Singular Kalman Filtering: New Aspects Based on an Alternative System Theory

René Arnošt and Pavel Žampa *
Department of Cybernetics & Cybernetic Systems Research Center
University of West Bohemia
Univerzitní 8,
30614 Pilsen, Czech Republic

ABSTRACT
This paper is based on an alternative approach to system theory and deals with optimal filters for continuous linear dynamic systems with measurements disturbed by a "colored" noise or a "white" noise vector with a singular covariance matrix. It is shown that the optimal filter proposed in this paper is a slight modification of the Kalman-Bucy filter which can generally contain integrators as well as (backward) differentiators. The classical formulation was, however, reformulated in order to obtain a tractable mathematical interpretation of stochastic differential equations describing given processes and random errors.

KEY WORDS
Singular filtering, Kalman-Bucy filter, System theory, Causality law

1 Introduction

The linear (optimal) filtering problem was successfully treated in the past for discrete-time systems [6], [9] as well as for continuous-time systems [3], [7], [8], [18] technically as a part of all classical approaches to system theory. In particular, the work of R.E. Kalman [6], [7], [8] has been accepted by professional and scientific community. Despite an optimal filter for a general continuous-time linear dynamical system was explicitly proposed [8] in case that every measurement contains additive white noise (i.e. additive random disturbances with correlation times short compared to times of interest in the system), there may exist practical systems in which the correlation times of the random measurement errors are not short compared to times of interest in the system (so called colored noise). Also, some measurement may be so accurate that it is sometimes reasonable to assume they are perfect (i.e. they contain no errors). Both of these cases (no or colored noise) are singular problems within the framework of the Kalman-Bucy filtering theory and they need a special treatment. Therefore, we need a solid background of a system theory that is in a good agreement with real observations and supported by modern scientific methods. With respect to the importance of optimal filtering, a significant effort was made on the singular problems in the past. Some of the proposed solutions are discussed in [12]. Other authors presented their work based on different approaches on the topic, e.g. [4], [5], [10], [11], [14], [15] and [16]. This paper deals with the singular filtering problem solution to which is proposed on the basis of an alternative, fundamentally new approach to system theory [22].

*This work was supported by Ministry of Education of The Czech Republic - Project No. MSM235200004.

2 System Theory

Although the general system theory is a natural basis of cybernetics [19], [20], [21], no theory has been generally accepted by professional community so far. The fact that there are strong contradictions with observations in known system theories is considered to be a main reason for it. This is also why some authors define a system rather by examples than by an exact definition and why others have abandoned the system approach completely and use ad hoc models with no claim to cope with general system properties at all. Cybernetics thus loses its firm foundation and become rather a set of service instructions than a serious theory.

A significant reason for unsatisfactory results of contemporary system theories can be seen in the fact that they do not discern directional and non-directional causation between system variables [22], [20]. The system can be meaningfully defined as an assemble of all interconnected subsystems including all signal sources. The recently submitted alternative system theory [22] is based on the assumption that

- the system is an abstract concept purpose of which is to model such a part of the real world that is not influenced by its surroundings
- the system processes are driven by an internal system mechanism governed by a newly viewed principle and law of causality
- the system behavior is generally stochastic and can be reduced to deterministic one only when stochastic properties can be really neglected

- the system contains only a finite number of variables taking on only a finite number of values. This assumption can be extended to infinite cases when some consistency conditions are fulfilled

- the system is composed of subsystems (corresponding to the systems in contemporary system theories) mutually interconnected only by directional connections, which can be canceled or restored by an external intervention.

These paradigms lead to axioms that define directly the discrete-time systems only. The continuous-time system is then defined as a limit of an suitable sequence of previously defined discrete-time systems for their time periods converging to zero. According to our knowledge of physics it is supposed that the system behavior is generally stochastic and may be approximated by deterministic behavior in justified cases.

3 Singular Kalman Filtering

As stated in the Introduction, there are two possible types of singular problems within the framework of the Kalman-Bucy filtering theory discussed in literature in the past. In the first case, some measurement may be so accurate that, from technical point of view, it is reasonable to assume they are perfect. Such measurements theoretically contain no noise (deterministic measurements) when the filtering problem is being formulated. Hence, the correlation matrix of the measurement noise is singular whilst the matter of deterministic systems existence is a part of the general system theory. Classical system theories enable to define such systems that contain measurement random errors with correlation times not short compared to times of interest in the system. This is the case of colored noise in some of the measurements; the noise can often be simulated by an auxiliary linear dynamic system with white noise inputs (sometimes called a shaping filter). With this approach the colored noise vector becomes part of an augmented state variables. From the systemic point of view within the alternative system theory, the stochastic causal system is to be modeled as a whole entire that primarily includes the shaping filter in the system model. Thus the shaping filter approach makes the (augmented) system appear as a stochastic causal system which contains perfect measurement. Hence, the correlation matrix of the measurement noise is again singular.

We shall discuss another aspect of the singularity of the measurement noise covariance matrix now. Apparently, when two or more measurements are disturbed by a linearly dependent combinations of the random errors then the covariance matrix of measurement noise appears as singular again. In such a case, the corresponding measurements are affected by the same disturbances (linear combinations of the "same" white noise), which is very unlikely from both the technical and theoretic point of view; obviously two or more gauges can never contain identical random disturbances. However, thanks to the linear dependency, this case of measurement can be transformed into the case that some measurements are perfect.

