Mathematical constructions in optimal linear filtering theory

Vladimir N. Fomin   Michael V. Ruzhansky

Department of Mathematics and Mechanics, St.Petersburg University
Department of Mathematics, Utrecht University

Abstract

The linear optimal filtering problems in infinite dimensional Hilbert spaces and their extensions are investigated. The quality functional is allowed to be a general quadratic functional defined by a possibly degenerate operator. We describe the solution of the stable and the causal filtering problems. In case of the causal filtering we establish the relation with a relaxed causal filtering problem in the extended space. We solve the last problem in continuous and discrete cases and give the necessary and sufficient conditions for the solvability of the original causal problem and conditions for the analogue of Bode–Shannon formula to define an optimal filter.

1 Introduction

We consider the linear optimal filtering problems in infinite dimensional Hilbert spaces and their extensions. Briefly, the problem is as follows. Let H', H'' be Hilbert spaces and $z = \begin{bmatrix} x \\ y \end{bmatrix}$ a random element in $H = H' \times H''$, where x and y are unobservable and observable components of z in H' and H'' respectively. The correlation operator of z is assumed to be bounded in H and we denote by $\mathbb{H}$ a subset of all linear operators $h : H'' \to H'$. The $\mathbb{H}$–optimal linear filtering problem is a problem of the estimations of the unobservable component x based on the realizations of the observable component y in the form

$$\hat{x} = hy \tag{1}$$

solving the minimization problem in $\mathbb{H}$

$$J(h) \to \inf_{h \in \mathbb{H}}, \tag{2}$$

where the quality functional J is defined by

$$J(h) = E||D(x - \hat{x})||^2 \tag{3}$$

with a suitable norm in (3) and a linear operator $D : H' \to H'$. If $\mathbb{H}$ consists of all continuous linear operators h, the problem (1), (2), (3) is called *stable*. If H' and H'' are Hilbert resolution spaces, one has a time structure in H and in its terms defines an "independent of the future" class of the causal continuous operators $\mathbb{H}$. In this case the problem is called *causal*.

If H' and H'' are finite dimensional and D is the identity matrix, the problem (1), (2), (3) is equivalent to the problem first being solved in [6], [9]. The causal operators become the upper triangular matrices and the solution is unique and can be efficiently represented in terms of the factorization of the correlation matrix R_y of y by the so called *Bode–Shannon formula* ([1]). Spectral factorization coincides with the well known Holetsky factorization of matrices in this case ([2]). The result was immediately applied in various branches of the filtering theory, for some applications see [9], [7].

However, in many applications one estimates only the specific components of x or their combination, which is represented by the degenerate matrix D in the quality functional (3) and the solutions of the generalized finite dimensional problems can be found in [8]. In this case the solution need not be unique and there are conditions on the degeneracy of D for which the Bode–Shannon formula still defines an optimal filter.

On the other hand the infinite dimensional applications required the development of the filtering theory in Hilbert ([3]) and sometimes Banach spaces ([4], [5]). For the applications of this theory to the problem of the linear estimation of the parameters of a signal based on the observations of its realizations see for example [4].

In this paper the stable filtering problem will be solved for the general quadratic quality functional (3). The solution of the causal filtering problem need not exist in general. We will establish necessary and sufficient conditions of the solvability by relaxing the problem allowing a slightly general class of the weight operators in (2). The relaxed problem can be solved and the analysis of its solution can be used for the construction of the minimizing sequences. The solutions will be given for continuous and discrete resolutions. The conditions for the analogue of the Bode–Shannon formula of [4] to define an optimal filter will be also given.

Some of the literature we are referring to is in Russian and for the sake of completeness the results of [4] needed in this paper will be briefly reviewed. We will not give the complete proves of them in order to avoid unimportant for the nature of the results of this paper technicalities.

In Section 2 we fix the notation related to the concept of the extended Hilbert space and random elements in it. In Section 3 we formulate and give the solution of the general linear filtering problem in extended Hilbert spaces in Theorem 3.1. The stable linear filtering problem is solved in Section 4 (Theorem 4.1). Section 5 is devoted to the causal filtering problem. In Subsections 5.1, 5.2 we discuss Hilbert resolution spaces, their extensions and linear operators in extended spaces. In Subsection 5.3 we formulate the problem. The corresponding relaxed problem is solved in Subsection 5.4 (Theorem 5.1). In Subsection 5.5 we treat the case of the discrete resolution of the identity (Theorem 5.2) and give necessary and sufficient conditions of the solvability of the original problem (Theorem 5.3). In Section 6 the concept of the spectral factorization will be discussed and the conditions for the Bode–Shannon formula to define an optimal filter will be given in Theorem 6.2.

2 Preliminaries

2.1 Extended Hilbert spaces and linear operators

Here we will review a concept of the extended Hilbert space. Let H be a complex Hilbert space with an inner product $\langle \cdot, \cdot \rangle$ and $F \subset H$ be a linear dense subset of H. We introduce the notion of F-weak convergence in H:

Definition 2.1 *A sequence $\psi_l \in H$ is called F–*weakly Cauchy sequence *if*

$$\lim_{l,m\to\infty} \langle \psi_l - \psi_m, \phi \rangle = 0, \forall \phi \in F.$$

Let Ψ be a set of all F-weakly Cauchy sequences $\psi = \{\psi_l\}$, $\psi_l \in H$, $l \in \mathbb{N}$. Let $\overset{F}{\cong}$ be an equivalence relation in Ψ: $\psi \overset{F}{\cong} \xi$ if $\lim_{m\to\infty}\langle \psi_m - \xi_m, \phi \rangle = 0$ for all $\phi \in F$. Then it is not difficult to check that the quotient space $H_F \equiv \Psi / \overset{F}{\cong}$ of Ψ with respect to the equivalence relation $\overset{F}{\cong}$ is a linear Hausdorff topological space.

Every element $\bar{\psi} \in H_F$ defines a functional $\bar{\psi}^* : F \to \mathbb{C}$ by $\bar{\psi}^*(\phi) = \lim_{l\to\infty}\langle \psi_l, \phi \rangle$, where $\{\psi_l\}$ is a sequence from $\bar{\psi}$. This duality is an extension of the inner product in H and we will denote this also by $\bar{\psi}^*(\phi) = \langle \bar{\psi}, \phi \rangle$. The following relation is obvious:

Proposition 2.1 *H_F is complete in F-weak topology and $F \subset H \subset H_F$.*

The pair (F, H_F) is called *an equipment* of H and H with such an equipment *an equipped Hilbert space*. We will also use a construction which gives a space equivalent to H_F.

Definition 2.2 *The space F^* is a space of all the elements f for which there exists a sequence $f_l \in H^*$ such that $f(\phi) = \lim_{l\to\infty} f_l(\phi)$ for all $\phi \in F$.*

Obviously, F^* is complete with respect to the topology of componentwise convergence on F. This implies the completeness of H_F in view of the

Proposition 2.2 (i) *F^* is isomorphic to H_F.*

(ii) *Let $\bar{\psi} \in H_F$. Define for the corresponding $\bar{\psi}^* \in F^*$ a "norm" $|\bar{\psi}^*|_{F^*} = \sup_{\phi\in F} \frac{|\langle \bar{\psi}, \phi \rangle|}{|\phi|_H}$. Then $\bar{\psi} \in H$ if and only if $|\bar{\psi}^*|_{F^*} < \infty$. In this case $|\bar{\psi}^*|_{F^*} = |\bar{\psi}|_H$.*

The proof easily follows from the definitions above. Let $A : H \to H$ be a linear operator defined in a dense subspace $D(A)$ of H. Recall the following

Definition 2.3 *Let $D_* \subset H$ be a space of all elements $\phi \in H$ for which there exist $f(\phi) \in H$ such that $\langle A\psi, \phi \rangle = \langle \psi, f(\phi) \rangle$ for all $\psi \in D(A)$. The operator adjoint to A is defined as $A^* : D_* \to H$, such that $A^*\phi = f(\phi)$.*

One has then $D_* = D(A^*)$. An operator A is called *symmetric* if $D(A) \subset D(A^*)$ and for every $\phi, \psi \in D(A)$ holds $\langle A\psi, \phi \rangle = \langle \psi, A\phi \rangle$. A symmetric operator A is called *self-adjoint* if $D(A) = D(A^*)$.

