Non-linear and signal energy optimal asymptotic filter design

Josef Hrušák
Department of Applied Electronics
University of West Bohemia
Univerzitní 8, 306 14 Plzeň, Czech Republic
E–mail: hrusak@kae.zcu.cz

Václav Černý
Department of Cybernetics
University of West Bohemia
Univerzitní 8, 306 14 Plzeň, Czech Republic
E–mail: vcerny@kky.zcu.cz

***Abstract*— The paper studies some connections between the main results of the well known Wiener-Kalman-Bucy stochastic approach to filtering problems based mainly on the linear stochastic estimation theory and emphasizing the optimality aspects of the achieved results and the classical deterministic frequency domain linear filters such as Chebyshev, Butterworth, Bessel, etc. A new non-stochastic but not necessarily deterministic (possibly non-linear) alternative approach called asymptotic filtering based mainly on the concepts of signal power, signal energy and a system equivalence relation plays an important role in the presentation. Filtering error invariance and convergence aspects are emphasized in the approach. It is shown that introducing the signal power as the quantitative measure of energy dissipation makes it possible to achieve reasonable results from the optimality point of view as well. The property of structural energy dissipativeness is one of the most important and fundamental features of resulting filters. Therefore, it is natural to call them asymptotic filters. The notion of the asymptotic filter is carried in the paper as a proper tool in order to unify stochastic and non-stochastic, linear and non-linear approaches to signal filtering.**

***Index Terms*— Causality, Invariance, Structure, Convergence, Signal power, Signal energy, Equivalence**

I. Introduction

Filtering is a large field in signal processing having a wide sort of different applications and long history. Thus, it is not very surprising that a lot of approaches have independently been proposed and developed. In fact, two of them dominate. The first one, the well known and broadly used frequency domain approach based mainly on deterministic linear time-invariant input-output system representations, seems to be the most natural. If all the important characteristics of the problem are known in the stochastic sense, then the stochastic version of the input-output approach known as the Wiener filtering theory can be considered as a proper tool for that. The second approach can be characterized as state space oriented. The well known Wiener-Kalman-Bucy linear filtering theory is its stochastic version [1], [2], [3], [4]. It can be seen as the substantial generalization of the Wiener filtering theory already mentioned above.

Let us briefly summarize some main ideas and important features of both the deterministic and stochastic approaches to signal filtering in this paragraph in order to point differences and similarities between them later. In signal processing [5], [6], [7], [8] the main function of a filter is to remove unwanted parts of a signal such as a random noise and other measurement errors or to extract useful parts of the signal such as its certain components lying within a specific frequency range. Hence, it seems to be natural to start with the general theory of stochastic processes [9], [10] and stochastic estimation theory [11], [12], [13], [14], [15], [16] if the randomness of an unwanted signal uncertainty can be considered as the main attribute of reality and a reasonable quantitative model of the uncertainty can be assumed. Such an approach leads to the well known stochastic non-linear filtering problem concisely introduced in the sections II and III [3], [11], [15], [17], [18], [19], [20]. On the other hand, the frequency domain characterization of both the useful and unwanted parts of the signal seems to be more acceptable in many practical situations and therefore the frequency domain approaches based on the concept of an ideal frequency filter are often preferred. The ideal frequency filter would have a rectangular magnitude response. Unfortunately, it is non-causal. However, there are practical filter design techniques that approximate the ideal frequency filter characteristics and they are realizable. Each of the major types - Butterworth, Chebyshev, Bessel, etc. - optimizes a different aspect of the approximation [5], [6], [7], [8].

The main aim of the paper consists in demonstrating that it is possible to make compatible to each other both the stochastic and frequency domain approaches and their results. Further, they can be seen as a special case of the proposed asymptotic filtering philosophy [21], [22], [23], [24], [25], [26], [27], [28], [29], [30] in a certain way. The paper also tries to explain similarities between the developed asymptotic and optimal filtering. This problem has already been discussed for example in [13], [20], [21], [24]. Nevertheless, any straightforward answer has not been provided yet. Finally, some simulation experiments and their results are presented to illustrate these mentioned points.

II. Stochastic formulation of non-linear filtering problem

Let us shortly characterize a standard stochastic approach to a continuous-time non-linear filtering problem [10], [11], [15], [17].

Consider two vector processes $x(t)$ and $y(t)$. The problem is to estimate the unknown process $x(t)$ from observation of the related process $y(t)$ under the assumption that both processes are described by the following non-linear stochastic differential equations:

$$\begin{aligned} dx(t) &= f[x(t)]dt + g[x(t)]dw(t) & (1)\\ dy(t) &= h[x(t)]dt + m[x(t)]dv(t) & (2)\\ x(t_0) &= x^0. & (3) \end{aligned}$$

The state process $x(t)$ and the output process $y(t)$ evolve on n and p dimensional manifolds. Driving signals $w(t)$ and $v(t)$ are supposed to be represented by r and p dimensional independent standard Wiener processes. The initial state vector x^0 is assumed to be a random variable independent of $w(t)$ and $v(t)$. If mappings f, g, h, m and a probability distribution of the initial state vector x^0 are known, it is said that a local coordinate process representation using Itô differentials is given.

Having the process representation (1), (2), (3), the stochastic non-linear filtering problem is to compute in real time the conditional probability distribution of the current state $x(t)$ given the past observations $y(s)$, $t_0 \leq s < t$.

Stochastic non-linear filtering is an extremely important and very nice but extremely difficult *theoretical* problem as well. It is also known that the problem is critically sensitive to small perturbations of problem specifications. Thus, its direct utilization for practical filter design is relatively limited.

R. E. Kalman and R. S. Bucy [1], [2] discovered the only broad class of stochastic process representations for which an efficient algorithm is known. The class of the representations is a special case of (1), (2), (3) and can be described by the following linear stochastic differential equations:

$$\begin{aligned} dx(t) &= F(t)x(t)dt + G(t)dw(t) & (4)\\ dy(t) &= H(t)x(t)dt + M(t)dv(t) & (5)\\ x(t_0) &= x^0, & (6) \end{aligned}$$

where both the probability distribution of the initial state x^0 and the conditional probability distribution of the state $x(t)$ are Gaussian and thus completely described by their means and covariances.