Subsequently, it follows from the above discussed, that the case of some theoretically perfect measurements is sufficient to be considered as the reason for the singular covariance matrix of the measurement noise. This is merely a consequence of the proper definition of the stochastic causal system within the framework of the alternative system theory.

Before we proceed to the problem formulation, we shall make some remarks on the known solutions mentioned in the Introduction. Some of the authors who were involved in the matter of singular filtering problem were interested in spectral factorization and z-transformation, see e.g. [15], [16]. However, should the authors of this paper be interested in an application of the alternative system theory to the singular filtering problem, they are more interested in solution derived in time domain. From the problem analyses given e.g. in [4] and [12], it follows that using continuous differentiators to produce ideal derivatives of perfect measurements, in order to obtain required estimations, is inevitable. Merely, this was rather intuitively postulated than based on a deeper analysis. Moreover, some scholars are reluctant to the use of differentiators from the reasons for their practical inapplicability. The goal of this paper is to present a solution (based on the alternative system theory with the application of a continualization method – see [1]) that should clarify the need for differentiators. As we shall see later, the presented method can lead to both derivative and integrative character of the optimal filter, see example in section 7.2.

4 Problem formulation

A standard problem of linear filtering is the following: it is required to estimate the state variables of a continuous process in the usual state-space form

$$\dot{x}(t) = Fx(t) + GW(t), \tag{1}$$

where $x(t)$ is the state vector of dimension n, F is a matrix of known time functions describing dynamics of the system, G is a square matrix, $W(t)$ is a white noise vector of dimension n. The observations are given by

$$z(t) = Hx(t) + JV(t), \tag{2}$$

where $z(t)$ is an $m-$dimensioned observation vector, H and J are matrices of appropriate dimensions, $V(t)$ is a white noise vector of dimension m. It is also assumed that the initial condition $x(t_0)$ is a Gaussian random variable of zero mean and known covariance matrix P_0 and that the noise characteristics are given as

$$E\{W(t)\}{=}0, \quad E\{GW(t)W^T(\tau)G^T\}{=}GG^T\delta(t-\tau),$$
$$E\{V(t)\}{=}0, \quad E\{JV(t)V^T(\tau)J^T\}{=}JJ^T\delta(t-\tau),$$

where δ is the Dirac delta function. If the matrix JJ^T is positive definite, the solution is simply a Kalman-Bucy filter [8]. If JJ^T is singular, the filtering problem becomes singular. The problem stated as above is not formulated in a precise mathematical way. Note, the state-variables and measurements are disturbed by additive white noise, i.e. random errors with infinite standard deviations. To obtain a tractable mathematical formulation of (1) and (2) we shall use the Itô interpretation of stochastic differential equations [13] and reformulate the filtering problem.

5 Problem reformulation

As discussed in [13] the Itô interpretation of (1) and (2) is

$$\mathrm{d}x(t) = Fx(t)\mathrm{d}t + G\mathrm{d}w(t), \tag{3}$$

$$\mathrm{d}y(t) = Hx(t)\mathrm{d}t + J\mathrm{d}v(t), \tag{4}$$

where $w(t)$ is $n-$dimensional Brownian motion independent of $x(t_0)$, $v(t)$ is $m-$dimensional Brownian motion independent of $w(t)$ and $x(t_0)$[1]. Note that only the signal $y(t)$ is measured in the equation (4). The reformulated problem is again singular if matrix JJ^T is singular.

The equations (3) and (4) can be generalized as a system that contains measurements,

$$\mathrm{d}s(t) = \boldsymbol{F}s(t)\mathrm{d}t + \boldsymbol{G}\mathrm{d}\boldsymbol{w}(t), \tag{5}$$

where $s(t) = [x^T(t), y^T(t)]^T$, matrices $\boldsymbol{F}$, $\boldsymbol{G}$ are obviously block-composed of matrices F, H and G, J respectively and $\boldsymbol{w}(t) = [w^T(t), v^T(t)]^T$ is a vector of fictive random errors[2]. Moreover, following the system paradigms of the alternative system theory [22], the estimation problem should be reformulated as the synthesis of an estimation system Σ (Figure 1) which is composed of a given subsystem $\Sigma^{(1)}$ and an optimal estimator $\Sigma^{(2)}$. The filtering problem then consists in finding an optimal estimator from a class of admissible subsystems $\Sigma^{(2)}$ so that the system Σ is well defined. $\Sigma^{(2)}$ produces required optimal estimations $\mu(t) = E\{x(t)|y(\tau)\}$, $\tau \leq t$ of the state vector $x(t)$ in the least square sence,

$$\mu^*(t) = \arg\min_{\mu(t)} E\{|x(t) - \mu(t)|^2\}. \tag{6}$$

Having found the mathematical formulation of the filtering problem, which coincides with the correct definition of the stochastic causal system within the alternative system theory, we can start to study the properties of the optimal estimation $\mu(t)$.

