in adaptive control. The price paid for this simplicity is that the parametrization is nonlinear.

In essence it suffices to use in parallel a number of parameter identification algorithm in conjunction with a switching rule [11] to decide which parameter identification algorithm produces the estimate for control purposes. Despite the fact that the parametrization is nonlinear we establish that the classical gradient descent based identification algorithms may be used.

[1]Department of Applied Mathematics, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands, e-mail: k.arent@math.utwente.nl (corresponding author)

[2]Department of Engineering Australian National University, ACT 0200, Australia, e-mail: iven.mareels@anu.edu.au

[3]Department of Applied Mathematics, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands, e-mail: j.w.polderman@math.utwente.nl

In this contribution we focus on the system parametrization rather than its application in an adaptive control algorithm. We use tools from the behavioral approach to system theory [15] to expound our ideas.

The paper is organised as follows. In the next section we introduce the system parametrization and explore some of its properties. In Section 3 we describe briefly how this parametrization could be exploited in an adaptive control context. Some conclusion are provided in Section 4. For a more complete description we refer to [6].

II. System representation via continued fraction

Consider a dynamical system represented as:

$$A(\sigma)y = B(\sigma)u \tag{1}$$

where u is the system input, y is the measurable output, σ is the shift operator, for any integer t_0, $(\sigma^{t_0}w)(t) = w(t+t_0)$. $A(\xi)$, $B(\xi) \in \Re[\xi]$ are polynomials:

$$\begin{aligned} A(\xi) &= \xi^n + \cdots + a_1\xi + a_0 \\ B(\xi) &= b_m\xi^m + \cdots + b_1\xi + b_0 \quad n > m \end{aligned} \tag{2}$$

The coefficients of $A(\xi)$, $B(\xi)$ can be reparameterized as follows:

Definition II.1 *Continued fraction parameterization* Consider $A(\xi)$, $B(\xi) \in \Re[\xi]$ as in the system description (2). Let $k \in N$, $r_i \in \Re$ and $\Gamma_i(\xi) \in \Re[\xi]$ in the form

$$\Gamma_i(\xi) = \xi^{d_i} + \gamma_{i,d_i-1}\xi^{d_i-1} + \cdots + \gamma_{i,0} \quad d_i > 0 \tag{3}$$

be such that:

$$A(\xi) = R_0(\xi) \quad B(\xi) = r_1 R_1(\xi) \tag{4}$$

$$R_i(\xi) = \Gamma_{i+1}(\xi)R_{i+1}(\xi) + r_{i+2}R_{i+2}(\xi) \tag{5}$$

$$R_k(\xi) = 1 \quad R_{k-1}(\xi) = \Gamma_k(\xi) \tag{6}$$

The coefficients $\gamma_{i,j}$ and r_i are referred to as *the continued fraction parameters* for the pair of polynomials $A(\xi)$, $B(\xi)$. □

Remark II.2 If $A(\xi)$ and $B(\xi)$ are coprime then $\Gamma_1(\xi),\ldots,\Gamma_k(\xi)$, $r_1,\ldots,r_k$ can be obtained in the process of the continued fraction expansion of $A(\xi)/B'(\xi)$, where $B'(\xi) = B(\xi)/b_0$. This is where the notion *continued fraction* parameterization comes from. The above algorithmic definition of the continued fraction parameters follows from the Euclidean algorithm. □

Remark II.3 The importance of the continued fraction parameters (II.1) for the system (1) stems from the observation that the system is controllable iff all of the r_i continued fraction parameters are non-zero. This fact will be established in the sequel. □

A. Controllability issues

The continued fraction parameters can be used to represent the systems' behavior as follows:

Theorem II.4: Let $A(\xi)$, $B(\xi)$, $\Gamma_1(\xi)$, ..., $\Gamma_k(\xi)$ $r_1,\ldots,r_k$ be as in Definition II.1. Consider the dynamical system

$$\begin{bmatrix} r_1 & -\Gamma_1(\sigma) & \cdots & 0 \\ 0 & r_2 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & -\Gamma_k(\sigma) \end{bmatrix} \begin{bmatrix} u \\ y \\ e_1 \\ \vdots \\ e_{k-1} \end{bmatrix} = 0 \quad (7)$$

The system description (7) represents the same external behavior (in terms of the variables u and y) as the system description (1) [15], [14].