Let $\bar{A} : D(\bar{A}) \to H_F$ be an extension of A. Assume $F \cap D(\bar{A})$ to be dense in H. Similar to the definition above define D_* as a space of all $\phi \in F$ for which there exist $f(\phi) \in F$ such that $\langle \bar{A}\bar{\psi}, \phi \rangle = \langle \bar{\psi}, f(\phi) \rangle$ for all $\bar{\psi} \in D(\bar{A})$. Let $\bar{A}^* : D_* \to H_F$ be an operator defined by $\langle \bar{A}\bar{\psi}, \phi \rangle = \langle \bar{\psi}, \bar{A}^*\phi \rangle$ for all $\bar{\psi} \in D(\bar{A})$, $\phi \in D_*$. The F–weak closure of $\bar{A}^*$ is called *the adjoint* to $\bar{A}$ in H_F and will be also denoted by $\bar{A}^*$. As above, $\bar{A}$ is called *symmetric* if $D(\bar{A}) \subset D(\bar{A}^*)$ and A is symmetric. A symmetric operator $\bar{A}$ is called *self–adjoint* if $D(\bar{A}) = D(\bar{A}^*)$.

Example 2.1 *Let M be a smooth manifold and $H = L^2(M)$ with respect to some positive smooth density on M. Let F_k be a space of k–times continuously differentiable compactly supported functions in M. Then the F_k–weak completion of H is a space of the distributions of order k in M.*

Example 2.2 *Let $H = l^2(\mathbb{N})$ with its standard inner product and let $A : l^2 \to l^2$ be defined by $(A\phi)_n = \sum_{k=1}^{\infty} A_{nk}\phi_k$ with suitable conditions on A_{nk}. Let $F \subset H$ be a set of all the sequences consisting of finite number of nonzero elements. Then F is dense in H and $F \subset D(A) \subset H$. For the extension of H with respect to F one has $H_F = \mathbb{R}^{\mathbb{N}}$, the space of all the sequences with values in $\mathbb{R}$. One readily checks that the above extension of A yields a linear operator $\bar{A}$ in $\bar{l}^2 = H_F$ defined by a matrix A_{nk} with $D(\bar{A}) = \{\phi : \sum_{k=1}^{\infty} |A_{nk}\phi_k|^2 < \infty, n \in \mathbb{N}\}$.*

2.2 Generalized random elements

Let F, H, H_F be as above and let $(\Omega, \mathcal{A}, P)$ be a probability space, P a complete measure.

Definition 2.4 *A mapping $z : \Omega \to H_F$ is called a* random H_F–element *if for every $\phi \in F$ holds:*

(i) *$z^*\phi = \langle z, \phi \rangle : \Omega \to \mathbb{C}$ is a random variable.*

(ii) *$Ez^*\phi = (Ez)^*\phi$ for some $Ez \in H_F$.*

(iii) *there exists c such that $E|(z - Ez)^*\phi|^2 \le c|\phi|_H^2$ uniformly in $\phi \in F$.*

Without loss of generality we will consider the centralized elements: $Ez^*\phi = 0$ for all $\phi \in F$. Then $E\overline{z^*\phi}z^*\phi = E|z^*\phi|^2 = \langle \phi, R_z\phi \rangle$ is a quadratic form in F. A linear operator R_z is called the *correlation operator* of z. Property (iii) of the definition implies that R_z is a continuous operator on F and, therefore, it can be extended to a continuous self-adjoint operator in H, $R_z \in \mathcal{L}(H)$. Thus, we have proved

Proposition 2.3 *$R_z \in \mathcal{L}(H)$, $R_z^* = R_z$ and $R_z \ge 0$ in the sense of quadratic forms.*

In analogy to the classical case we will write $R_z = Ezz^*$. Note, that Definition 2.4 is equivalent to the condition that $z : \Omega \to H_F$ is measurable, centralized and has continuous correlation operator, where Ω and H_F are equipped with σ–algebras $\mathcal{A}$ and one generated by the open sets of F–weak topology in H_F respectively.

Example 2.3 *The random processes can be interpreted as random generalized elements of a Hilbert space. Let $T > 0$ and $H = L^2(m, T)$ be a Hilbert space of m–vector functions on $[0, T]$ equipped with the standard inner product. Let $w = \{w(t), 0 \le t \le T\}$ be a Gauss process, such that almost all realizations of w are elements of $L^2(m, T)$ and $E\int_o^T |w(t)|^2 dt < \infty$. Then R_w is a nuclear operator. If $T = +\infty$, then almost all realizations of $w \notin L^2(m, T)$, but for $F = \mathcal{C}_{\text{comp}}([0, T])$ the realizations of w are elements of $L^2(m, T)_F$.*

3 Linear filtering

Let $H = H' \times H''$, where H' and H'' are Hilbert spaces with inner products $\langle \cdot, \cdot \rangle_{H'}$ and $\langle \cdot, \cdot \rangle_{H''}$ respectively. Let $F' \subset H'$ and $F'' \subset H''$ be linear dense subsets. The elements $\phi \in H$ can be interpreted as $\phi = \begin{bmatrix} \phi' \\ \phi'' \end{bmatrix}$ with $\phi' \in H'$, $\phi'' \in H''$. Let $F = F' \times F''$. We will consider random H_F elements $z = \begin{bmatrix} x \\ y \end{bmatrix}$, with x and y random $H'_{F'}$– and $H''_{F''}$– elements respectively. The correlation operator R_z will be assumed continuous on H, which is natural in view of Proposition 2.3 and have the following block form:

$$R_z = \begin{bmatrix} R_x & R_{xy} \\ R_{yx} & R_y \end{bmatrix},$$

where we write $R_x = Exx^*$, $R_y = Eyy^*$, $R_{xy} = R_{yx}^* = Exy^*$.

Let $h : H''_{F''} \to H'_{F'}$ be linear. We assume now that there exist an operator $h^* : H'_{F'} \to H''_{F''}$ defined on the whole of $H'_{F'}$, such that for every $\phi' \in F'$, $\phi'' \in F''$ one has

$$(h\phi'')^*\phi' = (\phi'')^*(h^*\phi'). \tag{4}$$

Relation (4) defines h^* uniquely and h^* is the adjoint to h operator.