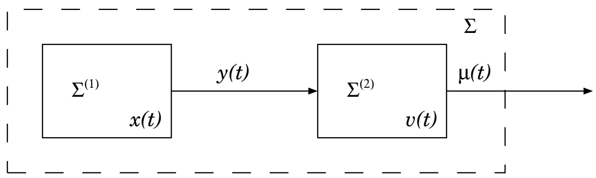

Figure 1. Estimation system

6 Problem solution

To be able to focus on the main ideas in the solution to the singular filtering problem, we will first consider 1-dimensional measurement case. The extension to multi-dimensional case of equation (4) is technical, but does not require any essentially new ideas.

From now on, we will assume with no loss in generality that the subsystem Σ_1 is given in the observability canonical form (there is always a regular transformation to the observability canonical form for every linear system). For a single output, the filtering problem is singular iff J in equation (4) is a zero row,

$$J = [0, 0, \cdots, 0]. \tag{7}$$

With respect to the fundamental ideas of the alternative system theory, the solution will be presented in the following steps:

6.1 Sequence of discrete subsystems $\Sigma_k^{(1)}$

At this stage we need to find an appropriate sequence of time-invariant discrete-time subsystems $\Sigma_k^{(1)}$ that will for $k \to \infty$ converge to the given time-invariant continuous-time subsystem $\Sigma^{(1)}$ described by equation (5), see e.g. [2] for details. Such a sequence can be described by

$$\begin{aligned} s_k(t+h_k) =& e^{\boldsymbol{F}h_k} s_k(t) + \int_0^{h_k} e^{\boldsymbol{F}(h_k-\tau)} \boldsymbol{G}\mathrm{d}\boldsymbol{w}(\tau) = \\ =& \widetilde{\boldsymbol{A}}_k s_k(t) + \widetilde{\boldsymbol{\Gamma}}_k \xi_k(t+h_k), \end{aligned} \tag{8}$$

with a generally stochastic initial condition $s_k(t_0)$ of the Gaussian (normal) distribution,

$$s_k(t_0) = s(t_0) \sim \mathcal{N}\{\boldsymbol{m}\ ,\ R\}, \tag{9}$$

$t,\ t + h_k \in T_k$, $h_k = \frac{h_{k-1}}{2}$, where $\widetilde{\boldsymbol{A}}_k = e^{\boldsymbol{F}h_k}$, $\widetilde{\boldsymbol{\Gamma}}_k \widetilde{\boldsymbol{\Gamma}}_k^T = \int_0^{h_k} e^{\boldsymbol{F}(h_k-\tau)} \boldsymbol{G} \cdot \boldsymbol{G}^T e^{\boldsymbol{F}(h_k-\tau)^T} \mathrm{d}\tau$ (Itô isometry), $\xi_k(t + h_k) = \frac{\Delta w(t+h_k)}{\sqrt{h_k}}$, $t + h_k \in T_k$ is a sequence

---

[1] A precise mathematical treatment is presented in [13]. There is also given an explanation of the need to transform measurement $z(t)$ into $y(t) = \int_0^t z(s)\mathrm{d}s$ with no loss or gain in information and thereby to obtain the stochastic integral representation of observation $y(t)$ in (4).

[2] The stochastic characteristics of stochastic causal systems are given by probability density functions or by cumulative distribution functions. Vectors of random variables do not represent any state variables and do not correspond to any attributes of a given system. They were introduced into system description to model stochastic characteristics within state-space models as part of transformation from probability description to state-space description of causal systems, see e.g. [1], [22]. Naturally, causal dependencies of systems or their properties cannot be changed only by different description of a given system.

of fictive[3] random, mutually independent variables (also independent of the initial condition $s_k(t_0)$) with normal distribution $\xi_k(t+h_k) \sim \mathcal{N}\{0\,,\,I\,\}$. For a very small h_k ($h_k = \frac{h_0}{2^k} \to 0$ for $k \to \infty$) we can approximate ($\widetilde{\boldsymbol{A}}_k \doteq \boldsymbol{A}_k, \widetilde{\boldsymbol{\Gamma}}_k \doteq \boldsymbol{\Gamma}_k$)

$$\boldsymbol{A}_k = I + \boldsymbol{F}\cdot h_k, \tag{10}$$

$$\boldsymbol{\Gamma}_k\boldsymbol{\Gamma}_k^T = \boldsymbol{G}\boldsymbol{G}^T\cdot h_k. \tag{11}$$

The state vector $s_k(t)$ naturally consists of two components $s_k(t) = [x_k^T(t), y_k^T(t)]^T$, where $x_k(t)$ is again of dimension n and $y_k(t)$ is a scalar (single) output. The sequence of discrete subsystems $\Sigma_k^{(1)}$ can then be rewritten for given

$$\boldsymbol{A}_k = \begin{bmatrix} A_k & B_k \\ C_k & D_k \end{bmatrix}, \quad \boldsymbol{\Gamma}_k = \begin{bmatrix} \Gamma_k & 0 \\ 0 & 0 \end{bmatrix}$$

as

$$x_k(t+h_k) = A_k x_k(t) + \Gamma_k\cdot\xi_k(t+h_k), \tag{12}$$

$$y_k(t+h_k) = C_k x_k(t) + D_k y_k(t), \tag{13}$$

where

$$A_k = I + F\cdot h_k, \quad D_k = 1, \quad B_k = 0,$$
$$C_k = H\cdot h_k = \begin{bmatrix} 0 & 0 & \cdots & 0 & h_k \end{bmatrix},$$
$$\boldsymbol{\Gamma}_k = \begin{bmatrix} \Gamma_k \\ 0 \end{bmatrix} = \begin{bmatrix} G\cdot\sqrt{h_k} \\ 0 \end{bmatrix}.$$

Remark: It is easy to show that the sequence of discrete-time systems $\Sigma_k^{(1)}$ does converge to the originally given continuous-time system $\Sigma^{(1)}$ for $k \to \infty$, see e.g. [2].