Remark II.5 The variables $e_1,\ldots,e_{k-1}$ are referred to as auxiliary variables.

Proof: The proof is by induction on the auxiliary variables e. We will show that after elimination of $e_{k-1},\ldots,e_l,e_{l+1}$ the system (7) has the following representation:

$$\begin{bmatrix} r_1 & -\Gamma_1(\sigma) & \cdots & 0 \\ 0 & r_2 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & -\Gamma_l(\sigma) & -1 \\ 0 & \cdots & r_{l+1}R_{l+1}(\sigma) & -R_l(\sigma) \end{bmatrix} \begin{bmatrix} u \\ y \\ e_1 \\ \vdots \\ e_l \end{bmatrix} = 0 \quad (8)$$

where $R_i(\xi)$ is defined by (5, 6).

First consider $u, y, e_1, \ldots, e_{k-2}$ as manifest variables and e_{k-1} as a latent variable ([14], [15]). In this first step we eliminate the latent variable e_{k-1}. The procedure to eliminate the other auxiliary variables is completely analogous.

The system representation (7) may alternatively be written as

$$\begin{bmatrix} r_1 & -\Gamma_1(\sigma) & \cdots & 0 \\ 0 & r_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & -\Gamma_{k-1}(\sigma) \\ 0 & \cdots & \cdots & r_k \end{bmatrix} \begin{bmatrix} u \\ y \\ e_1 \\ \vdots \\ e_{k-2} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ \Gamma_k(\sigma) \end{bmatrix} e_{k-1} \quad (9)$$

Define the matrix $U_k(\xi)$ in the following way:

$$U_k(\xi) = \begin{bmatrix} I_{k-2} & 0 & 0 \\ 0 & \Gamma_k(\xi) & -1 \\ 0 & 1 & 0 \end{bmatrix} \quad (10)$$

$U_k(\xi)$ is a unimodular matrix. Left multiplication of (9) by $U_k(\sigma)$ yields an equivalent representation [14], [15]:

$$\begin{bmatrix} r_1 & -\Gamma_1(\sigma) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & r_{k-1}\Gamma_k(\sigma) & -\Gamma_k(\sigma)\Gamma_{k-1}(\sigma) - r_k \\ 0 & \cdots & r_{k-1} & -\Gamma_{k-1}(\sigma) \end{bmatrix} \times \begin{bmatrix} u \\ y \\ e_1 \\ \vdots \\ e_{k-2} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{bmatrix} e_{k-1}$$

According to [14], the behavior in terms of $(u, y, e_1, \ldots, e_{k-2})$ is described by those rows in the left hand side of the above expression that corresponds to zeroes in the right hand side. This in conjunction with (6) yields (8) with $l = k-2$.

Along the same lines we can show that (8) holds for all $l = 0,\ldots,k-1$. It follows that (8) is input/output equivalent to (7) for $l = 0,\ldots,k-1$. The description (8) implies that elimination of all variables e_i yields

$$(\Gamma_1(\sigma)R_1(\sigma) + r_2R_2(\sigma))y = r_1R_1(\sigma)u \quad (11)$$

The result now follows from (5) with $i = 0$. □

We now investigate the controllability properties of the behaviors described by (1) and (7). We start with the system description (7).