The most important feature of the stochastic linear filtering problem is that the conditional covariance is independent of the observation process $y(t)$. Consequently, it can be precomputed using *a'priori* knowledge only. Hence, the only real time (observation data dependent) computation needed is that of conditional mean.

There have been several attempts to find other stochastic models than (4), (5), (6) for which the resulting partial differential equation characterizing the conditional probability distribution of the state $x(t)$ would reduce to a finite set of ordinary differential equations driven by the observation process $y(t)$. Unfortunately, those studies have not provided any new wide enough class of finite dimensional filters yet.

It is worthwhile to notice that the mappings f and h are typically derived from physical laws and/or at least closely related to fundamental attributes of reality such as *causality*, *energy conservation principle*, etc. On the other hand, the description of noise effects, especially magnitudes of g and m and their dependence on $x(t)$ is merely the result of educated guesses and simulations. Even in the linear case, there is not any consensus on how to choose the matrices $G(t)$ and $M(t)$.

From the practical filter design point of view, it is of crucial importance to overcome the principal difficulties discussed above. Consequently, the natural question of some new *non-stochastic non-linear filter design paradigm* arises. This point can be seen as one of the main motivations for introducing the concept of *asymptotic filtering* in [21], [22], [23], [24], [30].

III. Basic results of Wiener-Kalman-Bucy linear filtering theory

For later comparison, let us briefly summarize the main results of the well known Wiener-Kalman-Bucy linear stochastic filtering theory [1], [2], [3], [4], [18], [19] in this section.

Consider a linear stochastic time-varying signal generating system representation:

$$\begin{aligned} \mathcal{R}\{\mathcal{S}\} : \frac{dx(t)}{dt} &= A(t)x(t) + B(t)[u(t) + \xi(t)] & (7)\\ y(t) &= C(t)x(t) + \eta(t), & (8) \end{aligned}$$

where $x(t) \in R^n$ is a state vector, $u(t) \in R^r$ is an input and $y(t) \in R^p$ is an output. Matrices $A(t)$, $B(t)$, $C(t)$ and the input signal $u(t)$ are supposed to be known in a *deterministic sense*. Both a driving noise $\xi(t)$ and an observation noise $\eta(t)$ are supposed to be white, Gaussian, zero mean, independent of each other, independent of an initial state x^0 and known in a *stochastic sense*. The initial state x^0 is supposed to be a Gaussian random vector with zero mean and known in the stochastic sense:

$$\begin{aligned} E\{x^0\} &= 0, \quad E\{x^0(x^0)^T\} = P_0 & (9)\\ E\{\xi(t)\} &= 0, \quad E\{\xi(t)\xi^T(\varepsilon)\} = Q(t)\delta(t-\varepsilon) & (10)\\ E\{\eta(t)\} &= 0, \quad E\{\eta(t)\eta^T(\varepsilon)\} = R(t)\delta(t-\varepsilon), & (11) \end{aligned}$$

where $\delta(t-\varepsilon)$ is a Dirac function and $R(t)$, $Q(t)$ and P_0 are supposed to be known symmetric matrices with $R(t) > 0$ and $Q(t)$, $P_0 \geq 0$.

Optimal filtering problem:
Find the estimate $\hat{x}(t|\tau)$ of the state vector $x(t)$ for $t = \tau$; t, $\tau \in [0, T]$ based on the observation $y(s)$, $0 \leq s \leq \tau$, which minimizes the conditional covariance of the state error $\tilde{x}(t|\tau) = x(t) - \hat{x}(t|\tau)$. It means that

$$\hat{x}(t|\tau) = argmin\{E[q\tilde{x}(t|\tau)\tilde{x}^T(t|\tau)q^T|y(s)], 0 \leq s \leq \tau\} \quad (12)$$

for any $1 \times n$ non-vanishing vector q and $\tau = t$.

General solution:
Standard statistical results imply that the solution of the minimum conditional covariance estimation problem mentioned above is given by the conditional mean:

$$\hat{x}(t|\tau) = E[x(t)|y(s), 0 \leq s \leq \tau] \quad (13)$$

with $t = \tau$ for filtering, $t > \tau$ for prediction and $t < \tau$ for smoothing and referred as *optimal* in the minimum variance sense.

Structure of Wiener-Kalman-Bucy filter:
From the practical point of view, it is important to have an efficient method for explicit computing the solution $\hat{x}(t|\tau)$ from the observation $y(s)$ or equivalently a realizable device called an *optimal* filter in order to generate the filtered state $\hat{x}(t|\tau)$ and/or a filtered output $\hat{y}(t)$ on line. Such a device is known as the Wiener-Kalman-Bucy filter and it is given by the following *structure* for the signal generating system representation described above:

$$\begin{aligned} \mathcal{R}\{\mathcal{F}\}: \frac{d\hat{x}(t|\tau)}{dt} &= A(t)\hat{x}(t|\tau) + B(t)u(t) + \\ &+ K(t)[y(t) - C(t)\hat{x}(t|\tau)] \quad (14) \\ \hat{y}(t) &= C(t)\hat{x}(t|\tau). \quad (15) \end{aligned}$$

Parametrization of Wiener-Kalman-Bucy filter:
The matrices $A(t)$, $B(t)$ and $C(t)$ are supposed to be known. Thus, the only unknown matrix is so called the *Kalman gain matrix* $K(t)$:

$$K(t) = P(t)C^T(t)R^{-1}(t) \quad (16)$$

depending on the matrix solution $P(t)$ of the well known *Riccati differential equation*:

$$\begin{aligned} \frac{dP(t)}{dt} &= A(t)P(t) + P(t)A^T(t) + B(t)Q(t)B^T(t) - \\ &- P(t)C^T(t)R^{-1}(t)C(t)P(t) \quad (17) \end{aligned}$$

corresponding to the given initial condition

$$P(0) = P_0, \quad (18)$$

where P_0 is the known covariance matrix of the initial state x^0.

Note that the procedure of the filter parametrization is completely independent of the observed data $y(t)$ and of the estimated real process represented by $x(t)$ as well.

