Hokkaido Mathematical Journal Vol. 31 (2002) p. 615–628

A geometric proof of the fluctuation-dissipation theorem for the KM_2O-Langevin equation

Maciej KLIMEK, Erlendur KARLSSON, Masaya MATSUURA and Yasunori OKABE

(Received May 23, 2001)

Abstract. We give a short proof, based on the geometry of inner product spaces, of the fluctuation-dissipation theorem that asserts applicability of the Whittle-Wiggins-Robinson algorithm in the context of the KM_2O-Langevin equations also in degenerate and non-stationary cases.

Key words: non-stationarity property; degeneracy property; update property; fluctuation-dissipation theorem; KM_2O-Langevin equation.

1. Introduction

For several decades the Levinson-Durbin algorithm ([7], [2]) and its higher dimensional version the Whittle-Wiggins-Robinson algorithm ([29], [28]) have played a major role in signal processing. The history of the closely related Schur algorithm goes back even further and, in fact, its origins can be found in the realm of pure mathematics in a paper by J. Schur on complex analytic functions published in two installments in [25] and [26]. Multidimensional (or multichannel) version of Schur's algorithm was introduced nearly half a century later in [24]. From the countless papers on these algorithms we have chosen a few most relevant to this note: [1], [3]–[6], [27], [30].

On the other hand, in time series analysis, the KM_2O-Langevin equation constitutes the key element around which Yasunori Okabe and the mathematicians working with him have built their approach to the subject (see [8]–[23]). The Whittle-Wiggins-Robinson algorithm applies in calculation of the coefficients of the KM_2O-Langevin equation. It has been shown in [9] that the method works also with degenerate and non-stationary time series. The purpose of this paper is to derive in terms of elementary geometry of inner product spaces an alternative proof of the fluctuation-dissipation theorem which asserts the applicability of the Whittle-Wiggins-Robinson

---

1991 Mathematics Subject Classification : Primary 60G25; Secondary 60G12, 82C05.

algorithm in the context of the KM_2O-Langevin equation under the same assumptions as in [9]. The geometric approach offers additional insight into the properties of the KM_2O-Langevin equation.

2. Multichannel orthogonal projections

In what follows H will denote a real Hilbert space. If d is a positive integer, then the symbol H^d will stand for the Cartesian power of H of order d. Elements of H^d will be treated as column arrays. If $\mathbf{x}_j = (x_{1j}, \dots, x_{dj})^T \in H^d$ for $j = 1, \dots, m$, then by $\mathbf{Span}(\mathbf{x}_1, \dots, \mathbf{x}_m)$ we will denote the subspace of H generated by the vectors $\{x_{ij};\ i = 1, \dots, d,\ j = 1, \dots, m\}$. The elements of H^d can be multiplied from the left by $d \times d$-matrices according to the following formula:

$$C\mathbf{x} = \left(\sum_{j=1}^{d} c_{1j} x_j, \dots, \sum_{j=1}^{d} c_{dj} x_j \right)^T \in H^d,$$
$$\mathbf{x} = (x_1, \dots, x_d)^T \in H^d, \quad C = (c_{ij}) \in \mathbf{R}^{d \times d}.$$

Therefore,

$$\mathbf{y} \in (\mathbf{Span}(\mathbf{x}_1, \dots, \mathbf{x}_m))^d$$

if and only if there exist matrices $C_1, \dots, C_m \in \mathbf{R}^{d \times d}$ such that

$$\mathbf{y} = \sum_{k=1}^{m} C_k \mathbf{x}_k.$$

If in addition we assume that all the vectors x_{ij} are linearly independent, then the matrices $C_1, \dots, C_m$ are uniquely determined.