Let x and y be the unobservable and observable components of z respectively. We define the random $H'_{F'}$–element $\hat{x}$ by

$$\hat{x} = hy. \tag{5}$$

One readily checks that $\hat{x}$ is a random element in the sense of Definition 2.4 in view of our assumptions on h. Then $R_{\hat{x}} = hR_yh^* : H' \to H'$ is involutive in H' as the correlation operator of a random element $\hat{x}$. The element $\hat{x}$ is interpreted as a linear estimate of the nonobservable component x of a random H_F element z, based on the realizations of its observable component y. The relation (5) is called *a linear filter with weight operator h.*

Let a linear operator $D : {H'}_{F'} \to {H'}_{F'}$ have an adjoint D^*. We define *the quality functional* as

$$J_{\phi'}(h) = E|\langle \phi', D(x - \hat{x})\rangle|^2,\ \phi' \in F'. \tag{6}$$

Let $\mathbb{H}$ be a given subset of linear operators $h : {H''}_{F''} \to {H'}_{F'}$. Then the $\mathbb{H}$–*optimal filtering problem* is defined as a problem of the minimization of the functionals

$$J_{\phi'}(h) \to \inf_{h \in \mathbb{H}}, \tag{7}$$

defined by (6), (5) for every $\phi' \in F'$.

We will need a notion of the pseudo inversion of an operator. Let $A : H \to H$ be a linear operator in a Hilbert space H. Let Q_A be an orthogonal projection on the image of A, $Q_A : H \to \mathrm{Im}A$. The space $Q_A H$ is invariant for A and we write $A^{-1}Q_A$ for the inverse of A in $Q_A H$. The operator

$$A^+ = Q_A \circ A^{-1}Q_A \circ Q_A$$

is called the *pseudo inverse of* A. It follows that

$$A^+A = AA^+ = A. \tag{8}$$

One readily checks that (8) determines A^+ uniquely and

Proposition 3.1 *If A is a Hermitian operator, then the solution of*

$$\langle Ag - f, Ag - f\rangle \to \inf_{g \in H},\ f \in H$$

with minimal norm defines a linear functional of f which is given by $g = A^+f$.

We will not prove this fact here since we will not use it explicitly. Assume that $\mathbb{H}$ is a space of all linear operators $h : {H''}_{F''} \to {H'}_{F'}$. Then the solution of the $\mathbb{H}$–optimal filtering problem is given by

Theorem 3.1 *Let the correlation operator R_z of a random H element z be continuous in H and let R_y^+ denote the pseudo inverse operator for the correlation operator R_y of y in H''. Then the minimization problem (7) in the class $\mathbb{H}$ of all weight operators $h : {H''}_{F''} \to {H'}_{F'}$ is solvable and any solution is of the form*

$$h_{\mathrm{opt}} = R_{xy}R_y^+ + Q, \tag{9}$$

where $Q : {H''}_{F''} \to {H'}_{F'}$ is any linear operator satisfying $DQ = 0$. Moreover, one has

$$\inf_{h \in \mathbb{H}} J_{\phi'}(h) = J_{\phi'}(h_{\mathrm{opt}}) = \langle \phi', D[R_x - R_{xy}R_y^+ R_{xy}^*]D^*\phi'\rangle.$$

The proof follows the lines of the proof of Theorem 4.1, which is given in the next section. The existence of D^* assures the decomposition (12), from which the statement of Theorem 3.1 follows.

4 Linear stable filtering

If $\mathbb{H}$ is a space of all *continuous* linear operators from H'' to H', then the linear filters of the form (5) with weight operator in $\mathbb{H}$ are called *stable* and $\mathbb{H}$–optimal filtering problem is called the *stable filtering problem.* In this case one allows $\phi' \in H'$ in (6) and the minimization problem can be reformulated for scalar functionals

$$J(h) = \sup_{\phi' \in H'} \frac{E|\langle \phi', D(x - \hat{x})\rangle|^2}{|\phi'|^2_{H'}}. \tag{10}$$

Now we are ready to describe the solution of the linear stable filtering problem. Let us assume that R_y is continuously invertible in its image $R_y H''$, which means that there exist a neighborhood U of zero such that $\sigma(R_y) \cap U = \{0\}$, $\sigma(R_y)$ being the spectrum of R_y. We will also assume that the operator D in the quality functional (6) is continuous in H' and has an adjoint D^*.

Theorem 4.1 *Let the correlation operator R_z of a random H element z be continuous in H and assume that the correlation operator R_y of y has the continuous pseudo inverse operator R_y^+ in H''. Then the minimization problem (7) in the class $\mathbb{H}$ of all continuous weight operators $h : H'' \to H'$ is solvable and any solution is of the form*

$$h_{\text{opt}} = R_{xy} R_y^+ + Q, \tag{11}$$

where $Q : H'' \to H'$ is any linear continuous operator satisfying $DQ = 0$. Moreover, one has

$$\inf_{h \in \mathbb{H}} J_{\phi'}(h) = J_{\phi'}(h_{\text{opt}}) = \langle \phi', D[R_x - R_{xy} R_y^+ R_{xy}^*] D^* \phi' \rangle_{H'}.$$

The operators (11) are also optimal in the problem with quality functional (10) and

$$\inf_{h \in \mathbb{H}} J(h) = J(h_{\text{opt}}) = |D[R_x - R_{xy} R_y^+ R_{xy}^*] D^*|_{H'}.$$

Proof First we rewrite the quality functionals (6) as

$$J_{\phi'}(h) = \langle \phi', R_{D(x-hy)} \phi' \rangle,$$

where $R_{D(x-hy)}$ is the correlation operator of $D(x - hy)$ and using the existence of D^* and h^* we have

$$\begin{aligned} R_{D(x-hy)} &= E[D(x-hy)][D(x-hy)]^* = DE(x-hy)(x-hy)D^* \\ &= D[R_x - R_{xy}h^* - hR_{yx} + hR_y h^*]D^*. \end{aligned}$$

This means

$$\begin{aligned} J_{\phi'}(h) &= \langle \phi', D[R_x - R_{xy}h^* - hR_{yx} + hR_y h^*]D^* \phi' \rangle \\ &= \langle \phi', D[R_x - R_{xy} R_y^+ R_{xy}^*] D^* \phi' \rangle + \\ &\quad \langle \phi', D(h - R_{xy} R_y^+) R_y (h - R_{xy} R_y^+)^* D^* \phi' \rangle. \end{aligned} \tag{12}$$

Here only the second term depends on h and it is a nonnegative quadratic form attaining its minimum if and only if $D(h - R_{xy}R_y^+) = 0$ in view of Proposition 2.3. The set of all continuous h satisfying this equation is precisely the set of h_{opt} in (11) for all linear continuous $Q : H'' \to H'$ satisfying $DQ = 0$. For such h_{opt} the second term in (12) is zero, implying the second statement of the theorem. It follows from (12) that functionals (11) are also optimal for the problem (10) and one readily verifies the last statement of the theorem. The proof is complete.

Remark 4.1 *If the kernel of R_y is nontrivial ($Q_{R_y} \neq I_{H''}$), then one has $h = h_{\text{opt}} + \tilde{h}(I_{H''} - Q_{R_y})$, where Q_{R_y} is the orthogonal projection on the image of R_y, and $J_{\phi'}(h) = J_{\phi'}(h_{\text{opt}})$ for every linear continuous operator $\tilde{h} : H'' \to H'$. If R_y is bijective ($Q_{R_y} = I_{H''}$), then $R_y^+ = R_y^{-1}$ and $h_{\text{opt}} = R_{xy}R_y^{-1} + Q$, $DQ = 0$. If D is bijective, then the only solution of (7) is $h_{\text{opt}} = R_{xy}R_y^+$.*

Remark 4.2 *If R_y can be rewritten in the diagonal form by a suitable choice of a basis in H'', the problem of finding a pseudo inverse operator of R_y simplifies. In stationary case the filtering problem can be also reformulated in frequency terms. These methods can be applied for the problems of linear estimation of the parameters of a signal based on the observations of its realizations.*

We will not discuss it here, but the reader can consult [4] for the detailed application.