IV. Signal power balance relations for non-stochastic process representations

Consider two vector processes $x(t)$ and $y(t)$ described by the equations (1), (2), (3) as before, but with the *additional assumption* that no reliable quantitative information about process uncertainties is accessible. On the contrary, it is supposed that the vector field f and the vector mapping h can typically be specified by using fundamental physical principles with acceptable precision. The situation is by no means deterministic in such cases, but the stochastic process representation based on Itô differentials can not be completely specified and effectively used. Thus, it seems to be reasonable to reduce the (incomplete) stochastic process representation (1), (2), (3) into the form of the following ordinary vector non-linear differential equation describing state evolution:

$$\begin{aligned} \frac{dx(t)}{dt} &= f[x(t)] + g[x(t)]u(t) \quad (19) \\ x(t_0) &= x^0 \quad (20) \end{aligned}$$

with the non-linear vector output relation:

$$y(t) = h[x(t)]. \quad (21)$$

describing the observation process.

The state process $x(t)$ and the output process $y(t)$ evolve on n and p dimensional manifolds as before. A driving process $u(t)$ is supposed to be an arbitrary *non-stochastic* square integrable function and the initial state x^0 is assumed to be a completely unknown arbitrary point of the state process manifold. The representation (19), (20), (21) can be considered as a given *non-stochastic* local coordinate process representation if the mappings f, g and h are known. It is obvious that non-linear stochastic filtering is not well posed problem under such conditions and any stochastic concept becomes meaningless.

There arises a question whether an alternative non-linear filtering problem formulation can be found which could be expected to provide a reasonable (finite dimensional) solution. Unfortunately, it is not very easy to find any answer to it at this stage. Anyway, we will try to make the first step here and introduce some concepts which seem to be fundamental for developing a new "time-frequency-signal energy metric" approach. Its main idea is inspired by the time-frequency localization problem of the wavelet transformation theory [31], [32], [28], which can be interpreted as time-varying generalization of the classical Fourier-Laplace frequency signal decomposition. Especially the role of a properly generalized form of the classical Parseval's relation and the discovery of "time-frequency Heisenberg-like un-certainty equivalence principle" [32] are extraordinary challenging in the *non-linear* filtering context, too. Later, we will analyze similar situations from a *signal power balance relation* point of view with the objective to find such a time-energy-(frequency)-signal decomposition, which could be effectively used in *non-linear filtering*. Therefore, we summarize some obvious but very important facts now.

In signal processing [6], [7], [8] the total *signal energy* is defined by

$$E = \int_{-\infty}^{\infty} P(t)dt, \quad (22)$$

where the instantaneous value of *signal power* $P(t)$ is defined by

$$\begin{aligned} P_c(t) &= \|u(t)\|^2 \quad (23) \\ P_o(t) &= \|y(t)\|^2. \quad (24) \end{aligned}$$

for the external processes $u(t)$ and $y(t)$ related to a causal system S and its representation $\mathcal{R}\{\mathcal{S}\}$.

Let us postulate the following input and output *power balance relations* [24], [30] for the causal system representation:

$$\begin{aligned} \frac{dE_c(t)}{dt} &= P_c(t) \quad (25) \\ \frac{dE_o(t)}{dt} &= -P_o(t), \quad (26) \end{aligned}$$

where $E_c(t)$ represents the instantaneous value of input process (control signal) energy and $E_o(t)$ represents the instantaneous value of output process (observation signal) energy. It is easy to show that physical correctness [33], minimality [34] and asymptotic stability [35], [36], [37], [38] of the causal system representation are closely related to the postulated signal power balance relations (25), (26).

V. Basic structure of non-stochastic asymptotic non-linear filtering problem

Some fundamental ingredients of an approach to signal filtering based on the concepts of *signal power*, *signal energy* and *system equivalence* motivated by [37], [38] are introduced in this section. Only a continuous-time version is considered here. A discrete-time modification is discussed for example in [25], [27].

A. Problem formulation

Consider a non-linear time-varying signal generating system representation (SGS):

$$\mathcal{R}\{\mathcal{S}\} : \frac{dx(t)}{dt} = f[x(t),t] + g[x(t),t]u(t) \qquad (27)$$

$$y(t) = h[x(t),t], \qquad (28)$$

where $x(t) \in R^n$ is a state vector, $u(t) \in R^r$ is an input and $y(t) \in R^p$ is an output. A vector field f, a vector field matrix g and a vector mapping h are supposed to be known in the deterministic sense. The input and output signals $u(t)$ and $y(t)$ are supposed to be continuously measuered (perhaps with an uncertainty) and the state vector $x(t)$ with its initial state x^0 are supposed to be completely unknown.

Asymptotic filtering problem:
Find a *structure* and proper *parametrization* of a realizable system, which will be called an *asymptotic* filter:

- The filter structure should have a *strict causality property* (expressing a realizability demand) and a *state filtering error invariance property* (expressing the independence requirement of a state filtering error with respect to the input signal $u(t)$, the output signal $y(t)$, the unknown state $x(t)$, a filtered output $\hat{y}(t)$ and a filtered state $\hat{x}(t)$ generated by the filter).
- The filter parametrization should have a *filtering error convergence property*. It means that both the state filtering error and an output filtering error will be *uniformly convergent* to zero and it will be possible to choose the proper *convergence rate and/or mode* of them.

Basic structure of asymptotic filtering problem:
We assume that the state $x(t)$ of the SGS is not accessible for measurement at all. However, its input and output signals $u(t)$ and $y(t)$ are and therefore both of them can be used as the inputs of the filter (fig. 1). As we can see from the fig. 1,

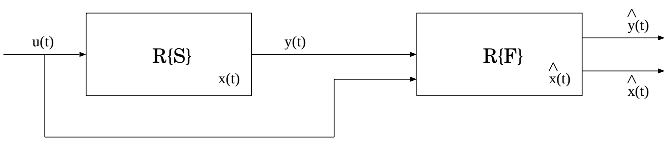

Fig. 1. The basic structure of the asymptotic filtering problem

both the filtered output $\hat{y}(t)$ and the filtered state $\hat{x}(t)$ are considered as the outputs of the filter.