6.2 Synthesis of discrete estimation system

The synthesis of the discrete-time estimation system consists in the synthesis of a standard discrete Kalman filter for a given sequence of discrete subsystems $\Sigma_k^{(1)}$ in every step $k = 0, 1, 2, \cdots$ for each $h_k = \frac{h_{k-1}}{2}$, crucially for a very small time-period h_k. Thus, we assume that $\Sigma_k^{(1)}$ is given by equations (12) and (13). The equations of standard Kalman filter are given as

$$\widehat{x}_k(t+h_k) = A_k\mu_k(t),$$
$$\widehat{y}_k(t+h_k) = C_k\mu_k(t) + D_k y_k(t),$$
$$\mu_k(t) = \widehat{x}_k(t) + K_k(t)\cdot\Big(y_k(t) - \widehat{y}_k(t)\Big),$$

where $\mu_k(t) = E\{x_k(t)|y_k(0..t)\}$, $\widehat{x}_k(t+h_k) = E\{x_k(t+h_k)|y_k(0..t)\}$, $\widehat{y}_k(t+h_k) = E\{y_k(t+h_k)|y_k(0..t)\}$ with appropriate initial conditions, $t, t+h_k \in T_k$. The error covariance matrix is then, for the perfect measurement $y(t)$ from equation (13), given by

$$P_k(t+h_k) = A_k P_k(t) A_k^T + \Gamma_k\Gamma_k^T - \tag{14}$$
$$- \big(A_k P_k(t) C_k^T\big)\cdot\big(C_k P_k(t) C_k^T\big)^{-1}\cdot\big(C_k P_k(t) A_k^T\big),$$

[3]See footnote No. 2.

with an initial condition $P_k(t_0)$ and similarly the Kalman gain is given as

$$K_k(t+h_k) = \big(A_k P_k(t) C_k^T\big)\cdot\big(C_k P_k(t) C_k^T\big)^{-1} \tag{15}$$

for $t > 0$ and with an initial condition $K_k(t_0)$ for $t = t_0$. Taking an advantage of the observability canonical form of $\Sigma^{(1)}$ and the perfect measurement (13) it will be useful to study the optimal estimation of the $n-$the component of the state vector $x_k(t)$ of $\Sigma_k^{(1)}$; therefore let $x_k(t)$ be partitioned as

$$x_k(t) = \begin{bmatrix} x_k^{(1)}(t) \\ \vdots \\ x_k^{(N_x-1)}(t) \\ ----- \\ x_k^{(N_x)}(t) \end{bmatrix} = \begin{bmatrix} x_k^{(1\ldots N_x-1)}(t) \\ ----- \\ x_k^{(N_x)}(t) \end{bmatrix}. \tag{16}$$

Hence, we can also partition P_k, A_k, $\mu_k(t)$, C_k, $\Gamma_k\Gamma_k^T$ with corresponding dimensions,

$$P_k(t) = \begin{bmatrix} P_k^{(1,1)}(t) & P_k^{(1,2)}(t) \\ P_k^{(1,2)^T}(t) & P_k^{(2,2)}(t) \end{bmatrix}, A_k = \begin{bmatrix} A_k^{(1,1)} & A_k^{(1,2)} \\ A_k^{(2,1)} & A_k^{(2,2)} \end{bmatrix},$$

$$\mu_k(t) = \begin{bmatrix} \mu_k^{(1\ldots N_x-1)}(t) \\ \mu_k^{(N_x)}(t) \end{bmatrix}, C_k = \begin{bmatrix} 0 & \cdots & 0 & h_k \end{bmatrix},$$

$$\Gamma_k\Gamma_k^T = \begin{bmatrix} \Gamma_k^{(1)}\Gamma_k^{(1)^T} & \Gamma_k^{(1)}\Gamma_k^{(2)^T} \\ \Gamma_k^{(2)}\Gamma_k^{(1)^T} & \Gamma_k^{(2)}\Gamma_k^{(2)^T} \end{bmatrix} \text{ for } \Gamma_k = \begin{bmatrix} \Gamma_k^{(1)} \\ \Gamma_k^{(2)} \end{bmatrix}.$$