5 Linear causal filtering

In this section we will give the solution of the generalized linear causal filtering problem. However, we need some preliminary notions and results first.

5.1 Hilbert resolution spaces and causal operators

Let H be a Hilbert space, $\mathbb{T} = (t_s, t_f)$, $-\infty \leq t_s < t_f \leq +\infty$ and let $\mathbb{P}_T = \{P_t, t \in \mathbb{T}\}$ be a family of commutative projectors $P_t : H \to H$, $P_t^2 = P_t$, $P_tP_s = P_sP_t$, $t, s \in \mathbb{T}$. Let $\mathbb{P}_T$ satisfy the following two properties

(i) *monotonicity*: $P_tP_s = P_s$ for $t \geq s$, $t, s \in \mathbb{T}$.

(ii) *completeness*: $\lim_{t\to t_s} P_t = 0_H$, $\lim_{t\to t_f} P_t = I_H$, where the limits are taken in the strong operator topology.

Note, that condition (i) is equivalent to the fact that $P_sH \subset P_tH$, $t \geq s$. We assume the family $\mathbb{P}_T$ to be bounded uniformly in t: $\sup_{t\in\mathbb{T}} |P_t| < \infty$ and strongly continuous from the left: $\lim_{\epsilon\to 0+} P_{t-\epsilon}\phi = P_t\phi$ for every $\phi \in H$. Such family $\mathbb{P}_T$ is called a *resolution of the identity of* H and $(H, \mathbb{P}_T)$ is called a *Hilbert resolution space*. If $\mathbb{P}_T$ consists of the orthogonal projectors: $P_t = P_t^*$, then it is called a *Hermitian resolution of the identity*. In this case the condition of the uniform boundedness in t is automatically satisfied since $|P_t| \leq 1$.

Let $H = H' \times H''$, where $(H', \mathbb{P}'_T)$, $(H'', \mathbb{P}''_T)$ are Hilbert resolution spaces. Then H may be equipped with the canonical resolution of the identity

$$P_t = \begin{bmatrix} P'_t & 0_{12} \\ 0_{21} & P''_t \end{bmatrix}, \; t \in \mathbb{T}, \tag{13}$$

where $0_{12} : H'' \to H'$, $0_{21} : H' \to H''$ are zero operators.

Definition 5.1 *Let $A : D(A) \to H$ be a linear densely defined operator. A is called* finite from above *if there exists a measurable, essentially bounded function $\tau : \mathbb{T} \to \mathbb{T}$, such that for almost all $t \in \mathbb{T}$ the operator $P_t A$ is bounded in H and if $t - \tau(t) \in \mathbb{T}$, then*

$$P_t A = P_t A P_{t-\tau(t)} \tag{14}$$

on $D(A) \cap P_{t-\tau(t)} D(A)$. The function $\tau = \tau_+(\cdot)$ is called the upper characteristic of A. *A finite from above operator A with characteristic $\tau_+(\cdot)$ is called* τ–causal *or* τ_+–finite.

The space of all τ_+–finite operators will be denoted by $\mathbb{A}^\tau$ and $\mathbb{A}^0 = \cup_\tau \mathbb{A}^\tau$. 0–causal operators are called *causal.* For $\phi \in H$ one can consider a trajectory $\{P_t\phi, \; t \in \mathbb{T}\}$ connecting ϕ and zero in H. Then (14) means that a τ–causal operator A considered as a shift operator along these trajectories does not depend on a future with respect to the resolution, namely it follows from the completeness of $\mathbb{P}_T$ that $P_t A\phi$ is independent of $P_s\phi$ for $s > t - \tau(t)$. One has also a notion of finiteness from below, given in the following

Definition 5.2 *Let $A : D(A) \to H$ be a linear densely defined operator. A is called* finite from below *if there exists a measurable, essentially bounded function $\tau : \mathbb{T} \to \mathbb{T}$, such that for almost all $t \in \mathbb{T}$ the operator $(I_H - P_t)A$ is defined in $D(A)$ and if $t - \tau(t) \in \mathbb{T}$, then*

$$(I_H - P_t)A = (I_H - P_t)A(I_H - P_{t-\tau(t)}) \tag{15}$$

on $D(A) \cap (I_H - P_{t-\tau(t)})D(A)$. The function $\tau = \tau_-(\cdot)$ is called the lower characteristic of A. *A finite from below operator A with characteristic $\tau_-(\cdot)$ is called* τ–anticausal.

Note, that for the Hermitian resolution of the identity A is finite from below if and only if A^* is finite from above and $\tau_-(A) = \tau_+(A^*)$. 0–anticausal operators are called *anticausal* and one writes $\mathbb{A}_\tau$ for all $\tau = \tau_-$–anticausal operators, $\mathbb{A}_0 = \cup_\tau \mathbb{A}_\tau$. If $A \in \mathbb{A}^0 \cap \mathbb{A}_0$ and $\tau_+ = \tau_- = \tau$, then A is called *τ–local.* 0–local operator is called *local.* Every operator commuting with $\mathbb{P}_T$ is local. We will need the following property

Lemma 5.1 ([4]) *Let* $A \in \mathbb{A}^\tau$ (resp.$\mathbb{A}_\tau$), $A' \in \mathbb{A}^{\tau'}$(resp.$\mathbb{A}_{\tau'}$). *Then $B = AA'$ (if exists) is finite from above (resp. below) with characteristic $\beta(t) = \tau(t) + \tau'(t - \tau(t))$.*

Example 5.1 *Let* $H = L^2(\mathbb{R})$ *and* $(h\phi)(t) = \int_{-\infty}^{+\infty} h(t,s)\phi(s)ds$. *Hilbert space* L^2 *becomes a resolution space when equipped with a family* $\mathbb{P}_T$ *defined by*
$P_t\phi(s) = \begin{cases} \phi(s), & \text{if } s \leq t \\ 0, & \text{if } s > t \end{cases}$. *One readily sees that* h *is* $\tau-$*causal (anticausal) if and only if* $h(t,s) = 0$ *for* $s > \min(t, t-\tau(t))$, $(s < \max(t, t-\tau(t)))$. *In particular,* h *is local if and only if* $h = 0$.

5.2 Extended Hilbert resolution spaces

Assume now that H is infinite dimensional and $t_f = +\infty$. An element $\phi \in H$ is called *finite* if there exist $t_*(\phi) \in \mathbb{T}$, $t_* < \infty$, such that $P_t\phi = \phi$ for all $t \geq t_*$. Let F be a space of all finite elements of H. Then F is dense in H and $\bar{H} = H_F$ is called the $t-$*extension* or $t-$*completion* of H. We write $t - \lim_{n\to\infty} \phi_n = \phi$ if for every $t \in \mathbb{T}$ one has $\lim_{n\to\infty} P_t\phi_n = P_t\phi$; $\phi_n, \phi \in H$. This defines $t-$*convergence* in H and the associated Hausdorff topology is weaker than the canonical inner product topology of H. Note that H is not complete with respect to $t-$convergence. One readily checks the following simple

Proposition 5.1 *The completion of* H *with respect to* $t-$*topology is isomorphic to* H_F, *the* $F-$*weak completion of* H.