B. Filtering error invariance and filter structure determination

The *strict causality property* alone implies that the class of filter representations can be identified with the following class of non-linear time-varying representations:

$$\hat{\mathcal{R}}\{\mathcal{F}\} : \frac{d\hat{x}(t)}{dt} = \hat{F}[\hat{x}(t),u(t),y(t),t] \qquad (29)$$

$$\hat{y}(t) = \hat{H}[\hat{x}(t),t]. \qquad (30)$$

From the *state filtering error invariance property* for the state filtering error $\tilde{x}(t) = x(t) - \hat{x}(t)$ expressed by:

$$\tilde{\mathcal{R}}\{\mathcal{F}\} : \frac{d\tilde{x}(t)}{dt} = \tilde{f}[\tilde{x}(t),t] \qquad (31)$$

we get the *structure* of the filter:

$$\mathcal{R}\{\mathcal{F}\} : \frac{d\hat{x}(t)}{dt} = f[\hat{x}(t),t] + g[\hat{x}(t),t]u(t) + \hat{K}[y(t),\hat{x}(t),t] \qquad (32)$$

$$\hat{y}(t) = h[\hat{x}(t),t]. \qquad (33)$$

Consider for now that the representation $\mathcal{R}\{\mathcal{S}\}$ is linear:

$$\mathcal{R}\{\mathcal{S}\} : \frac{dx(t)}{dt} = A(t)x(t) + B(t)u(t) \qquad (34)$$

$$y(t) = C(t)x(t). \qquad (35)$$

Then

$$\tilde{\mathcal{R}}\{\mathcal{F}\} : \frac{d\tilde{x}(t)}{dt} = [A(t) - \hat{K}(t)C(t)]\tilde{x}(t) \qquad (36)$$

$$\tilde{y}(t) = C(t)\tilde{x}(t), \qquad (37)$$

where $\tilde{y}(t) = y(t) - \hat{y}(t)$ is the output filtering error, and

$$\mathcal{R}\{\mathcal{F}\} : \frac{d\hat{x}(t)}{dt} = A(t)\hat{x}(t) + B(t)u(t) + \hat{K}(t)[y(t) - C(t)\hat{x}(t)] \qquad (38)$$

$$\hat{y}(t) = C(t)\hat{x}(t). \qquad (39)$$

In fact, if $A(t) = A$, $B(t) = B$, $C(t) = C$ and $\hat{K}(t) = \hat{K}$, then the filter structure is closely related to the well known Luenberger observer [39], [40].

C. Filter parametrization for linear time-varying signal generating system representations

Assume that the signal generating system representation has an asymptotical stability property and is of the minimal order n. It means that it is controllable and observable. In such a case controllability and observability Grammian matrices $W_c(t)$ and $W_o(t)$ exist, are symmetric, positive definite and satisfy the following Lyapunov equations:

$$A(t)W_c(t) + W_c(t)A^T(t) + \frac{dW_c(t)}{dt} = -B(t)B^T(t) \qquad (40)$$

$$A^T(t)W_o(t) + W_o(t)A(t) + \frac{dW_o(t)}{dt} = -C^T(t)C(t). \qquad (41)$$

Let us start with the group of linear time-varying state transformations:

$$\bar{x}(t) = T(t)x(t),\ x(t) = T^{-1}(t)\bar{x}(t). \quad (42)$$

The equivalent representation of the error signal generating system (EGS) is now given by:

$$\begin{aligned}\tilde{\bar{\mathcal{R}}}\{\mathcal{F}\} : \frac{d\tilde{\bar{x}}(t)}{dt} &= [\bar{A}(t) - \hat{\bar{K}}(t)\bar{C}(t)]\tilde{\bar{x}}(t) & (43)\\ \tilde{\bar{y}}(t) &= \bar{C}(t)\tilde{\bar{x}}(t) & (44)\\ \tilde{\bar{v}}(t) &= \bar{B}^T(t)\tilde{\bar{x}}(t), & (45)\end{aligned}$$

where $\tilde{\bar{v}}(t)$ is a dual output filtering error, and

$$\begin{aligned}\bar{A}(t) &= [T(t)A(t) + \frac{dT(t)}{dt}]T^{-1}(t) & (46)\\ \bar{B}(t) &= T(t)B(t) & (47)\\ \hat{\bar{K}}(t) &= T(t)\hat{K}(t) & (48)\\ \bar{C}(t) &= C(t)T^{-1}(t). & (49)\end{aligned}$$

It is natural to identify the *filtering error convergence demand* with a form of a *filtering error signal dissipativeness condition*. If the instantaneous value of the *output filtering error signal power* is given by:

$$\tilde{\bar{P}}_o(t) = \|\tilde{\bar{y}}(t)\|^2 \quad (50)$$

and the instantaneous value of the *error signal generating system energy* $\tilde{\bar{E}}(t)$ *accumulated at a time instant* t in the error signal generating system is defined by [38]:

$$\tilde{\bar{E}}(t) = \tilde{\bar{E}}[\tilde{\bar{x}}(t)] = \delta\|\tilde{\bar{x}}(t)\|^2 \quad (51)$$

with $\delta > 0$ as an *energy scaling parameter*, then the *signal energy conservation principle* can be expressed in the form of the *error signal power balance relation* [21]:

$$\frac{d\tilde{\bar{E}}(t)}{dt} = -[\rho^{-1}(t)\|\tilde{\bar{y}}(t)\|^2 + \sigma(t)\|\tilde{\bar{v}}(t)\|^2] \quad (52)$$

with $\rho^{-1}(t) > 0$ and $\sigma(t) > 0$ as *design parameters*. By the help of them we can specify the *required degree of dissipativeness* (rate and/or mode of convergence) or *some prior knowledge* about measures of *input and output uncertainties*. Computing the time derivative $\frac{d\tilde{\bar{E}}(t)}{dt}$ along the given EGS representation and comparing with the relation (52) we get the special form of a Lyapunov equation for the equivalent EGS representation:

$$\begin{aligned}\delta[\bar{A}(t) + \bar{A}^T(t)] &= \delta[\hat{\bar{K}}(t)\bar{C}(t) + \bar{C}^T(t)\hat{\bar{K}}^T(t)] - \\ &- [\rho^{-1}(t)\bar{C}^T(t)\bar{C}(t) + \sigma(t)\bar{B}(t)\bar{B}^T(t)].\end{aligned} \quad (53)$$

It would be used for determining the *gain matrix* $\hat{K}(t)$ if the proper state transformation matrix $T(t)$ was known.