Lemma II.6: Consider system (7). Assume that Γ_i and r_i are as in Definition II.1. Let $\mathfrak{B} := \{(u, y, e_1, \ldots, e_{k-1}) \mid \text{s.t. } (7) \text{ holds}\}$. $\mathfrak{B}$ is controllable iff $r_i \neq 0\ \forall i$.

Proof: Directly from the Hautus test, see also [6], [14], [15]. □

Next we observe that the introduction/elimination of the auxiliary variables has no influence on the controllability.

Lemma II.7: Consider the behavior $\mathfrak{B}$ as in Lemma II.6. For $l = 1, \dots k-1$, let

$$\begin{aligned}\mathfrak{B}_l &= \{(u, y, e_1, \dots, e_l) | \\ &\quad \exists e_{l+1}, \dots, e_{k-1} \\ &\quad \text{s.t.} (u, y, e_1, \dots, e_{k-1}) \in \mathfrak{B}\}\end{aligned}$$

Let

$$\begin{aligned}\mathfrak{B}_0 &= \{(u, y) | \exists e_1, \dots, e_{k-1} \\ &\quad \text{s.t. } (u, y, e_1, \dots, e_{k-1}) \in \mathfrak{B}\}\end{aligned}$$

$\mathfrak{B}_l$ is controllable iff $\mathfrak{B}_{l-1}$ is controllable, for all $l = 1, 2, \dots, k-1$.

Proof: Consider the proof of Theorem II.4. It follows that $\mathfrak{B}_l$ is parameterized by (8). Denote by $M_l(\xi)$ the matrix in (8). Equations (5, 6)) imply that there exists a unimodular matrix $U_l(\xi)$ such that

$$U_l(\xi) M_l(\xi) = \begin{bmatrix} M_{l-1}(\xi) & 0 \\ * & 1 \end{bmatrix} \tag{12}$$

Hence it follows that

$$\text{rank} M_l(\lambda) = \text{rank} M_{l-1}(\lambda) + 1 \text{ for all } \lambda \in C$$

The result is now a direct consequence of the generalized Hautus test for controllability. □

We summarize as follows:

Theorem II.8: Consider the dynamical system (1) with continued fraction parameters as given in Definition (II.1). The system is controllable iff $r_i \neq 0 \ \forall i$.

B. Representation issues

The link between the coefficients in the polynomials A and B and their continued fraction parameters defines a non trivial parameter transformation. In order to be able to construct an identification algorithm for the continued fraction parameters it is essential to understand this transformation in some detail. First we provide an example, then the general theory.

Example II.9 Consider a pair of polynomials

$$A(\xi) = \xi^2 + a_1 \xi + a_0 \quad B(\xi) = b_1 \xi + b_0 \tag{13}$$

where $a_i, \ b_i \in \Re$.

The continued fraction parameterization of $A(\xi)$, $B(\xi)$ is as follows:

$$\begin{aligned}A(\xi) &= \xi^2 + (\gamma_{1,0} + \gamma_{2,0})\xi + (\gamma_{1,0}\gamma_{2,0} + r_2) \\ B(\xi) &= r_1 \xi + r_1 \gamma_{2,0} \\ &\quad \text{for } b_1 \neq 0 \\ &\quad \text{or} \\ A(\xi) &= \xi^2 + \gamma_{1,1}\xi + \gamma_{1,0} \\ B(\xi) &= r_1 u \\ &\quad \text{for } b_1 = 0\end{aligned} \tag{14}$$

Let Ω in the parameter space $\{a_1, a_0, b_1, b_0\}$ represent polynomials with common factors. Define the corresponding set in the parameter space $\{\gamma_{1,0}, \gamma_{2,0}, r_1, r_2\}$ as Ω^{cfp}.

$$\begin{aligned}\Omega &= \{(a_1, a_0, b_1, b_0) \mid \\ &\quad b_0 = \tfrac{1}{2} b_1 (a_1 + \sqrt{a_1^2 - 4a_0}\ ) \\ &\quad \text{or} \\ &\quad b_0 = \tfrac{1}{2} b_1 (a_1 - \sqrt{a_1^2 - 4a_0}\ )\} \\ \Omega^{\text{cfp}} &= \{(\gamma_{1,0}, \gamma_{2,0}, r_1, r_2) \mid r_1 = 0 \text{ or } r_2 = 0\}\end{aligned}$$

Ω^{cfp} has a simple structure (as proven above in Theorem II.8). This is the main reason to introduce this parametrization.