Consequently, the error covariance matrix in equation (14) can be rewritten as

$$P_k(t+h_k) = A_k\cdot S_k(t+h_k)\cdot A_k^T + \Gamma_k\Gamma_k^T, \tag{17}$$

where

$$S_k(t+h_k) = \begin{bmatrix} S_k^{(1,1)}(t+h_k) & 0 \\ 0 & 0 \end{bmatrix}$$

and

$$S_k^{(1,1)}(t+h_k) = P_k^{(1,1)}(t) - P_k^{(1,2)}(t) P_k^{(2,2)^{-1}}(t) P_k^{(1,2)^T}(t).$$

Hence, we can calculate the error covariance matrix

$$P_k(t) = \begin{bmatrix} A_k^{(1,1)} S_k^{(1,1)}(t) A_k^{(1,1)^T} + \Gamma_k^{(1)}\Gamma_k^{(1)^T}, & A_k^{(1,1)} S_k^{(1,1)}(t) A_k^{(2,1)^T} + \Gamma_k^{(1)}\Gamma_k^{(2)^T} \\ A_k^{(2,1)} S_k^{(1,1)}(t) A_k^{(1,1)^T} + \Gamma_k^{(2)}\Gamma_k^{(1)^T}, & A_k^{(2,1)} S_k^{(1,1)}(t) A_k^{(2,1)^T} + \Gamma_k^{(2)}\Gamma_k^{(2)^T} \end{bmatrix}$$

with respect to

$$S_k^{(1,1)}(t+h_k) = A_k^{(1,1)} S_k^{(1,1)}(t) A_k^{(1,1)^T} + \Gamma_k^{(1)}\Gamma_k^{(1)^T} -$$
$$- \Big(A_k^{(1,1)} S_k^{(1,1)}(t) A_k^{(2,1)^T} + \Gamma_k^{(1)}\Gamma_k^{(2)^T}\Big)\cdot$$
$$\cdot \Big(A_k^{(2,1)} S_k^{(1,1)}(t) A_k^{(2,1)^T} + \Gamma_k^{(2)}\Gamma_k^{(2)^T}\Big)^{-1}\cdot$$
$$\cdot \Big(A_k^{(2,1)} S_k^{(1,1)}(t) A_k^{(1,1)^T} + \Gamma_k^{(2)}\Gamma_k^{(1)^T}\Big)$$

where the initial condition $S_k^{(1,1)}(t_0)$ equals to the given $P_k^{(1,1)}(t_0)$.

Remark: The last equation is the discrete Riccati equation of the (n-1)th order.

Similarly, the Kalman gain can be computed as

$$K_k(t) = \begin{bmatrix} \widetilde{K}_k^{(1)}(t) \\ \widetilde{K}_k^{(2)}(t) \end{bmatrix} \cdot \frac{1}{h_k},$$

where

$$\widetilde{K}_k^{(1)}(t) = \\ = A_k^{(1,1)} \cdot \Big(A_k^{(1,1)} S_k^{(1,1)}(t-h_k) A_k^{(2,1)T} + \Gamma_k^{(1)} \Gamma_k^{(2)T} \Big) \cdot \\ \cdot \Big(A_k^{(2,1)} S_k^{(1,1)}(t-h_k) A_k^{(2,1)T} + \Gamma_k^{(2)} \Gamma_k^{(2)T} \Big)^{-1} + A_k^{(1,2)}$$

and

$$\widetilde{K}_k^{(2)}(t) = \\ = A_k^{(2,1)} \cdot \Big(A_k^{(1,1)} S_k^{(1,1)}(t-h_k) A_k^{(2,1)T} + \Gamma_k^{(1)} \Gamma_k^{(2)T} \Big) \cdot \\ \cdot \Big(A_k^{(2,1)} S_k^{(1,1)}(t-h_k) A_k^{(2,1)T} + \Gamma_k^{(2)} \Gamma_k^{(2)T} \Big)^{-1} + A_k^{(2,2)}.$$

The optimal discrete-time estimations are then given as

$$\mu_k^{(1\dots N_x-1)}(t) = \widehat{x}_k^{(1\dots N_x-1)}(t) + \\ + \widetilde{K}_k^{(1)}(t) \cdot \Big(\frac{y_k(t) - y_k(t-h_k)}{h_k} - H\mu_k(t-h_k) \Big), \\ \mu_k^{(N_x)}(t) = \widehat{x}_k^{(N_x)}(t) + \\ + \widetilde{K}_k^{(2)}(t) \cdot \Big(\frac{y_k(t) - y_k(t-h_k)}{h_k} - H\mu_k(t-h_k) \Big).$$

Having found the discrete-time solution, we can proceed to continuous-time filter now.

6.3 Continuous filter

To derive the continuous-time solution to the singular filtering problem, we need to follow the principles of the continualization method presented e.g. in [2].