A densely defined operator A in H is called $t-$*continuous* if for every sequence $\phi_n \in D(A)$ with $t - \lim_{n\to\infty} \phi_n = 0$ one has $t - \lim_{n\to\infty} A\phi_n = 0$. Note that a general continuous operators in H need not be $t-$continuous. Now we collect the further properties following [4] (see also [3]).

Lemma 5.2 *The following holds:*

(i) *Every finite from above continuous operator in* H *is* $t-$*continuous.*

(ii) *If* A *is* $t-$*continuous, then by Proposition 5.1 it allows an extension to an operator* $\bar{A}$ *in* $\bar{H}$: $\bar{A}|_H = A$. *In particular, every* $P_t \in \mathbb{P}_T$, *being a local operator, allows an extension to* $\bar{P}_t$ *in* $\bar{H}$. *The family* $\overline{\mathbb{P}_T}$ *is a resolution of the identity in* $\bar{H}$. *One can generalize the notions of causality for* $\bar{H}$, *in particular* $\bar{P}_t$ *are local in* $\bar{H}$.

(iii) *For every* $t \in \mathbb{T}$ *and* $\bar{\phi} \in \bar{H}$ *holds* $\bar{P}_t\bar{\phi} \in P_tH$.

(iv) *A restriction of* $\tau-$*causal operator* $\bar{A}$ *in* $\bar{H}$ *to* H *defines a* $\tau-$*causal operator* A *in* H.

(v) *If* $|\bar{\phi}|_{\bar{H}} = \sup_{t\in\mathbb{T}} |\bar{P}_t\bar{\phi}|_H$, *then* $|\bar{\phi}|_{\bar{H}} < \infty$ *if and only if* $\bar{\phi} \in H$. *In this case* $|\bar{\phi}|_{\bar{H}} = |\bar{\phi}|_H$.

(vi) *Let* $\bar{A} : \bar{H} \to \bar{H}$ *be linear* $\tau-$*causal. Then there exists an operator* $\bar{A}^* : \bar{H} \to \bar{H}$ *uniquely defined by*

$$(\bar{P}_t\bar{A}\bar{P}_{t-\tau(t)}\bar{\phi})^*\psi = (\bar{P}_{t-\tau(t)}\bar{\phi})^*\bar{A}^*\bar{P}_t\psi$$

for every $\psi \in F$, $\bar{\phi} \in \bar{H}$, $t \in \mathbb{T}$. *The operator* $\bar{A}^*$ *is the adjoint to* $\bar{A}$ *and is* $(-\tau)-$*anticausal.*

Definition 5.3 *An operator* $\bar{A} : \bar{H} \to \bar{H}$ *is called* $\tau-$bounded *for a measurable function* $\tau : \mathbb{T} \to \mathbb{T}$*, if*

$$\sup_{\bar{\phi}\in\bar{H}} \sup_{t\in\mathbb{T}} \frac{|\bar{P}_t\bar{A}\bar{\phi}|_H}{|\bar{P}_{t-\tau(t)}\bar{\phi}|_H} < \infty.$$

$0-$*bounded operators are called* stable ([3]).

We collect the properties of $\tau-$bounded operators in

Lemma 5.3 *The following holds:*

(i) *Let* $\bar{A} : \bar{H} \to \bar{H}$ *be* $\tau-$*bounded. Then* H *is invariant subspace for* $\bar{A}$ *and the restriction* $\bar{A}|_H$ *is continuous.*

(ii) *Let* $\bar{A} : \bar{H} \to \bar{H}$ *be* $\tau-$*bounded for* $\tau \geq 0$*. Then* $\bar{A}$ *is* $\tau-$*causal with respect to* $\overline{\mathbb{P}_T}$*.*

(iii) *An operator* $\bar{A} : \bar{H} \to \bar{H}$ *is stable if and only if*

  (a) $\bar{A}$ *is causal.*

  (b) H *is an invariant subspace of* $\bar{A}$*.*

  (c) *The restriction* $\bar{A}|_H$ *is continuous in* H*.*

5.3 Linear causal filtering problem

Let $(H', \mathbb{P}'_T)$, $(H'', \mathbb{P}''_T)$ be Hermitian resolution spaces. Let $H = H' \times H''$ be equipped with the resolution defined by (13). We denote by $\mathbb{H}^\tau$ the space of all linear continuous $\tau-$causal operators $h : H'' \to H'$. Let $D : H' \to H'$ be continuous with the adjoint $D^* : H' \to H'$. Then *the optimal linear causal filtering problem* is the minimization problem

$$J_{\phi'}(h) \to \inf_{h\in\mathbb{H}^\tau} \tag{16}$$

for every $\phi' \in H'$, where $J_{\phi'}(h)$ is defined by

$$J_{\phi'}(h) = E|\langle\phi', D(x - hy)\rangle|^2,\ h \in \mathbb{H}^\tau. \tag{17}$$

It turns out that the condition of the continuity of weight operators is very restrictive for the solution of the problem (16). We will apply the methods presented in [4], namely first we relax the problem (16) allowing h to be unbounded. Analyzing the solution of the relaxed problem we derive the conditions for the solvability of (16).

5.4 Generalized linear causal filtering problem

Let $\bar{H}'$, $\bar{H}''$ be the $t-$completions of H' and H'' respectively. Let $\bar{\mathbb{H}}^\tau$ be the space of all linear $\tau-$causal operators $\bar{h} : \bar{H}'' \to \bar{H}'$, such that for every $t \in \mathbb{T}$ the operators $\bar{P}'_t\bar{h}\bar{h}^*\bar{P}'_t : P'_tH' \to P'_tH'$ are continuous. Assume z to be a random $\bar{H}$ element, and, therefore, $R_z = Ezz^*$, $z = \begin{bmatrix} x \\ y \end{bmatrix}$, is bounded on the space F of finite elements in H and can be then continuously extended to the whole of H. The problem is to find linear estimates of a random H' element x based on the realizations of a random H'' element y of the form

$$\hat{x} = \bar{h}y, \tag{18}$$

minimizing for every $t \in \mathbb{T}$ the functional

$$J^{(t)}(\bar{h}) = E|D\bar{P}_t(x - \hat{x})|^2_{H'}. \tag{19}$$

Note that $J^{(t)}(\bar{h})$ is finite for $\bar{h} \in \bar{\mathbb{H}}^\tau,\ t \in \mathbb{T}$, therefore the problem of the minimization

$$J^{(t)}(\bar{h}) \to \inf_{\bar{h}\in\bar{\mathbb{H}}^\tau} \tag{20}$$

for every $t \in \mathbb{T}$ is correctly posed. Let us reformulate the problem (20) now. For $\phi' \in H'$ we define

$$\begin{aligned} J^{(t)}_{\phi'}(\bar{h}) &= E|\langle \phi', DP'_t(x - \hat{x})\rangle_{H'}|^2 = E|\langle \phi', DP'_t(x - \bar{h}x)\rangle_{H'}|^2 \\ &= \langle \phi', DP'_t[R_x - R_{xy}\bar{h}^* - \bar{h}R_{yx} + \bar{h}R_y\bar{h}^*]P'_tD^*\phi'\rangle_{H'}. \end{aligned} \tag{21}$$

Now, the problem (20) is equivalent to the problem

$$J^{(t)}_{\phi'}(\bar{h}) \to \inf_{\bar{h}\in\bar{\mathbb{H}}^\tau} \tag{22}$$

for every $\phi' \in H'$.