The second interesting property of the continued fraction representation is that the parameters of the polynomials (14) are uniquely determined by the continued fraction coefficients of (13) and vice versa.

In particular, if $b_1 \neq 0$ we have

$$\begin{array}{ll} a_1 = -\gamma_{1,0} - \gamma_{2,0} & \gamma_{1,0} = -a_1 + \dfrac{b_0}{b_1} \\ a_0 = \gamma_{1,0}\gamma_{2,0} + r_2 & \gamma_{2,0} = -\dfrac{b_0}{b_1} \\ b_1 = r_1 & r_1 = b_1 \\ b_0 = -r_1 \gamma_{2,0} & r_2 = (\dfrac{b_0}{b_1})^2 - a_1(\dfrac{b_0}{b_1}) + a_0 \end{array} \tag{15}$$

If $b_1 = 0$ we have

$$a_1 = \gamma_{1,1} \quad a_0 = \gamma_{1,0} \quad b_0 = r_0 \tag{16}$$

The above calculations (15,16) show clearly that to parameterize the polynomials (13) two sets of continued fraction parameters, one for $b_1 \neq 0$ and another for $b_1 = 0$ are required. This complicates significantly the identification aspect of the continued fraction parameters! □

The following Lemma II.10 relates the parametrizations in the polynomial framework in general.

Lemma II.10: Let

$$\mathcal{PC}_n^{\text{cfp}} = \{ (\Gamma_1(\xi),\ldots,\Gamma_k(\xi),r_1,\ldots,r_k) | \; \Gamma_i(\xi) \in \Re[\xi] \text{ is monic }, k=1,\ldots,n; \; \deg \Gamma_i(\xi) \geq 1 \; \forall i, \textstyle\sum_{i=1}^k \deg \Gamma_i(\xi) = n \; r_i \neq 0 \text{ for } i=1,\ldots,k-1 \}$$

and

$$\mathcal{PC}_n = \{ (A(\xi),B(\xi)) | A(\xi) \in \Re[\xi] \; A(\xi) \text{ is monic, } \deg A(\xi) = n \; B(\xi) \in \Re[\xi], \deg B(\xi) < n \}$$

The equations (4,5,6) define a bijective mapping $h: \mathcal{PC}_n^{\text{cfp}} \to \mathcal{PC}_n$.

Proof: The fact that $h(\mathcal{PC}_n^{\text{cfp}}) \subset \mathcal{PC}_n$ follows from (6, 5, 4) and from the fact that $\sum_{i=1}^k \deg \Gamma_i(\xi) = n$.

Step 1: h is injective.

Suppose that this is not the case. There exist $(\Gamma_1,\ldots,\Gamma_k, r_1,\ldots,r_k) \in \mathcal{PC}_n^{\text{cfp}}$, $(\Gamma'_1,\ldots,\Gamma'_{k'},r'_1,\ldots,r'_{k'}) \in \mathcal{PC}_n^{\text{cfp}}$ such that

$$\begin{aligned}(A(\xi),B(\xi)) &= h(\Gamma_1,\ldots,\Gamma_k,r_1,\ldots,r_k) \\ &= h(\Gamma'_1,\ldots,\Gamma'_{k'},r'_1,\ldots,r'_{k'})\end{aligned} \tag{17}$$