The continuous error covariance matrix is given by an equation gained from its discrete version (17) by limiting for $h_k \to 0$,

$$P(t) = \lim_{k\to\infty} P_k(t+h_k) = \begin{bmatrix} P^{(1,1)}(t) & 0 \\ 0 & 0 \end{bmatrix},$$

differential equation of which is given by

$$\dot{P}(t) = \begin{bmatrix} \dot{P}^{(1,1)}(t) & 0 \\ 0 & 0 \end{bmatrix},$$

where

$$\dot{P}^{(1,1)}(t) = \lim_{k\to\infty} \frac{P_k^{(1,1)}(t+h_k) - P_k^{(1,1)}(t)}{h_k} = \\ = F^{(1,1)} P^{(1,1)}(t) + P^{(1,1)}(t) F^{(1,1)T} + G^{(1)} G^{(1)T} - \\ - \Big(P^{(1,1)}(t) F^{(2,1)T} + G^{(1)} G^{(2)T} \Big) \cdot \Big(G^{(2)} G^{(2)T} \Big)^{-1} \cdot \\ \cdot \Big(F^{(2,1)T} \cdot P^{(1,1)}(t) + G^{(2)} G^{(1)T} \Big)$$

and G was partitioned similarly to partition (16), $G = \begin{bmatrix} G^{(1)} \\ G^{(2)} \end{bmatrix}$. Subsequently, we can compute the Kalman gain of the continuous filter by the following limitation

$$\widetilde{K}(t) = \lim_{k\to\infty} \begin{bmatrix} \widetilde{K}_k^{(1)}(t) \\ \widetilde{K}_k^{(2)}(t) \end{bmatrix} = \\ = \begin{bmatrix} \Big(P^{(1,1)}(t) F^{(2,1)T} + G^{(1)} G^{(2)T} \Big) \cdot \Big(G^{(2)} G^{(2)T} \Big)^{-1} \\ 1 \end{bmatrix}.$$

Finally, the continuous optimal estimations of state vector $x(t)$ are given as

$$\mu^{(1\dots N_x-1)}(t) = \underset{k\to\infty}{\text{l.i.m.}}\, \mu_k^{(1\dots N_x-1)}(t) = \\ = \widehat{x}^{(1\dots N_x-1)}(t) + \widetilde{K}^{(1)}(t) \cdot \Big(\dot{y}(t-) - H\mu_k(t-) \Big), \tag{18}$$

$$\mu_k^{(N_x)}(t) = \underset{k\to\infty}{\text{l.i.m.}}\, \mu_k^{(N_x)}(t) = \dot{y}(t-), \tag{19}$$

where $\dot{y}(t-) = \frac{\mathrm{d}y(t)}{\mathrm{d}t^-} = \text{l.i.m.}_{k\to\infty} \frac{y_k(t) - y_k(t-h_k)}{h_k}$ (l.i.m. denotes the "limit in the mean").

Due to the perfect measurement $y(t)$ and the observability canonical form, estimation of the state $x^{(N_x)}(t)$ is precise ($P^{(2,2)}(t) = 0$) and given by a "backward" derivative of the measurement $y(t)$. Therefore, "backward" differentiators are needed as a part of the filter. Note, that this is possible only if the measurement is really perfect.

In the next section, we shall demonstrate advantages of the new approach to singular filtering problem on some examples.

7 Examples

7.1 Example 1

Consider a continuous subsystem $\Sigma^{(1)}$

$$\mathrm{d}s(t) = \mathbb{F} \cdot s(t) \cdot \mathrm{d}t + \mathbb{G} \cdot \mathrm{d}w(t), \tag{20}$$

with an initial condition $s(0) \sim \mathcal{N}\{\mathbf{m}\ ,\ R\}$, $s(t) \in \mathbb{R}^3$, $t \in T = \langle 0\ ,\ \vartheta \rangle$, $\mathrm{d}w(t)$ represents the Wiener Lèvy process, $\mathrm{E}\{\mathrm{d}w(t)\} = 0$, $\mathrm{E}\{\mathrm{d}w(t)^2\} = \mathrm{d}t$. The system is given in the observability canonical form with matrices

$$\mathbb{F} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \qquad \mathbb{G} = \begin{bmatrix} g \\ 0 \\ 0 \end{bmatrix}.$$

The state vector consists of unmeasurable variables $x(t) \in \mathbb{R}^2$ and a scalar measurement $y(t) \in \mathbb{R}$,

$$\begin{aligned} \mathrm{d}x^{(1)}(t) &= & & g \cdot \mathrm{d}w(t) \\ \mathrm{d}x^{(2)}(t) &= x^{(1)}(t) \cdot \mathrm{d}t & & \\ \mathrm{d}y(t) &= & x^{(2)}(t) \cdot \mathrm{d}t, & \end{aligned}$$

where the apparent matrices F, G, H (see e.g. equations (3), (4)) and their dimensions are given by partitioning $\mathbb{F}$ and $\mathbb{G}$ for the partitioned state vector $s(t) =$

$\left[x^T(t), y^T(t)\right]^T$. An appropriate sequence of discrete systems converging to the given continuous plant can be given as

$$\begin{aligned} x_k(t+h_k) &= A_k x_k(t) + \Gamma_k \xi_k(t+h_k) \\ y_k(t+h_k) &= C_k x_k(t) + D_k y_k(t), \end{aligned}$$

with initial conditions $s_k(0) = \left[x_k^T(0), y_k^T(0)\right]^T \sim \mathcal{N}\{\mathbf{m}\,,\, R\}$, where $\mathbb{A}_k = I + \mathbb{F}\cdot h_k$ and $\mathbf{\Gamma}_k = \sqrt{h_k}\cdot\mathbb{G}$; hence $A_k = I + F\cdot h_k$, $\Gamma_k = \sqrt{h_k}\cdot G$, $C_k = H\cdot h_k$, $D_k = 1$ and $\xi_k(t) \sim \mathcal{N}\{0\,,\, I\}$ (independent of the initial condition $s_k(0)$).