If S is a closed subspace of H, then by $\mathbf{Proj}_S : H \longrightarrow S$ we will denote the orthogonal projection of H onto S. We will use the same symbol $\mathbf{Proj}_S$ to denote componentwise orthogonal projection of H^d onto S^d:

$$\mathbf{Proj}_S(\mathbf{x}) = (\mathbf{Proj}_S(x_1), \dots, \mathbf{Proj}_S(x_d))^T,$$
$$\mathbf{x} = (x_1, \dots, x_d)^T \in H^d.$$

It will be convenient to extend the inner product to a matrix-valued bilinear mapping on $H^d \times H^d$ by the formula

$$\langle \mathbf{x}, \mathbf{y} \rangle = \begin{bmatrix} \langle x_1, y_1 \rangle & \langle x_1, y_2 \rangle & \cdots & \langle x_1, y_d \rangle \\ \langle x_2, y_1 \rangle & \langle x_2, y_2 \rangle & \cdots & \langle x_2, y_d \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle x_d, y_1 \rangle & \langle x_d, y_2 \rangle & \cdots & \langle x_d, y_d \rangle \end{bmatrix} \in \mathbf{R}^{d \times d},$$

where $\mathbf{x} = (x_1, \ldots, x_d)^T \in H^d$ and $\mathbf{y} = (y_1, \ldots, y_d)^T \in H^d$. It is clear that $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle^T$. Moreover $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ if and only if x_i is orthogonal to y_j for all possible choices of i and j. We will use the following notation:

$$\mathbf{x}^\perp = \mathbf{Span}(\mathbf{x})^\perp \subset H \quad \text{for} \quad \mathbf{x} \in H^d.$$

Note that

$$A\langle \mathbf{x}, \mathbf{y} \rangle B = \langle A\mathbf{x}, B^T\mathbf{y} \rangle, \qquad A, B \in \mathbf{R}^{d \times d}.$$

The above notation allows us to characterize orthogonal projections as follows.

The Multichannel Projection Formula *Let* $\mathbf{x}_j = (x_{1j}, \ldots, x_{dj})^T \in H^d$ *for* $j = 1, \ldots, m$, *and let*

$$S = \mathbf{Span}(\mathbf{x}_1, \ldots, \mathbf{x}_m).$$

Suppose that $\mathbf{y} = (y_1, \ldots, y_d)^T \in H^d$ *and* $C_1, \ldots, C_m$ *are real* $d \times d$*-matrices. Then*

$$\mathbf{Proj}_S(\mathbf{y}) = \sum_{k=1}^{m} C_k \mathbf{x}_k \tag{1}$$

if and only if

$$\begin{bmatrix} \langle \mathbf{x}_1, \mathbf{x}_1 \rangle & \langle \mathbf{x}_1, \mathbf{x}_2 \rangle & \cdots & \langle \mathbf{x}_1, \mathbf{x}_m \rangle \\ \langle \mathbf{x}_2, \mathbf{x}_1 \rangle & \langle \mathbf{x}_2, \mathbf{x}_2 \rangle & \cdots & \langle \mathbf{x}_2, \mathbf{x}_m \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \mathbf{x}_m, \mathbf{x}_1 \rangle & \langle \mathbf{x}_m, \mathbf{x}_2 \rangle & \cdots & \langle \mathbf{x}_m, \mathbf{x}_m \rangle \end{bmatrix} \begin{bmatrix} C_1^T \\ C_2^T \\ \vdots \\ C_m^T \end{bmatrix} = \begin{bmatrix} \langle \mathbf{x}_1, \mathbf{y} \rangle \\ \langle \mathbf{x}_2, \mathbf{y} \rangle \\ \vdots \\ \langle \mathbf{x}_m, \mathbf{y} \rangle \end{bmatrix}. \tag{2}$$

If, additionally, we assume that the vectors $\{x_{ij}\,;\, i = 1, \ldots, d,\ j = 1, \ldots, m\}$ *are linearly independent, then the choice of* $\mathbf{y}$ *determines uniquely the matrices* $C_1, \ldots, C_m$.

Proof. The equation (2) can be re-written as follows:

$$\left\langle \mathbf{x}_j, \mathbf{y} - \sum_{k=1}^{m} C_k \mathbf{x}_k \right\rangle = 0 \quad \text{for} \quad j = 1, \ldots, m.$$