Consider the sequences $\{R_i\}_{i=0}^k$ and $\{R'_i(\xi)\}_{i=0}^{k'}$ corresponding to $(\ldots,\Gamma_i(\xi),\ldots,r_i,\ldots)$ and $(\ldots,\Gamma'_i(\xi),\ldots,r'_i,\ldots)$. It follows from (6, 5) and from the fact: $\deg \Gamma_i(\xi) \geq 1 \; \forall i$ that

$$\begin{aligned}&R_i(\xi),\; R'_i(\xi) \text{ are monic} \\ &\deg R_i(\xi) > \deg R_{i+1}(\xi) \\ &\deg R'_i(\xi) > \deg R'_{i+1}(\xi)\end{aligned} \tag{18}$$

Equations (17) and definitions (6, 5, 4) imply

$$\begin{aligned}B(\xi) &= r_1 R_1(\xi) = r'_1 R'_1(\xi) \\ A(\xi) &= \Gamma_1(\xi)R_1(\xi) + r_2 R_2(\xi) \\ &= \Gamma'_1(\xi)R'_1(\xi) + r'_2 R'_2(\xi)\end{aligned} \tag{19}$$

It follows from the equations (18,19) that

$$\begin{aligned}&r_1 = r'_1, \quad R_1(\xi) = R'_1(\xi) \\ &\Gamma_1(\xi) = \Gamma'_1(\xi), \quad r_2 = r'_2, \quad R_2(\xi) = R'_2(\xi)\end{aligned} \tag{20}$$

Next consider

$$\begin{aligned}R_i &= \Gamma_{i+1}(\xi)R_{i+1}(\xi) + r_{i+2}R_{i+2}(\xi) \\ R_i &= \Gamma'_{i+1}(\xi)R_{i+1}(\xi) + r'_{i+2}R'_{i+2}(\xi)\end{aligned} \tag{21}$$

It follows from (19),(20) that $(\ldots,\Gamma_i(\xi),\ldots,r_i,\ldots)$, $(\ldots,\Gamma'_i(\xi),\ldots,r'_i,\ldots)$ satisfy (21) for $i=0$.

Consider the case $i > 0$. Observe that because of (18) the pair of equations (21) has only one solution, $\Gamma_{i+1}(\xi) = \Gamma'_{i+1}(\xi)$, $R_{i+2}(\xi) = R'_{i+2}(\xi)$, $r_{i+2} = r'_{i+2}$. Hence h is injective.

Step 2: h is surjective.

Take any $(A(\xi),B(\xi)) \in \mathcal{PC}_n$ and perform (6, 5, 4) backward such that (18) holds. It is easy to see that $(\Gamma_1,\ldots,\Gamma_k, r_1,\ldots,r_k) \in \mathcal{PC}_n^{\text{cfp}}$ and $h(\Gamma_1,\ldots,\Gamma_k,r_1,\ldots,r_k) = (A(\xi),B(\xi))$ □

Lemma II.11: Let

$$\mathcal{PC}^{\text{cfp}}_{n,k,\{d_1,\ldots,d_k\}} = \{(\Gamma_1(\xi),\ldots,\Gamma_k(\xi),r_1,\ldots,r_k) | \; \deg \Gamma_i(\xi) = d_i \; \text{for } i=1,\ldots,k \}$$

and $\mathcal{PC}_{n,k,\{d_1,\ldots,d_k\}} := h(\mathcal{PC}^{\text{cfp}}_{n,k,\{d_1,\ldots,d_k\}})$

The families $\mathcal{PC}^{\text{cfp}}_{n,k,\{d_1,\ldots,d_k\}}$ and $\mathcal{PC}_{n,k,\{d_1,\ldots,d_k\}}$ define partitions of $\mathcal{PC}_n^{\text{cfp}}$ and $\mathcal{PC}_n$ respectively.