Energy conservation principle and determination of state transformation

Certainly, any real-world SGS has to satisfy a form of the *signal energy conservation law*. If it is expressed in a proper form, it gives the state transformation matrix $T(t)$. It is natural to assume the existence of such a coordinate system where an *appropriate energy function takes the same form as it is defined for the error signal generating system*. Thus, we have for the equivalent representation of the SGS:

$$\begin{aligned}\bar{E}(t) &= \bar{E}[\bar{x}(t)] = \delta\|\bar{x}(t)\|^2 & (54)\\ \frac{d\bar{E}(t)}{dt} &= \rho^{-1}(t)\|\bar{y}(t)\|^2 - \sigma(t)\|\bar{v}(t)\|^2, & (55)\end{aligned}$$

where $\delta > 0$, $\rho^{-1}(t) > 0$, $\sigma(t) \geq 0$, $\bar{x}(t)$ is a state vector, $\bar{y}(t)$ is an output signal and $\bar{v}(t)$ is a dual output signal. Computing the time derivative $\frac{d\bar{E}(t)}{dt}$ along the equivalent SGS representation and comparing with the relation (55) we get the special form of the Lyapunov equation for the equivalent SGS representation:

$$\delta[\bar{A}(t)+\bar{A}^T(t)] = \rho^{-1}(t)\bar{C}^T(t)\bar{C}(t) - \sigma(t)\bar{B}(t)\bar{B}^T(t). \quad (56)$$

Combining the relations (53) and (56) and performing some elementary modifications we obtain the expression for the equivalent gain matrix $\hat{\bar{K}}(t)$:

$$\hat{\bar{K}}(t) = \delta^{-1}\bar{C}^T(t)\rho^{-1}(t). \quad (57)$$

Hence, the *gain matrix* $\hat{K}(t)$ in the original coordinates is given by:

$$\hat{K}(t) = \delta^{-1}[T^T(t)T(t)]^{-1}C^T(t)\rho^{-1}(t). \quad (58)$$

Using the signal power balance relations we get the *matrix differential equation* for the state transformation matrix $T(t)$:

$$\begin{aligned}\frac{dT(t)}{dt}T^{-1}(t) &+ [T^T(t)]^{-1}\frac{dT^T(t)}{dt} = T(t)A(t)T^{-1}(t) - \\ &- \frac{\rho^{-1}(t)}{\delta}[T^T(t)]^{-1}C^T(t)C(t)T^{-1}(t) + \\ &+ \frac{\sigma(t)}{\delta}T(t)B(t)B^T(t)T^T(t) + \\ &+ [T(t)A(t)T^{-1}(t)]^T. \end{aligned} \quad (59)$$

The isometry condition [38] implies that

$$\forall t,\ \forall x(t) : E[x(t)] = \bar{E}[\bar{x}(t)] \text{ for } \bar{x}(t) = T(t)x(t). \quad (60)$$

Subsequently, it follows for the Lyapunov energy function $E[x(t)]$ of the SGS representation in the original coordinates that

$$E[x(t)] = \delta\|T(t)x(t)\|^2 = x^T(t)S(t)x(t), \quad (61)$$

where $S(t) = \delta T^T(t)T(t)$.

D. Relation to stochastic case

The parameter δ is positive, the state transformation matrix $T(t)$ is invertible and hence the matrix $S(t)$ is always positive definite. It means that the error signal generating system is *structurally dissipative*. Let us define the symmetric positive definite matrix $P(t)$ by the following relation

$$P(t) = S^{-1}(t) = \delta^{-1}[T^T(t)T(t)]^{-1} \quad (62)$$

and the parameters δ, $\rho^{-1}(t)$ and $\sigma(t)$ choose as follows:

$$\delta = 1,\ \rho^{-1}(t)\cdot I = R^{-1}(t),\ \sigma(t)\cdot I = Q(t). \quad (63)$$

Then

$$\hat{K}(t) = P(t)C^T(t)R^{-1}(t) \quad (64)$$

and the *matrix differential equation* (59) for the state transformation matrix $T(t)$ becomes to

$$\begin{aligned}\frac{dP(t)}{dt} &= A(t)P(t) + P(t)A^T(t) + B(t)Q(t)B^T(t) - \\ &- P(t)C^T(t)R^{-1}(t)C(t)P(t). \qquad (65)\end{aligned}$$

It can be seen that the relations (16) and (64) for the *gain matrix* as well as the *matrix differential equations* (17) and (65) are *equivalent* if (63) holds.

Finally, it has been demonstrated that there is not any essential difference between the results of *stochastic* and *non-stochastic* filtering in linear case from the structural point of view. It has been shown that the *structure* of both the filters is the same and the *asymptotic* filter is *optimal* in the *minimum variance sense for the choice of the design parameters* (63). Thus, the main results of both the approaches are *equivalent* in such a case.

VI. Non-stochastic error signal energy optimal asymptotic filter design

The *asymptotic* properties of the filter have only been analyzed and any *optimality* arguments have not explicitly been used in the previous section. On the other hand, any reasonable solution of any problem can be considered as the *optimal* one from a point of view. It has been shown in the section V that the *non-stochastic asymptotic* filter is closely related to the *optimal* one in the stochastic sense. Therefore, an example of *non-stochastic asymptotic filter optimization* will briefly be characterized here [35], [36], [37], [38].

We will consider a time-invariant case in order to get more explicit results. It means that an appropriate error signal generating system representation has the following form for $\tilde{v}(t) = 0$:

$$\begin{aligned}\tilde{\mathcal{R}}\{\mathcal{F}\} : \frac{d\tilde{x}(t)}{dt} &= \tilde{A}\tilde{x}(t) = [A - \hat{K}C]\tilde{x}(t) \qquad (66)\\ \tilde{y}(t) &= C\tilde{x}(t). \qquad (67)\end{aligned}$$

Optimality criterion:
The output filtering error signal energy has been chosen as the optimality criterion:

$$\tilde{E}(t) = \int_t^\infty \|\tilde{y}(\tau)\|^2 d\tau,\ t = t_0. \qquad (68)$$

Then

$$\tilde{A} = argmin\ \tilde{E}(t), \qquad (69)$$

where $\tilde{A} = A - \hat{K}C$.