But this means precisely the same as (1). To see that the second statement is true, note that the leftmost matrix in (2) is in fact the Gram matrix for the vectors

$$x_{11}, \ldots, x_{d1}, x_{12}, \ldots, x_{d2}, \ldots, x_{1m}, \ldots, x_{dm}$$

and hence it is non-singular under the assumption of linear independence. □

Corollary 1 *If* $\mathbf{x}, \mathbf{y} \in H^d$ *and* C *is a real* $d \times d$*-matrix, then*

$$C\langle \mathbf{x}, \mathbf{x}\rangle = \langle \mathbf{y}, \mathbf{x}\rangle \iff \mathbf{Proj}_{\mathbf{Span}(\mathbf{x})}(\mathbf{y}) = C\mathbf{x}. \tag{3}$$

In particular, if the components of the d*-tuple* $\mathbf{x}$ *are linearly independent, then*

$$\mathbf{Proj}_{\mathbf{Span}(\mathbf{x})}(\mathbf{y}) = \langle \mathbf{y}, \mathbf{x}\rangle \langle \mathbf{x}, \mathbf{x}\rangle^{-1}\mathbf{x}.$$

Our main geometric tool will be the following elementary property of orthogonal projections.

The Update Property *Let* S *be a closed subspace of the Hilbert space* H *and let* $\mathbf{x}, \mathbf{y} \in H^d$*. Define*

$$\hat{\mathbf{x}} = \mathbf{Proj}_{S^\perp}(\mathbf{x}) \quad \textit{and} \quad \hat{\mathbf{y}} = \mathbf{Proj}_{S^\perp}(\mathbf{y})$$

and

$$S_{\mathbf{y}} = S + \mathbf{Span}(\mathbf{y}).$$

Let π *denote an arbitrary* $d \times d$*-matrix. The following conditions are equivalent:*

(UP1) $\pi\langle \hat{\mathbf{y}}, \hat{\mathbf{y}}\rangle = \langle \hat{\mathbf{x}}, \hat{\mathbf{y}}\rangle$;

(UP2) $\mathbf{Proj}_{S_{\mathbf{y}}}(\mathbf{x}) = [\mathbf{Proj}_S(\mathbf{x}) - \pi\mathbf{Proj}_S(\mathbf{y})] + \pi\mathbf{y}$;

(UP3) $\mathbf{Proj}_{S_{\mathbf{y}}}(\mathbf{x}) - \pi\mathbf{y} \in S^d$;

(UP4) $\mathbf{Proj}_{S_{\mathbf{y}}^\perp}(\mathbf{x}) = \hat{\mathbf{x}} - \pi\hat{\mathbf{y}}$.

Proof. The implications (UP2)$\Longrightarrow$(UP4)$\Longrightarrow$(UP1) and (UP2)$\Longrightarrow$(UP3) are obvious. Corollary 1 applied to $\hat{\mathbf{y}}$, $\hat{\mathbf{x}}$ and π, yields the implication (UP1)$\Longrightarrow$(UP2).

Finally, we have (UP3)$\Longrightarrow$(UP2) because if (UP3) is assumed, then

$$\begin{aligned}\mathbf{Proj}_{S_{\mathbf{y}}}(\mathbf{x}) - \pi\mathbf{y} &= \mathbf{Proj}_{S}\left(\mathbf{Proj}_{S_{\mathbf{y}}}(\mathbf{x}) - \pi\mathbf{y}\right)\\ &= \mathbf{Proj}_{S}(\mathbf{x}) - \pi\mathbf{Proj}_{S}(\mathbf{y}).\end{aligned}$$

□

3. The KM$_2$O-Langevin equations

Let X be a finite sequence of elements of H^d identified with the function

$$X : \{N_-, \ldots, N_+\} \longrightarrow H^d,$$

where $N_-, N_+ \in \mathbf{Z}$ and $N_+ - N_- \geq 2$. For any two integers p, q such that $N_- \leq p \leq q \leq N_+$ we define

$$M_p^q = \mathbf{Span}(X(p), X(p+1), \ldots, X(q-1), X(q)).$$

Given $m \in \{N_-, \ldots, N_+\}$ define

$$\begin{aligned}&\nu_+(m,0) = X(m),\\ &\nu_+(m,n) = \mathbf{Proj}_{(M_m^{m+n-1})^\perp}(X(m+n)),\\ &\qquad\qquad\qquad\qquad \text{if } \ m \neq N_+ \ \text{ and } \ n = 1, \ldots, N_+ - m;\\ &\nu_-(m,0) = X(m),\\ &\nu_-(m,n) = \mathbf{Proj}_{(M_{m-n+1}^m)^\perp}(X(m-n)),\\ &\qquad\qquad\qquad\qquad \text{if } \ m \neq N_- \ \text{ and } \ n = 1, \ldots, m - N_-.\end{aligned}$$