Proof: The fact that the family $\mathcal{PC}^{\text{cfp}}_{n,k,\{d_1,\ldots,d_k\}}$ is a partition of $\mathcal{PC}_n^{\text{cfp}}$ is obvious. Because h is bijective (see Lemma II.10), the family $\mathcal{PC}_{n,k,\{d_1,\ldots,d_k\}}$ is a partition of $\mathcal{PC}_n$. □

This allows the following characterization of the parameter transformation:

Theorem II.12: Consider the transformation h defined by (4,5,6)

1. h induces unique bijective transformations $t_{n,k,\{d_1,\ldots,d_k\}}$ from $\mathcal{PC}^{\text{cfp}}_{n,k,\{d_1,\ldots,d_k\}}$ to $(A(\xi),B(\xi)) \in \mathcal{PC}_{n,k,\{d_1,\ldots,d_k\}}$.
2. $t_{n,k,\{d_1,\ldots,d_k\}}$ is a multilinear map.
3. If $d_i \neq d'_i$ for some i or $k \neq k'$ then $\text{Im}\, t_{n,k,\{d_1,\ldots,d_k\}} \cap \text{Im}\, t_{n,k',\{d'_1,\ldots,d'_{k'}\}} = \emptyset$.

Proof:

1. Follows from the fact that $\mathcal{PC}^{\text{cfp}}_{n,k,\{d_1,\ldots,d_k\}}$ is isomorfic to $\Re^n \times \Re^{k-1} \setminus \{0\} \times \Re$ and from the bijectivity of h (see Lemma II.10).
2. Follows immediately from the Definition II.1, especially (4, 5, 6).
3. Follows from Lemma II.10, Lemma II.11 and the first part of Theorem II.12.

□

The Lemma II.11 and Theorem II.12 imply that in the context of identifying a single system many distinct continued fraction parametrization need to be considered. The natural question is what is the number of such transformations.

Lemma II.13: Define $\Upsilon_{n,k} := \sharp\{(d_1,\ldots,d_k) \mid d_i \in N,\ \sum_{i=1}^{k} d_i = n\}$ and
$\Upsilon_n := \sharp\{(d_1,\ldots,d_k) \mid d_i \in N,\ \sum_{i=1}^{k} d_i = n;\ k = 1,\ldots,n\}$ where $\sharp$ stands for the number elements of a set. Then $\Upsilon_{n,k} = \sum_{l=k-1}^{n-1} \Upsilon_{l,k-1}$ with $\Upsilon_{n,1} = 1\ \forall n$ and $\Upsilon_n = \sum_{k=1}^{n} \Upsilon_{n,k}$

In the context of Lemma II.11 and Theorem II.12 Υ_n reflects the number of all parameter transformations.

Observe that if the set $\{(d_1,\ldots,d_k) \mid d_i \in N,\ \sum_{i=1}^{k} d_i = n\}$ consisted of non-ordered k-tuples then $\Upsilon_{n,k}$ would be a partition number [3], [4].

One can evaluate Υ_n as an explicit algorithm of n. It can be shown that Υ_n grows exponentially fast with n.

Remark II.14 Observe that each element in the set $\mathcal{PC}_{n,k,\{d_1,\ldots,d_k\}}$ is parameterized by n variables γ and k variables r. Moreover, the corresponding polynomials $(A(\xi),B(\xi)) \in \mathcal{PC}_{n,k,\{d_1,\ldots,d_k\}}$ have relative degree $(\deg(A)-\deg(B))$ d_1 and the number of coefficients is $2n-d_1$. It follows that if $k < n-d_1$ then $\mathcal{PC}_{n,k,\{d_1,\ldots,d_k\}}$ is a measure zero subset of the set $\mathcal{PC}_n$.