Determination of actual value of error signal generating system energy
It is obvious that

$$\tilde{y}(\tau) = C\tilde{x}(\tau),\ \tilde{x}(\tau) = e^{\tilde{A}(\tau - t)}\tilde{x}(t),\ \tau \geq t. \qquad (70)$$

Substituting (70) to (68) we get

$$\tilde{E}(t) = \tilde{E}[\tilde{x}(t)] = \tilde{x}^T(t)\tilde{W}_o\tilde{x}(t), \qquad (71)$$

where

$$\tilde{W}_o = \int_t^\infty e^{\tilde{A}^T(\tau - t)}C^TCe^{\tilde{A}(\tau - t)}d\tau. \qquad (72)$$

The energy function $\tilde{E}[\tilde{x}(t)]$ can be considered as a Lyapunov function generated by the observability Grammian matrix $\tilde{W}_o$ satisfying the Lyapunov equation:

$$\tilde{A}^T\tilde{W}_o + \tilde{W}_o\tilde{A} = -C^TC. \qquad (73)$$

Actual energy minimization
The energy function $\tilde{E}[\tilde{x}(t)]$ can be expressed in the metric equivalent form:

$$\tilde{\tilde{E}}[\tilde{\tilde{x}}(t)] = \delta\|\tilde{\tilde{x}}(t)\|^2, \qquad (74)$$

where $\delta = (2\omega_0)^{-1}$. Subsequently, it holds that

$$\begin{aligned}\delta[\tilde{\tilde{A}} + \tilde{\tilde{A}}^T] &= -\bar{C}^T\bar{C} \qquad (75)\\ \tilde{\tilde{A}} &= T[A - \hat{K}C]T^{-1}. \qquad (76)\end{aligned}$$

It is easy to specify the *optimal* solution by parametric minimization in the following general matrix form [35], [38]:

$$\tilde{\tilde{A}} = \omega_0 \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 \\ -1 & 0 & 1 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & -1 & 0 & 1 \\ 0 & \cdots & 0 & -1 & 0 \end{bmatrix},\ \omega_0 > 0 \qquad (77)$$

or equivalently in the recursively given normalized *optimal* filter transfer function form:

$$\begin{aligned}F(s) &= \frac{1}{P_n(s)} \qquad (78)\\ P_0(s) &= 1 \qquad (79)\\ P_1(s) &= s + \omega_0 \qquad (80)\\ P_k(s) &= sP_{k-1}(s) + \omega_0^2 P_{k-2}(s) \text{ for } k \in \{2, \ldots, n\}. \qquad (81)\end{aligned}$$

For example, a 7th order filter has the following transfer function:

$$F(s) = \frac{1}{P_7(s)}, \qquad (82)$$

where

$$\begin{aligned}P_7(s) &= s^7 + \omega_0 s^6 + 6\omega_0^2 s^5 + 5\omega_0^3 s^4 + 10\omega_0^4 s^3 + \\ &+ 6\omega_0^5 s^2 + 4\omega_0^6 s + \omega_0^7. \qquad (83)\end{aligned}$$

A. Relation to frequency domain case

Note that the positive design parameter ω_0 has the meaning of time scale transformation and can be used to adjust the required *bandwidth* of the filter. The integer n has been defined as the order of the minimal representation of the signal generating system and it can be interpreted as *signal complexity measure*. On the other hand, it also represents *measure of achievable degree of filter quality* in the sense of the best realizable approximation of the required *ideal* filter.

VII. Experimental results

Some properties of the asymptotic filters are illustrated by numerical examples in this section.

Example 1: Connection with the frequency domain approach:

- $n = 1$, $\omega_0 = 1$: $F(s) = \frac{1}{P_1(s)}$, $P_1(s) = s + 1$
- $n = 2$, $\omega_0 = 1$: $F(s) = \frac{1}{P_2(s)}$, $P_2(s) = s^2 + s + 1$
- $n = 7$, $\omega_0 = 1$: $F(s) = \frac{1}{P_7(s)}$
- $n = 14$, $\omega_0 = 1$: $F(s) = \frac{1}{P_{14}(s)}$
- $n = 21$, $\omega_0 = 1$: $F(s) = \frac{1}{P_{21}(s)}$

The corresponding frequency responses are shown on the fig. 2. It can be seen on the figure that the transfer properties of the filters converge to that of the *ideal low pass filter* with increasing signal complexity measure.

Example 2: Comparison of the stochastic optimal Wiener-Kalman-Bucy and non-stochastic error signal energy optimal asymptotic filters: Consider a linear second-order time-invariant signal generating system producing the output signal $y(t)$ shown at the fig. 6. The following situation is reflected:

- The output signal $y(t)$ of the system is disturbed by a white noise with the known mean $\mu = 0$ and variance $\sigma^2 = 2$. Additionally, the output signal is disturbed by a systematic error (an unknown constant is contained in the signal) as well (see the fig. 3).

Let us design both the filters. Their behaviour is shown on the fig. 4, 5.

It follows from the fig. 4, 5 that the W-K-B filter gives *better* results in the sense of eliminating the *white noise disturbance*. On the contrary, the asymptotic filter provides *better* results in the sense of eliminating an *initial state uncertainty* and the *systematic error*. The W-K-B filter does not even *converge* to zero at all.

VIII. Acknowledgments

This work was partly supported by The Ministry of Education of Czech Republic under the projects No. LN00B084 and LN00B096.

The paper [30] has been selected as the best one in the section Digital Signal Processing of the conference and it has subsequently been recommended for publication in the SCI Journal in an extended form.