It is to be noted that

$$M_m^{m+n} = \mathbf{Span}(\nu_+(m,0), \nu_+(m,1), \ldots, \nu_+(m,n)),\\ n = 0, \ldots, N_+ - m$$

and

$$\langle \nu_+(m,l), \nu_+(m,n) \rangle = 0, \qquad 0 \leq l \neq n \leq N_+ - m.$$

In other words, under the additional assumption that the vectors

$$\{X_i(j)\,;\, i = 1, \ldots, d, \ j = N_-, \ldots, N_+\} \tag{4}$$

are linearly independent, where $X(j) = (X_1(j), \ldots, X_d(j))^T$, one can view the vectors

$$\{\nu_+(m,0), \nu_+(m,1), \ldots, \nu_+(m, N_+ - m)\}$$

as a result of the block Gram-Schmidt orthogonalization of the vectors

$$\{X(m), X(m+1), \ldots, X(N_+)\}.$$

Similarly,

$$M_{m-n}^m = \mathbf{Span}(\nu_-(m,0), \nu_-(m,1), \ldots, \nu_-(m,n)),\\ n = 0, \ldots, m - N_-$$

and

$$\langle \nu_-(m,l), \nu_-(m,n)\rangle = 0, \qquad 0 \leq l \neq n \leq m - N_-.$$

In other words, the assumption of linear independence implies that the vectors

$$\{\nu_-(m,0), \nu_-(m,1), \ldots, \nu_-(m, m - N_-)\}$$

result from the block Gram-Schmidt orthogonalization of the vectors

$$\{X(m), X(m-1), \ldots, X(N_-)\}.$$

If $n = 1, \ldots, N_+ - m$, then there exists a system of matrices

$$\{\gamma_+(m,n,k)\,;\, k = 0, \ldots, n-1\} \subset \mathbf{R}^{d\times d}$$

such that

$$\mathbf{Proj}_{M_m^{m+n-1}}(X(m+n)) = -\sum_{k=0}^{n-1} \gamma_+(m,n,k) X(m+k).$$

Similarly, if $n = 1, \ldots, m - N_-$, then there exist matrices

$$\{\gamma_-(m,n,k)\,;\, k = 0, \ldots, n-1\} \subset \mathbf{R}^{d\times d}$$

such that

$$\mathbf{Proj}_{M_{m-n+1}^m}(X(m-n)) = -\sum_{k=0}^{n-1} \gamma_-(m,n,k) X(m-k).$$

Note that if we assume that the vectors (4) are linearly independent, then the matrices $\gamma_\pm(m,n,k)$ are uniquely determined.

It is convenient to introduce the function

$$R : \{N_-, \dots, N_+\} \times \{N_-, \dots, N_+\} \longrightarrow \mathbf{R}^{d\times d}$$

given by the formula

$$R(m,n) = \langle X(m), X(n)\rangle$$

for all $m,n \in \{N_-, \dots, N_+\}$. We will say that X is *weakly stationary*[1] if there exists a function

$$r : \{N_- - N_+, \dots, N_+ - N_-\} \longrightarrow \mathbf{R}^{d\times d}$$

such that for all $m,n \in \{N_-, \dots, N_+\}$

$$R(m,n) = r(m-n).$$

Observe that if both linear independence of (4) and weak stationarity of X are assumed, then the coefficients $\gamma_\pm(m,n,k)$ do not depend on the choice of m.

In the general case, we can write the equations

$$X(m \pm n) = -\sum_{k=0}^{n-1} \gamma_\pm(m,n,k) X(m \pm k) + \nu_\pm(m,n), \tag{5}$$

where $N_- \le m \le N_+ - n$ in the "plus" case and $N_- + n \le m \le N_+$ in the "minus" case. In the case when $m = 0$ and $N_- = -N_+$, these are called in the Japanese literature *the KM_2O-Langevin equations*, whereas the vectors $\nu_+(0,n)$ (resp. $\nu_-(0,n)$) are referred to as *the forward (resp. backward) KM_2O-Langevin fluctuation flows* (see [8]–[23]).