Theorem 5.1 *Let $R_z = Ezz^*$ satisfy*

(i) *The operators $R_z^{(t,t)} = \bar{P}_tR_z\bar{P}_t : P_tH \to P_tH$ are continuous for every $t \in \mathbb{T}$.*

(ii) *The operators $P''_tR_yP''_t : H'' \to H''$ are positive in the invariant subspace P''_tH'' for every $t \in \mathbb{T}$.*

Then for every $t \in \mathbb{T}$ there exist $\hat{x}_t \in P'_tH'$ such that for every $\phi \in H'$ one has

$$E|\langle \phi', D(x - \hat{x}_t)\rangle_{H'}|^2 = \inf_{h\in\bar{\mathbb{H}}^\tau} E|\langle \phi', D(x - P'_thy)\rangle_{H'}|^2.$$

The estimates $\hat{x}_t$ are given by

$$\hat{x}_t = R_{xy}^{(t,t-\tau(t))}(R_y^{(t-\tau(t),t-\tau(t))})^{-1}P''_{t-\tau(t)}y + Q_tP''_{t-\tau(t)}y, \tag{23}$$

where $R_{xy}^{(t,t-\tau(t))} = P'_tR_{xy}P''_{t-\tau(t)}$, $R_y^{(t,t)} = P''_tR_yP''_t$, $(R_y^{(t,t)})^{-1}$ means the inverse of $R_y^{(t,t)}$ in the invariant subspace $P''_{t-\tau(t)}H''$ and any $Q_t : H'' \to H'$ such that $DQ_t = 0$. Moreover,

$$E|\langle \phi', D(x-\hat{x}_t)\rangle_{H'}|^2 = \langle \phi', D[P'_tR_xP'_t - R_{xy}^{(t,t-\tau(t))}(R_y^{(t-\tau(t),t-\tau(t))})^{-1}R_{yx}^{(t-\tau(t),t)}]D^*\phi'\rangle_{H'}.$$

Proof In view of (23) we rewrite (21) as

$$J^{(t)}_{\phi'}(\bar{h}) = \langle \phi', D[P'_t R_x P'_t - R_{xy}^{(t,t-\tau(t))}\bar{h}^* - \bar{h}R_{yx}^{(t-\tau(t),t)} + \bar{h}R_y^{(t-\tau(t),t-\tau(t))}\bar{h}^*]D^*\phi'\rangle_{H'}. \tag{24}$$

The minimization problem (22) is now the same as the minimization of the functionals (24) in the invariant for $R_y^{(t-\tau(t),t-\tau(t))}$ subspace $H''_{t-\tau(t)} = P''_{t-\tau(t)}H''$. This is the minimization problem (7) for $H' = P'_t H'$ and $H'' = H''_{t-\tau(t)}$. Theorem 4.1 together with the invertability of $R_y^{(t-\tau(t),t-\tau(t))}$ by the assumption (ii) of Theorem 5.1 imply the solution of the problem in the form given by (23) and the last formula of the theorem.

The detailed discussion and the solutions of these problems for $D = I_{H'}$ can be found in [4], [5]. We will treat further the spaces with the discrete resolution of the identity. In general, the problems described above can be reduced to the discrete case by a suitable approximation of $\mathbb{P}_T$ by discrete resolutions of the identity, see [4] for the details.

5.5 Discrete resolutions of the identity

We assume now that $\mathbb{P}_T$ is a piecewise constant operator valued functional on $\mathbb{T}$ with at most countable number of discontinuity points without accumulations in $\mathbb{T}$. Let $\mathbf{t} = \{t_k, k \in \mathbb{K}\}$ be a finite or a countable ordered subset of $\mathbb{T}$ without accumulation points, $\mathbb{K} = \mathbb{Z} \cap (0, K)$, $t_0 = t_s$, $t_K = t_f$, K finite or $K = +\infty$. The *discrete resolution of the identity in H corresponding to* $\mathbf{t} \subset \mathbb{T}$ is the set $\mathbb{P}_{\mathbf{t}} = \{P_t, t \in \mathbf{t}\}$. The family of the orthogonal projectors $Q_k = P_{t_k} - P_{t_{k-1}}$, $k \in \mathbb{K}$ determines the resolution $\mathbb{P}_{\mathbf{t}}$ uniquely due to the relation $P_t = \sum_{k:t_k \le t} Q_k$. These projectors are mutually orthogonal: $Q_k Q_l = Q_l Q_k = 0_H$ for $k \neq l$.

Definition 5.4 *A family $\mathbb{Q}_K$ of the mutually orthogonal projectors Q_k is called* the orthogonal resolution of the identity *if $\mathbb{Q}_K$ is complete in a sense that $Q_k \to O_H$ for $k \to k_s$ and $\sum_{l \le k} Q_l \to I_H$ for $k \to k_f$. The pair $(H, \mathbb{Q}_K)$ is called the* discrete resolution space.

Every linear operator $R : H \to H$ can be decomposed with respect to $\mathbb{Q}_K$ into blocks $R_{kl} = Q_k R Q_l$ and $R = \sum_{k,l \in \mathbb{K}} R_{kl}$. The definitions of finiteness, causality and anticausality can be reformulated in terms of the discrete structure $\mathbb{Q}_K$. The function τ in Definitions 5.1, 5.2 is replaced by $\tau : \mathbf{t} \to \mathbf{t}$ with a property that $\tau(t_k) = t_l$, $k, l \in \mathbb{K}$ and the latter corresponds to a function $\kappa : \mathbb{K} \to \mathbb{K}$ such that $\tau(t_k) = t_{\kappa(k)}$. In analogy to the continuous case one has

Definition 5.5 *A linear operator $R : H \to H$ is called* κ–causal (strictly κ–causal, κ–anticausal) *if $R_{kl} = 0_H$ for $l > k - \kappa(k)$, ($l \ge k - \kappa(k)$, $l < k - \kappa(k)$) respectively. It is called* neutral *if its causal and anticausal.*

For a linear operator $R : H \to H$ we denote its κ–causal, anticausal and neutral components by $R_{[\kappa]} = \sum_{l \le k-\kappa(k)} R_{kl}$, $R_{[\bar{\kappa}]} = \sum_{l \ge k-\kappa(k)} R_{kl}$, $R_{[[\kappa]]} = \sum_{l = k-\kappa(k)} R_{kl}$ respectively.

Now we are ready to formulate the optimal causal filtering problem for the discrete resolution space $H = H' \times H''$, H', H'' equipped with the orthogonal resolutions of the identity $\mathbb{Q}'_K$ and $\mathbb{Q}''_K$ respectively. Let $\mathbb{H}^\kappa$ denote the space of all κ–causal continuous operators $h : H'' \to H'$ and $\hat{x}_k = Q'_k \hat{x}$, $y_k = Q''_k y$, $h_{kl} = Q'_k h Q''_l$. Then the problem is the linear estimation

$$\hat{x}_k = \sum_{l \leq k-\kappa(k)} h_{kl} y_l \tag{25}$$

minimizing the functionals

$$J_{\phi'}(h) = E|\langle \phi', D(x - hy)\rangle|^2 \to \inf_{h \in \mathbb{H}^\kappa} \tag{26}$$

for every $\phi \in H'$. Note, that this is the same as the minimization of

$$J_k(h) = E|D(x_k - \hat{x}_k)|^2 \to \inf_{h \in \mathbb{H}^\kappa} \tag{27}$$

for every $k \in \mathbb{K}$, where $x_k = Q'_k x$.