IX. Conclusions

It has been shown in the paper that the new approach to signal filtering based on the concepts of the *signal power*, *signal energy*, *signal power balance relation* and *suitably defined state equivalence transformation* can be used as a proper tool for the synthesis and design of so called the *asymptotic* filters. The main features of the approach are its *flexibility* with respect to *uncertainty modelling* and its relative *independence* of the standard *linearity* assumptions. It is shown in the special case of a linear signal generating system with the standard stochastic Gaussian uncertainty representation that the results are closely *related* to the well known Wiener-Kalman-Bucy linear stochastic *optimal* filtering theory. On the other hand, the same approach has been used as a proper tool for the non-stochastic asymptotic filter design by the parametric optimization in the linear non-stochastic case where no quantitative uncertainty model is considered. It can be seen that the frequency responses of such the *optimized* non-stochastic linear asymptotic filters are *comparable* with the well known *frequency domain* classical filters (such as Chebyshev, Butterworth, Bessel, etc.). The approach for the asymptotic filter design presented here can be used in *non-linear* observer design [22], [23], [26], [41], *non-linear* control [41], [42], [43], [44], [45] and elsewhere [46], [47].

References

[1] R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Basic Eng., A.S.M.E.*, vol. 82, pp. 201–205, 1960.
[2] R. S. Bucy and R. E. Kalman, "New results in linear filtering and prediction theory," *J. Basic Eng., A.S.M.E.*, vol. 83, pp. 95–108, 1961.
[3] R. S. Bucy and P. D. Joseph, *Filtering for Stochastic Processes with Applications to Guidance*. Interscience Publishers, John Wiley & Sons, Inc., New York, London, Sydney, Toronto, 1968.
[4] M. S. Grewal and A. P. Andrews, *Kalman Filtering, Theory and Practice Using MATLAB*. John Wiley & Sons, Inc., 2001.
[5] A. W. Oppenheim and R. W. Schafer, *Digital Signal Processing*. Prentice-Hall, Inc., 1975.
[6] E. C. Ifeachor and B. W. Jervis, *Digital Signal Processing, A Practical Approach*. Addison-Wesley Publishing Company, New York, 1996.
[7] C. B. Rorabangh, *DSP Primer*. McGraw-Hill, Inc., 1999.
[8] S. W. Smith, *The Scientist and Engineer's Guide to Digital Signal Processing*. California Technical Publishing, 1999.
[9] R. L. Stratonovich, "Conditional Markov Processes," Theory of Probability and Applications, 1960.
[10] K. Itô, "Lectures on Stochastic Processes," Lecture Notes, Tata Institute of Fundamental Research, Bombay, India, 1961.
[11] R. S. Bucy, "Nonlinear filtering," *IEEE Trans. Aut. Control*, 1965.
[12] D. L. Alspach and H. W. Sorenson, "Approximation of Density Function by a Sum of Gaussian Estimation," in *Proceedings of The 1st Symposium on Nonlinear Estimation Theory*, San Diego, USA, 1970.
[13] R. E. Mortensen, "Maximum likelihood recursive non-linear filtering," *J. Opt. Theory and Appl.*, vol. 2, no. 6, pp. 386–394, 1968.
[14] H. W. Sorenson, "An Overview of Filtering and Stochastic Control in Dynamic Systems," in *Advances in Theory and Applications*, vol. 12, 1976, pp. 1–61.
[15] S. I. Marcus, "Algebraic and Geometric Methods in Nonlinear Filtering," *SIAM J. Control and Optimization*, vol. 22, pp. 817–844, 1984.
[16] J. Hrušák, J. Mošna, and M. Šimandl, "Nonlinear Filtering: System Set Representations," in *Proceedings of The 10th IFAC World Congress*, vol. 9, Munich, Germany, 1987, pp. 255–260.
[17] A. J. Krener, "The Intristic Geometry of Dynamic Observations," M. Fliess and M. Hazewinkel (eds.), Algebraic and Geometric Methods in Nonlinear Control Theory, pp. 77–87, 1986.
[18] K. Wagner, J. Hrušák, and O. Pirohanič, "Investigation of Dynamic Behavior of U-D Factorized Kalman Filter," in *Proceedings of The 6th IFAC Symposium on Identification and System Parameter Estimation*, Washingtion, D. C., USA, 1982, pp. 397–402.
[19] A. J. Krener, *Minimum Covariance, Minimax and Minimum Energy Linear Estimators*. in Lecture Notes in Control and Information Sciences, Stochastic Control Theory and Stochastic Differential Systems, Springer-Verlag, Berlin, Heidelberg, New York, 1979, pp. 490-495.
[20] M. R. James, "Asymptotic nonlinear filtering and large deviations with application to observer design," Ph.D. dissertation, University of Maryland, College Park, MD 20742, USA, 1988.
[21] J. Hrušák and M. Štork, "On equivalence relations in stochastic and non-stochastic signal filtering," in *Proceedings of The International Conference WBU/IEEE: Applied Electronics*, Plzeň, Czech Republic, 2001, pp. 108–113.
[22] J. Hrušák and V. Černý, "Nonlinear state reconstruction of linear systems," in *Proceedings of The 4th International Scientific-Technical Conference Process Control*, Kouty nad Desnou, Czech Republic, June,11.–14. 2000.
[23] V. Černý and J. Hrušák, "Non-linear observer design via dissipation normal form," in *Proceedings of The International Scientific Conference of FME*, Ostrava, Czech Republic, September,5.–7. 2000.

[24] J. Hrušák and M. Štork, “Asymptotic filters: An unifying approach to signal filtering,” in *Proceedings of WBU*, Plzeň, Czech Republic, 2001, pp. 85–100.

[25] M. Štork and J. Hrušák, “Digital realization of the continuous-time minimum energy asymptotic filters,” in *Proceedings of The IEEE International Conference Applied Electronics*, Plzeň, Czech Republic, September, 11.–12. 2002.

[26] J. Hrušák and V. Černý, “Generalization of nonlinear observer design via dissipation normal form,” in *Proceedings of The 11th IFAC International Conference Control Applications of Optimization*, St. Petersburg, Russia, July,3.–6. 2000.

[27] J. Hrušák and D. Pánek, “Signal energy conservation and lattice structure of asymptotic filters,” in *Proceedings of The IEEE International Conference Applied Electronics*, Plzeň, Czech Republic, September, 10.–11. 2003.

[28] M. Štork and J. Hrušák, “On construction of wavelets based on continuous-time asymptotic filters,” in *Proceedings of The IEEE International Conference Applied Electronics*, Plzeň, Czech Republic, September, 10.–11. 2003.