The steady solution to equation (14) for the covariance error matrix $P_k(t)$ is given as

$$P_k(t) = \begin{bmatrix} 2\cdot g^2\cdot h_k & g^2\cdot h_k^2 \\ g^2\cdot h_k^2 & g^2\cdot h_k^3 \end{bmatrix}$$

and the steady-mode Kalman gain from eq. (15) as

$$K_k(t) = \begin{bmatrix} \frac{1}{h_k^2} \\ \frac{2}{h_k} \end{bmatrix} = \begin{bmatrix} \widetilde{\widetilde{K}}_k^{(1)}(t)\cdot\frac{1}{h_k} \\ \widetilde{K}_k^{(2)}(t) \end{bmatrix}\cdot\frac{1}{h_k}.$$

The optimal discrete estimations are then given by formulas

$$\begin{aligned} \mu_k^{(1)}(t) &= \frac{1}{h_k}\left(\frac{y_k(t)-y_k(t-h_k)}{h_k} - \frac{y_k(t-h_k)-y_k(t-2h_k)}{h_k}\right), \\ \mu_k^{(2)}(t) &= h_k\cdot\mu_k^{(1)}(t) + \frac{y_k(t)-y_k(t-h_k)}{h_k}. \end{aligned}$$

The optimal continuous filter will be derived by taking $k\to\infty$ ($h_k = \frac{h_0}{2^k} \to 0$). Merely, the steady error covariance matrix of the continuous filter is given by the limit

$$P(t) = \lim_{k\to\infty} P_k(t) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Similarly, we obtain the steady-mode Kalman gain

$$K(t) = \lim_{k\to\infty} \begin{bmatrix} \widetilde{\widetilde{K}}_k^{(1)}(t) \\ \widetilde{K}_k^{(2)}(t) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

The optimal continuous estimations are given by limits (in the square mean) as follows

$$\begin{aligned} \mu^{(1)}(t) &= \underset{k\to\infty}{\text{l.i.m.}}\, \mu_k^{(1)}(t) = \frac{\mathrm{d}y^2(t)}{\mathrm{d}t^{-2}} \\ \mu^{(2)}(t) &= \underset{k\to\infty}{\text{l.i.m.}}\, \mu_k^{(2)}(t) = \frac{\mathrm{d}y(t)}{\mathrm{d}t^{-}}, \end{aligned}$$

which represent "backward" derivatives of the perfect measurement $y(t)$.

7.2 Example 2

This example deals with the classical formulation of the singular filtering problem, stated e.g. in paragraph 4 or in [4], [5], [12]. Consider a continuous system $\Sigma^{(1)}$ given by the linear differential equation

$$\dot{x}(t) = F\cdot x(t) + G\cdot\dot{w}(t) \tag{21}$$

with measurement

$$y(t) = H\cdot x(t) + J\cdot\dot{v}(t). \tag{22}$$

Suppose the system matrices are given as

$$F = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}, \qquad G = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix},$$
$$H = [1\,,\,0]\,, \qquad J = [\lambda_3]\,.$$

The filtering problem is singular for $\lambda_3 = 0$. An appropriate sequence of discrete systems can be described by

$$x_k(t+h_k) = A\cdot x(t) + \Gamma_k\cdot\xi_k(t+h_k)$$

with measurement

$$y_k(t) = C_k\cdot x_k(t) + \Delta_k\cdot\eta_k(t),$$

where $A_k = I + F\cdot h_k$, $\Gamma_k = \sqrt{h_k}\cdot G$, $C_k = H$, $\Delta_k = \lambda_3$ and $\xi_k(t) \sim \mathcal{N}\{0\,,\, I\}$, $\eta_k(t) \sim \mathcal{N}\left\{0\,,\, I\cdot\frac{1}{h_k}\right\}$ ($\xi_k(t)$ and $\eta_k(t)$ are mutually independent and independent of the initial condition $x_k(0)$). The steady solutions for the error covariance matrix P_k and the Kalman gain K_k with taking $\lambda_3\to 0$ are given as

$$P_k = \begin{bmatrix} 0 & 0 \\ 0 & \frac{\lambda_2^2}{d(2+dh_k)} \end{bmatrix},\quad K_k = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

which gives the discrete estimations as

$$\begin{aligned} \mu_k^{(1)}(t) &= a\cdot h_k\cdot\mu_k^{(1)}(t-h_k) + y(t), \\ \mu_k^{(2)}(t) &= h_k\cdot\mu_k^{(1)}(t-h_k) + (1+d\cdot h_k)\cdot\mu_k^{(2)}(t-h_k). \end{aligned}$$

The limits of P_k and K_k for $k\to\infty$ are trivial. The continuous (steady-mode) estimation $\mu^{(1)}(t)$ of $x^{(1)}(t)$ is, according to the error covariance matrix P

$$P(t) = \lim_{k\to\infty} P_k(t) = \begin{bmatrix} 0 & 0 \\ 0 & \frac{\lambda_2^2}{2d} \end{bmatrix},$$

perfect and it is given in the (derivative) form

$$\mu^{(1)}(t) = \underset{k\to\infty}{\text{l.i.m.}}\, \mu_k^{(1)}(t) = y(t). \tag{23}$$