In analogy with the continuous case we will treat the relaxed problem first, replacing the condition of the continuity of h by the continuity of $h_k = Q'_k h = \sum_{l \in \mathbb{K}} h_{kl} : H'' \to H'$ for every $k \in \mathbb{K}$. The space of all linear κ–causal operators for which all the correspondent operators h_k are continuous will be denoted by $\bar{\mathbb{H}}^\kappa$. Note that because J_k are finite when R_z is bounded, the problem

$$J_k(h) = E|D(x_k - \hat{x}_k)|^2 \to \inf_{h \in \bar{\mathbb{H}}^\kappa}, \ k \in \mathbb{K} \tag{28}$$

is correctly posed. Note that if $\bar{H}', \bar{H}''$ are the completions of H', H'' in t–topology, then the space $\bar{\mathbb{H}}^\kappa$ is isomorphic to the space of all κ–causal operators from $\bar{H}''$ to $\bar{H}'$. The problem now becomes

$$J_k(h) = E|DQ'_k(x - \bar{h}y)|^2 \to \inf_{h \in \bar{\mathbb{H}}^\kappa}, \ k \in \mathbb{K}. \tag{29}$$

In analogy to Theorem 5.1 and Theorem 2.3 in [4] we have

Theorem 5.2 *Let $R_z = Ezz^*$ be continuous and R_y satisfy $P_{t_k} R_y P_{t_k} \geq \epsilon P_{t_k}$ for some $\epsilon > 0$ and for every $k \in \mathbb{K}$. Then all the solutions $\bar{h}_{\rm opt} : \bar{H}'' \to \bar{H}'$ of the problem (29) are given by*

$$\bar{h}_{\rm opt} = \sum_{k \in \mathbb{K}} Q'_k R_{xy} P''_{t_{k-\kappa(k)}} (P''_{t_{k-\kappa(k)}} R_y P''_{t_{k-\kappa(k)}})^{-1} P''_{t_{k-\kappa(k)}} + Q P''_{t_{k-\kappa(k)}}, \tag{30}$$

where $Q \in \bar{\mathbb{H}}^\kappa$ satisfies $DQ = 0$. One has

$$\inf_{h \in \bar{\mathbb{H}}^\kappa} J_k(\bar{h}) = J_k(\bar{h}_{\rm opt})$$

$$= |DP'_{t_k}[R_x - R_{xy} P''_{t_{k-\kappa(k)}} (P''_{t_{k-\kappa(k)}} R_y P''_{t_{k-\kappa(k)}})^{-1} P''_{t_{k-\kappa(k)}} R_{yx}] P'_{t_k} D^*|.$$

The proof is similar to the proof of Theorem 5.1 and is based on the calculations of $\hat{x}_k$ as optimal estimate in the subspace of H'' spanned by $y'_l = Q'_k y,\ l \leq k - \kappa(k)$. Similar to [4, Theorem 2.4] for the solution of the original problem (27) we have

Theorem 5.3 *Let the assumptions of Theorem 5.2 be satisfied. Then the problem (27) is solvable if and only if the solution $\bar{h} : \bar{H}'' \to \bar{H}'$ of (29) is κ–bounded. In this case the image of H'' under $\bar{h}$ is contained in H' and the restriction $\bar{h}|_{H''}$ is the solution of (27).*

Proof Under the assumptions of Theorem 5.3 formula (30) defines the optimal linear filters for (29). If $\bar{h}_{\rm opt}$ is κ–finite, the operator $\bar{h}_{\rm opt}|_{H''}$ is continuous and defines the weight operators for the solutions of (27). If $\bar{h}_{\rm opt}$ is not κ–bounded, the converse is also obvious from Lemma 5.3.

Note that taking finite partial sums in (30) one obtains minimizing sequences similar also in the case when $\bar{h}$ is not κ–bounded and the problem (27) is not solvable in $\mathbb{H}^\kappa$.

6 Bode–Shannon representation of the optimal filter

First we will briefly review the results on the spectral factorization of the operators which we need in order to discuss the application of Bode–Shannon theory (cf.[1],[4],[7],[8]) in our setting. The detailed discussion on various types of spectral factorization and separation of the operators can be found in [4].

Let $\mathbb{P}_T$ be a Hermitian resolution of the identity in H. As in the previous section we denote by $\bar{H}$ a t–completion of H and by $\mathbf{t}$ a discrete linearly ordered subset of $\mathbb{T}$. Let $\mathbb{G}_\mathbf{t}$ be a space of all bijective operators $\bar{G} : \bar{H} \to \bar{H}$ such that $\bar{P}_t\bar{G}\bar{P}_t$ and $\bar{P}_t\bar{G}^{-1}\bar{P}_t$ are continuous as operators from P_tH to P_tH for every $t \in \mathbf{t}$. Note, that $\mathbb{G}_\mathbf{t}$ contains the space of all causal, causally invertible operators in H.

Definition 6.1 *An operator $\bar{G} \in \mathbb{G}_\mathbf{t}$ is called* spectrally factorizable *if there exist a causal with respect to $\mathbb{P}_T$ operator $\bar{U} : \bar{H} \to \bar{H}$, such that the inverse of $\bar{U}$ exists and is causal in $\bar{H}$ and $\bar{G} = \bar{U}\bar{U}^*$, where $\bar{U}^*$ is the adjoint of $\bar{U}$.*

Let $\bar{G} = \bar{U}\bar{U}^*$ be a spectral factorization of $\bar{G}$. If $\bar{U}, \bar{U}^{-1}$ are stable (Definition 5.3), the restrictions $\bar{U}|_H, \bar{U}^{-1}|_H$ are causal and continuous in H in view of Lemma 5.3. This implies that the restriction $G = \bar{G}|_H$ is continuous in H and we can summarize it in the following

Definition 6.2 *A continuous operator $G : H \to H$ is called* strongly spectrally factorizable *if there exist a continuous causal operator $U : H \to H$ with continuous and causal inverse, such that $G = UU^*$, where U^* is the adjoint of U.*

We call operator $\bar{G} \in \mathbb{G}_\mathbf{t}$ *positive* if the operators $\bar{P}_t\bar{G}\bar{P}_t$, $\bar{P}_t\bar{G}^{-1}\bar{P}_t : P_tH \to P_tH$ are nonnegative for every $t \in \mathbf{t}$. The following is Theorems 2.5 and 2.6 in [4] (see also [5],[3]).

Theorem 6.1 (i) *Every positive operator $\bar{G} \in \mathbb{G}_{\mathfrak{t}}$ is spectrally factorizable. A causal with respect to the discrete resolution $\bar{P}_{\mathfrak{t}}$ operator $\bar{U}$ factorizing $\bar{G}$ is unique up to the multiplication from the right by a neutral unitary in $\bar{H}$ operator.*

(ii) *Let $\bar{G} \in \mathbb{G}_{\mathfrak{t}}$ and assume that the restriction $G = \bar{G}|_H$ is positive and continuous in H. Then G is strongly spectrally factorazible. A causal with respect to the discrete resolution $\mathbb{Q}_K$ operator U factorizing G is unique up to the multiplication from the right by a neutral unitary in H operator.*

It is convenient in the discrete case $(H, \mathbb{Q}_K)$ to denote by $\mathbb{G}_K$ the space of all bijective operators $\bar{G} : \bar{H} \to \bar{H}$ such that for every $k \in \mathbb{K}$ the operators $\sum_{l=0}^{k} \sum_{m=0}^{k} \bar{Q}_l \bar{G} \bar{Q}_m$, $\sum_{l=0}^{k} \sum_{m=0}^{k} \bar{Q}_l \bar{G}^{-1} \bar{Q}_m$ are continuous from $H^k = \bigoplus_{l=0}^{k} \bar{Q}_l H$ to H^k. $\bar{G} \in \mathbb{G}_K$ is called *positive* if the operators in the definition of $\mathbb{G}_K$ are nonnegative for every $k \in \mathbb{K}$.