Remark II.15 If the system (1) has $\deg(a) = n$ and $\deg(B) = m$ we could restrict ourselves further to the single class of systems described by continued fraction parameters in $\mathcal{PC}_{n,n-m,\{m,1,\ldots,1\}}$. This way we only sacrifice a measure zero set of systems in the classical parametrization. If only the $\deg(A) = n$ of the system (1) is known, we can similarly restrict attention to the n parameter transformations $t_{n,n-m,\{n-m,1,\ldots,1\}}$

III. Identification of systems in continued fraction parameterization

In the previous section the continued fraction parameterization for the polynomial pair $(A(\xi),B(\xi))$ has been introduced and analysed. Here we concentrate on the question if the continued fraction parameters are identifiable from input output data. The fact that there is a bijective, be it complicated, mapping between the classical and continued fraction parametrizations of a system suggests that identification in continued fraction parameter space may be feasible.

A. The gradient algorithm

Consider the system (1) where $(A(\xi),B(\xi)) \in \mathcal{PC}_{n,k,\{d_1,\ldots,d_k\}}$. This system can be expressed in the following compact notation

$$y_t = \theta^T(\vartheta)\phi_{t-1}, \quad t \in N \tag{22}$$

where

$$\vartheta = [\gamma_{1,d_1-1}\cdots\gamma_{1,0}\cdots\gamma_{k,d_k-1}\cdots\gamma_{k,0}r_1\cdots r_k]^T$$

$$\theta(\vartheta) = [a_{n-1}(\vartheta)\ldots a_0(\vartheta)b_{n-d_1}(\vartheta)\ldots b_0(\vartheta)]^T$$

$$\phi_t = [y_t\cdots y_{t-n+1}u_{t-d_1}\cdots u_{t-n+1}]^T$$

The notation: $a_i(\vartheta)$, $b_j(\vartheta)$ means that a_i or b_j is expressed in terms of the elements from ϑ.

Denote by ϑ_0 the vector of parameters of the true system in continued fraction parameterization and by $\hat{\vartheta}_t$ the vector of estimates of ϑ_0 at time instant t. The mismatch between the true model and its estimate can be detected via the one step ahead prediction error e:

$$e_t = y_t - \theta^T(\hat{\vartheta}_{t-1})\phi_{t-1} \tag{23}$$

e_t can be used for improving the estimate of ϑ_0. One of the standard methods in adaptive control is to update the parameter estimates in the direction of the negative gradient of e_t^2 i.e.

$$\hat{\vartheta}_{t+1} = \hat{\vartheta}_t - \frac{\mu}{f(\hat{\vartheta}_t,\phi_t)}\frac{\partial e_{t+1}^2}{\partial\vartheta}\Big|_{\vartheta=\hat{\vartheta}_t} \tag{24}$$

where f is suitably chosen real, positive valued function and $\mu \in \Re_+$.

It follows from (23) that

$$\frac{\partial e_t^2}{\partial\vartheta} = -K(\hat{\vartheta})\phi_{t-1}e_{t-1} \tag{25}$$

where

$$K(\vartheta) = \frac{\partial\theta(\vartheta)}{\partial\vartheta} \tag{26}$$

The formulation (25) is interesting in the context of analysis of invariant points of (24).

Theorem III.1: The matrix $K(\vartheta)$ defined by (26) has full row rank iff $r_i \neq 0$ for all i

Proof: See [6] for details. □

In general we expect that e_t^2 is a non-convex function of ϑ with non-trivial local minima. Observe however that in the generic case, when (1) can be parameterized by elements from $\mathcal{PC}_{n,d,\{d,1,\ldots,1\}}$, the matrix $K(\vartheta)$ is square. If $r_i \neq 0\ \forall i$ then we conclude, on the ground of Theorem III.1 that $K(\vartheta)$ is invertible. This implies that if $\hat{r}_{i,t} \neq 0 \ \forall i, t$ and $\phi_t \neq 0 \ \forall t$ then ϑ^* is an invariant point of (24). if and only if $e_t(\vartheta^*) \equiv 0$!

This in conjunction with (25) implies that every local minimum of e_t^2 is in fact global.