Note that the variable $x^{(1)}(t)$ is measured directly through the perfect measurement $y(t)$. However, the continuous estimation $\mu^{(2)}(t)$ of $x^{(2)}(t)$ is gained in the steady mode with a constant error given by $\frac{\lambda_2^2}{2d}$ as the variable $x^{(2)}(t)$ contains additive random errors. The formula for $\mu^{(2)}(t)$ is derived with respect to [1] as follows,

$$\dot{\mu}_k^{(2)}(t) = \text{l.i.m.}\,\frac{\mu_k^{(2)}(t)-\mu_k^{(2)}(t-h)}{h} = y(t) + d\cdot\mu^{(2)}(t),$$

which represents the integrative part of the continuous filter.

8 Conclusions

The objective of this paper is to introduce a new approach to singular Kalman-Bucy filtering problem and to make an attempt to overview other solutions. This work is based on an alternative system theory and gives results in a good agreement with real observations. The class of admissible estimators is supplemented with backward differentiators which can produce physically realizable (i.e. backward) derivatives of smooth signals when necessary. This is possible only after introducing the newly reviewed principle and law of causality as well as so called continualization method into general system theory. As shown here, the optimal filter can generally be of the integrative as well as derivative character. We shall point out, that the singularity of the problem is caused by technical simplifications when the measured data are so accurate that they appear to be perfect. However, practically every measurement contains random errors. Thus, if we claim no noise contained in a signal, the filter can propose derivatives in real time which leads to the use of backwards differentiators.

References

[1] Arnošt, R. (2003). The continualization method in the system theory and its application in estimation. PhD thesis. University of West Bohemia in Pilsen, Faculty of applied Sciences, Department of Cybernetics, Pilsen, Czech Republic.

[2] Arnošt, R. and P. Žampa (2004). On continuous-time systems in a new approach to system theory. In: *Proceedings of the Thirty-Sixth Southeastern Symposium on System Theory*. IEEE. Atlanta, Georgia, USA.

[3] Bode, H. W. and C. E. Shannon (1950). A simplified derivation of linear least square smoothing and prediction theory. *Proc. IRE, vol. 38*.

[4] Bryson, A.E. and D.E. Johansen (1965). Linear filtering for time-varying systems using measurements containing colored noise. *IEEE Transactions on Automatic Control*.

[5] Friedland, B. (1989). On the properties of reduced-order Kalman filters. *IEEE Transactions on Automatic Control*.

[6] Kalman, R.E. (1960). A new approach to linear filtering and prediction problems. *Trans. ASME, J. Basic Eng.*

[7] Kalman, R.E. (1963). New methods in Wiener filtering theory. In: *Proc. of 1st Symp. on Engrg. Application of Random Function Theory and Probability*. New York: Wiley.

[8] Kalman, R.E. and R. C. Bucy (1961). New results in linear filtering and prediction theory. *Trans. ASME, J. Basic Eng.*

[9] Kolmogorov, A. N. (1941). *Interpolation and extrapolation of stationary random sequences*. Bull. Acad. Sci. USSR, Math. Ser. vol. 5. English version published by RAND Corp., Santa Monica, California jako Memo. RM-3090-PR.

[10] Kwakernaak, H. and R. Sivan (1974). The maximally achiebable accuracy of linear optimal regulators and linear optimal filters. *IEEE Transactions on Automatic Control*.

[11] Lewis, F. (1981). A generalized inverse solution to the discrete-time singular Riccati equation. *IEEE Transactions on Automatic Control*.

[12] Moylan, P.J. (1974). A note on Kalman-Bucy filters with zero measurement noise. *IEEE Transactions on Automatic Control*.

[13] Oksendal, B. (2000). *Stochastic differential equations*. Springer-Verlag.

[14] Orfanidis, S.J. (1982). An exact solution of the time-invariant discrete Kalman filter. *IEEE Transactions on Automatic Control*.

[15] Shaked, U. (1985). Explicit solution to the singular discrete-time stationary linear filtering problem. *IEEE Transactions on Automatic Control*.

[16] Shaked, U. (1986). Solution in the z domain to the discrete-time stationary Kalman filtering problem for systems with perfect measurements. *IEEE Transactions on Automatic Control*.

[17] Wiener, N. (1949). *The Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications*. New York: Wiley.

[18] Kalman, R.E., Falb, P.L., Arbib, M.A., Topics in Mathematical System Theory. McGraw-Hill Book Company, New York, 1969.

[19] Willems, J.C., Paradigms and Puzzles in The Theory of Dynamical Systems. IEEE Transaction on Automatic Control, Vol. 36, No 3, March 1991.

[20] Zadeh, L.A., Desoer, C.A., Linear System Theory. The State Space Approach. McGraw-Hill Book Co., New York, 1963.

[21] Žampa, P. (1996). On a new system theory and its new paradigms. In: *Proceedings of the 13th European Meeting on Cybernetics and System Research*. Vienna, Austria.