[29] J. Teplý, “Adaptive asymptotic filter design,” Master’s thesis, University of West Bohemia, Department of Applied Electronics, Plzeň, Czech Republic, 2003.

[30] J. Hrušák and V. Černý, “Non-linear and signal energy optimal asymptotic filter design,” in *Proceedings of The 7th World Multiconference on Systemics, Cybernetics and Informatics*, Orlando, USA, July, 27.–30. 2003.

[31] J. C. Goswanni and A. K. Chan, *Fundamentals of wavelets, Theory, Algorithms and Applications*. John Wiley & Sons, Inc., 1999.

[32] Y. T. Chan, *Wavelet Basics*. Kluwer Academic Publishers, 1995.

[33] D. Mayer and J. Hrušák, “On correctness and asymptotic stability in causal system theory,” in *Proceedings of The 7th World Multiconference on Systemics, Cybernetics and Informatics*, Orlando, USA, July, 27.–30. 2003.

[34] A. C. Antoulas, *Mathematical System Theory*. Springer-Verlag, 1991.

[35] P. C. Parks, “A new proof of the Routh-Hurwitz stability criterion by the second method of Lyapunov with applications to optimum transfer functions,” in *Proceedings of The Joint Automatic Control Conference*, Minneapolis, USA, 1963.

[36] R. E. Kalman and J. E. Bertram, “Control System Analysis and Design via the Second Method of Lyapunov: I. Continuous-time Systems; II. Disctrete-time Systems,” *J. Basic Eng., A.S.M.E.*, vol. 82, no. 2, pp. 371–393, 394–400, 1960.

[37] J. Hrušák, “Anwendung der aquivalenz bei stabilitatsprufung,” Tagung uber die Regelungs-theorie, Mathematisches Forschungsinstitut, Oberwolfach, W. Germany, Tech. Rep., 1969.

[38] ——, “Isometric Transformation Method and some of its Applications,” Ph.D. dissertation, Technical University, Pilsen, Czechoslovakia, 1971.

[39] D. G. Luenberger, “Observing the state of a linear system,” *IEEE Trans. Mil. Electron.*, vol. 8, 1964.

[40] ——, “Observers for multivariable systems,” *IEEE Trans. Aut. Control*, *AC-11*, pp. 190–199, 1966.

[41] V. Černý, “Stabilization, separation and state reconstruction of nonlinear systems,” Ph.D. dissertation, University of West Bohemia, Department of Cybernetics, Plzeň, Czech Republic, 2002.

[42] V. Černý, A. Fialová, and J. Hrušák, “Stabilization of non-linear systems based on metric equivalence,” in *Proceedings of The IEEE 9th Mediterranean Conference on Control and Automation*, Dubrovnik, Croatia, June,27.–29. 2001.

[43] A. Fialová and V. Černý, “Synthesis of control law for non-linear systems in critical case,” in *Proceedings of The 1st IFAC/IEEE Symposium on System Structure and Control*, Prague, Czech Republic, August,29.–31. 2001.

[44] V. Černý and J. Hrušák, “Synthesis of control law for non-linear systems based on metric equivalence,” in *Proceedings of The IASTED International Conference Control and Applications*, Cancun, Mexico, May,20.–22. 2002.

[45] ——, “Separation principle for a class of non-linear systems,” in *Proceedings of The IEEE 11th Mediterranean Conference on Control and Automation*, Rhodes, Greece, June, 18.–20. 2003.

[46] J. Hrušák, J. Mošna, E. Janeček, and M. Šimandl, “Strip-Thickness Adaptive Control for Cold-Rolling Mills,” *Problems of Control and Information Theory*, vol. 11, no. 6, pp. 455–464, 1982.

[47] ——, “A New Adaptive Controller for Cold Rolling Mills,” in *Proceedings of IFAC/IFIP Symposium on Real Time Digital Control Applications*, Guadalajara, Mexico, 1983, England: Pergamon, pp. 69–74, 1984.

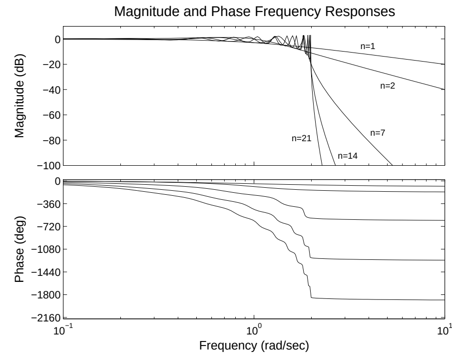

Fig. 2. The frequency responses of the non-stochastic error signal energy optimal asymptotic filters for $\omega_0 = 1$ and the different signal complexity measures $n = 1, 2, 7, 14, 21$

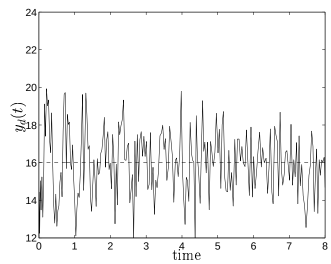

Fig. 3. The output signal disturbed by a white noise and a systematic error

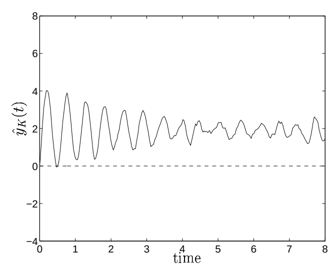

Fig. 4. The filtered output of the W-K-B filter

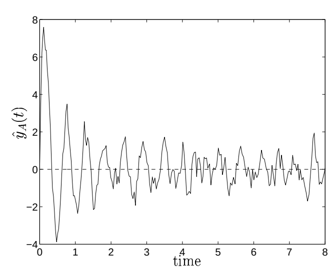

Fig. 5. The filtered output of the asymptotic filter

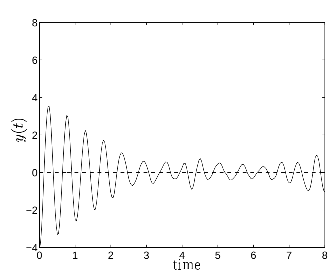

Fig. 6. The output signal without any measurement errors