Let $H = H' \times H''$ be equipped with the orthogonal resolution of the identity given by $Q_k = \begin{bmatrix} Q'_k & 0 \\ 0 & Q''_k \end{bmatrix}$, $Q'_k \in \mathbb{Q}'_K$, $Q''_k \in \mathbb{Q}''_K$. The κ–causal filters are given by

$$\hat{x}_k = \sum_{l=0}^{k-\kappa(k)} h_{kl} y_l,$$

where $h_{kl} : Q''_l H'' \to Q'_k H'$ are linear continuous. The corresponding filter $\bar{h} : \bar{H}'' \to \bar{H}'$ is defind by having its blocks equal to h_{kl}.

Theorem 6.2 *Assume that $R_z \in \mathbb{G}_K$, $R_y : H'' \to H'$ is positive and $\kappa \geq 0$. Then all optimal linear filters for the discrete generalized linear causal filtering problem (29) are of the form*

$$\bar{h}_{\rm opt} = [R_{xy}(U^{-1})^*]_{[\kappa]} U^{-1} + Q, \tag{31}$$

where U is a causal operator strongly factorizing R_y, $[R_{xy}(U^{-1})^*]_{[\kappa]}$ is the κ–causal component of $R_{xy}(U^{-1})^* : \bar{H}'' \to \bar{H}'$ and any $Q \in \bar{\mathbb{H}}^{\kappa}$ such that $DQ = 0$.

One has $\inf_{\bar{h} \in \bar{\mathbb{H}}^{\kappa}} J_k(\bar{h}) = J_k(\bar{h}_{\rm opt}) = |DQ_k[R_x - R_{xy} R_y^{-1} R_{yx} + [R_{xy}(U^{-1})^*]_{[\bar{\kappa}]}$ $([R_{xy}(U^{-1})^*]_{[\bar{\kappa}]})^*] Q_k D^*|$.

Proof Let $L = R_{xy}(U^{-1})^* - [R_{xy}(U^{-1})^*]_{[\kappa]}$ denote the strictly anticausal component of $R_{xy}(U^{-1})^*$. Rewriting $J_k(\bar{h})$ in analogy to (12) we have

$$\begin{aligned} E|\langle \phi', DQ'_k(x - \bar{h}y)\rangle|^2 &= \langle \phi', DQ'_k[R_x - R_{xy}\bar{h}^* - \bar{h}R_{yx} + \bar{h}R_y\bar{h}^*]Q'_k D^*\phi'\rangle \\ &= \langle \phi', DQ'_k[R_x - R_{xy}R_y^{-1}R_{yx}]Q'_k D^*\phi'\rangle + \\ &\quad \langle \phi', DQ'_k(\bar{h}U - M - L)(\bar{h}U - M - L)^* Q'_k D^*\phi'\rangle \\ &= \langle \phi', DQ'_k[R_x - R_{xy}R_y^{-1}R_{yx}]Q'_k D^*\phi'\rangle + \\ &\quad \langle \phi', DQ'_k(\bar{h}U - M)(\bar{h}U - M)^* Q'_k D^*\phi'\rangle + \\ &\quad \langle \phi', DQ'_k[(\bar{h}U - M)L^* - L(\bar{h}U - M)^*]Q'_k D^*\phi'\rangle, \end{aligned} \tag{32}$$

where M denotes $[R_{xy}(U^{-1})^*]_{[\kappa]}$, $R_{xy}(U^{-1})^* = L + M$. Note, that in the last equality in (32) the first term is constant in h, the second one is quadratic in $\bar{h}U - M$ and the third is linear. Now, an application of Lemma 5.1 yields the κ–causality of $\bar{h}U - M$

and in view of strict causality of L^* again by Lemma 5.1 the operator $(\bar{h}U - M)L^*$ is strictly κ–causal. It follows that $Q'_k(\bar{h}U - M)L^*Q'_k = 0$ if $\kappa \geq 0$. This means that the linear term in (32) vanishes and the minimum is attained if and only if the quadratic term in (32) is zero. This is the case of $\bar{h}U - M = S$ for any $S \in \bar{\mathbb{H}}^\kappa$ such that $DS = 0$. Multiplication by 0–causal operators U, U^{-1} does not change κ–causability in view of Lemma 5.1 and we obtain formula (31) with $Q = SU^{-1}$. The last formula of the theorem follows from the substitution of $\bar{h}_{\text{opt}}$ into the last expression of (32).

Corollary 6.1 *For $D = I_{H'}$ the only operator in (31) is obtain by taking $Q = 0$. This operator is called the* Bode–Shannon weight operator *and the filter (25) is called the* Bode–Shannon filter.

Corollary 6.2 *If R_z is stable, R_y^{-1} exists and is continuous in H and stable operator $R_{xy}(U^{-1})^*$ has the stable κ–causal component, then the original linear optimal causal filtering problem (27) is solvable and all optimal weight operators are the restrictions of $\bar{h}_{\text{opt}}$ in (31) to H''.*

The proof is similar to the proof of Theorem 5.3 and is left as an exercise. The reader can consult [4] for the application to the finite dimensional stationary processes, where the conditions of the Corollary 6.2 are reduced to the conditions in terms of analytic functions.

References

[1] Bode, H.W., Shannon, C.E. A simplified derivation of linear least square smoothing and prediction theory. Proc. IRE, 38, 1950, 417-425

[2] Faddeev, D.K., Faddeeva, V.N. *Calculation methods of linear algebra.* Fizmatgiz. 1963 (in Russian)

[3] Feintuch, A., Saeks, R. *System theory. A Hilbert space approach.* Academic Press, 1982

[4] Fomin, V.N. *Operator methods of the random processes linear filtering theory.* St.Petersburg University Publisher, 1995 (in Russian)

[5] Gorshkov, O.G., Fomin, V.N. Operator approaches to the time series filtering problem. Vestnik St.Petersburg University, Ser.1, 4 (22), 1993, 16-21 (in Russian)

[6] Kolmogorov, A.N. Interpolation and extrapolation of the stationary random sequences. Izvestia AN SSSR, Mathematics, 5, 1941, 3-14 (in Russian)

[7] Petrov, O.A., Fomin, V.N. *Linear filtering of random processes.* LGU publisher. 1991 (in Russian)

[8] Ruzhansky, M.V., Fomin, V.N. The optimal filter construction with the quadratic quality functional of general form. Vestnik St.Petersburg University. Ser. 1, 4 (22), 1995, 50-56 (in Russian)

[9] Wiener, N. *The Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications.* New York, 1949

V.N.Fomin
Department of Mathematics and Mechanics, St.Petersburg University
Bibliotechnaja pl. 2, 198904 St.Petersburg, Russia
e-mail: fomvn@niimm.spb.su

M.V.Ruzhansky
Mathematical Institute, University of Utrecht
P.O.Box 80.010, 3508 TA Utrecht, The Netherlands
e-mail: M.Ruzhansky@math.ruu.nl