Every invariant point of (24) specifies the (unique) system

$$(\sigma^n + a_{n-1}(\vartheta^*)\sigma^{n-1} + \cdots + a_0(\vartheta^*))y = (b_m(\vartheta^*)\sigma^m + \cdots + b_0(\vartheta^*))u \tag{27}$$

This equation necessarily explains a given data set (y, u). In [1] such a system is called an *unfalsified model.*

Unfalsified models have been carefully investigated in the context of adaptive pole placement control in [13]. It has been shown there that the pole placement controller based on an unfalsified model (27) behaves as if it was based on the true system. This goes a long way to certify that the above gradient based algorithm may be used to construct an adaptive algorithm that may overcome the pole-zero cancellation problem.

At present it is a conjecture that combining the continued fraction representation, with a gradient based algorithm complemented with a switching rule to guarantee that the r_i parameters remain of fixed sign indeed yields a working adaptive pole placement algorithm. Simulation evidence points this way [6].

IV. Conclusion

We introduced a novel parametrization for linear discrete time systems that can be used to aid in the solution of the pole-zero cancellation issue in pole placement adaptive control algorithms. Simulation evidence suggests that the identification and adaptation issues can be resolved. Further work is ongoing, see also [6].

References

[1] K.Arent, I.M.Y.Mareels , J.W.Polderman: The pole zero cancellation problem in adaptive control: a solution for minimum phase systems by approximate models. *Memorandum No. 1238, Dept. Appl. Math. University of Twente, The Netherlands, 1994.*

[2] R.G.Brocket Some Geometric Questions in the Theory of Linear Systems. *IEEE Trans. Aut. Contr.* AC-21, 449-455.

[3] K.L.Chung (1968) *A course in probability theory.* Harcourt, Brace & World, cop.

[4] G.H.Hardy and E.M.Wright (1960): *An introduction to the theory of numbers.* Oxford U.P.

[5] G.C.Goodwin and K.S.Sin (1984): *Adaptive Filtering Prediction and Control.* Prentice-Hall, Inc., Englewood Cliffs, N.J.

[6] K. Arent (1995) *Selected Topics in Adaptive Control,* PhD Thesis, Twente University.

[7] G.Kreisselmeier and M.C.Smith: Stable adaptive regulation of arbitrary n-th order plants. *IEEE Trans. Aut. Contr. AC-31*, 299-305, 1986.

[8] Ph. de Larminat: On the stabilizability condition in indirect adaptive control. *Automatica 20*, 793-795, 1984

[9] R.Lozano-Leal and G.C.Goodwin: A globally convergent adaptive pole placement algorithm without a persistence of excitation requirement, *IEEE Trans. Aut. Contr. AC-30*, 795-798, 1985.

[10] R.Lozano and X-I.Zhao: Adaptive Pole Placement Without Excitation Probing Signals *IEEE Trans. Aut. Contr. AC-39*, 47-58, 1994.

[11] F.M.Pait, S.A.Morse: A cyclic switching strategy for parameter adaptive control. *IEEE Trans. Aut. Contr. AC-39*, 1172-1183, 1994.

[12] J.W.Polderman: Avoiding the Non-Admissible Region of the Parameter Space in Indirect Adaptive Control Algorithms. *Proceedings of the Eighth International Conference on Analysis and Optimization of Systems.* Juan les Pins, France, 822-829, 1988

[13] J.W.Polderman: *Adaptive Control & Identification: Conflict or Conflux ?* -CWI Tract 67, Center for Mathematics and Computer Science, Amsterdam, The Netherlands, 1988.

[14] J.W.Polderman: Proper Elimination of Latent Variables. *Proceedings of 12th World Congress of Automatic Control* Sydney, Australia, 1993 Volume 10, 73-76.

[15] J.C. Willems (1988): Paradigms and Puzzles in the Theory of Dynamical Systems. *IEEE Trans. Aut. Contr.* AC-36, 259-294.