It is useful to define

$$V_\pm(m,n) = \langle \nu_\pm(m,n), \nu_\pm(m,n)\rangle ,$$

for $n = 0, 1, \dots, \pm(N_\pm - m)$. Observe that

$$V_\pm(m,0) = R(m,m) \tag{6}$$

[1]This definition is consistent with its probabilistic counterpart: if Ω is the event space with a probability measure P, then as H we take the orthogonal complement of the constant function 1 in the space of square integrable random variables $L^2(\Omega, P)$. All random variables in H have expected value 0.

and

$$V_\pm(m,n) = \langle \nu_\pm(m,n), X(m \pm n)\rangle = R(m \pm n, m \pm n) \\ + \sum_{k=0}^{n-1} \gamma_\pm(m,n,k) R(m \pm k, m \pm n) \tag{7}$$

for $n = 1, \ldots, \pm(N_\pm - m)$. In particular, if X is weakly stationary, then $V_\pm(m,n)$ is independent of m.

Theorem 1 *For any choice of $\gamma_\pm(m,n,k) \in \mathbf{R}^{d\times d}$, where $N_- + n \leq m \leq N_+ - n$, for which the KM$_2$O-Langevin equations* (5) *are satisfied the following relationships hold*:

$$\delta_\pm(m,1)V_\mp(m,0) = -R(m \pm 1, m), \tag{8}$$

$$V_\pm(m,n) = (I - \delta_\pm(m,n)\delta_\mp(m \pm n, n))\, V_\pm(m \pm 1, n-1), \tag{9}$$

$$\delta_\pm(m,n)V_\mp(m \pm n \mp 1, n-1) \\ = -R(m \pm n, m) - \sum_{k=0}^{n-2} \gamma_\pm(m \pm 1, n-1, k) R(m \pm k \pm 1, m) \tag{10}$$

for all admissible values of m *and* n, *where*

$$\delta_\pm(m,n) = \gamma_\pm(m,n,0).$$

Furthermore, there exist matrix coefficients $\gamma_\pm(m,n,k) \in \mathbf{R}^{d\times d}$ such that the KM$_2$O-Langevin equation (5) *is satisfied and*

$$\gamma_\pm(m,n,k) = \gamma_\pm(m \pm 1, n-1, k-1) \\ + \delta_\pm(m,n)\gamma_\mp(m \pm n \mp 1, n-1, n-k-1). \tag{11}$$

Proof. Formula (8) follows directly from (5) and the definition of $\nu_\pm$. Now we will check the "plus" case of (9) and (10). The "minus" case is similar.

Because (5) is fulfilled, we have

$$\mathbf{Proj}_{M_m^{m+n-1}}(X(m+n)) + \delta_+(m,n)X(m) \\ = -\sum_{k=1}^{n-1} \gamma_+(m,n,k) X(m+k) \in \left(M_{m+1}^{m+n-1}\right)^d.$$

Therefore (UP3) holds with

$$\begin{cases} S &= M_{m+1}^{m+n-1}, \\ \mathbf{x} &= X(m+n), \\ \mathbf{y} &= X(m), \\ \pi &= -\delta_+(m,n). \end{cases} \tag{12}$$

Similarly,

$$\begin{aligned} &\mathbf{Proj}_{M_{m+1}^{m+n}}(X(m)) + \delta_-(m+n,n)X(m+n) \\ &\quad = -\sum_{k=1}^{n-1} \gamma_-(m+n,n,k)X(m+n-k) \in \left(M_{m+1}^{m+n-1}\right)^d \end{aligned}$$

and hence (UP3) holds also with

$$\begin{cases} S &= M_{m+1}^{m+n-1}, \\ \mathbf{x} &= X(m), \\ \mathbf{y} &= X(m+n), \\ \pi &= -\delta_-(m+n,n). \end{cases} \tag{13}$$

We can check (9) as follows:

$$\begin{aligned} &V_+(m,n) \\ &= \langle \nu_+(m,n), X(m+n)\rangle \\ &= \langle \nu_+(m+1,n-1) + \delta_+(m,n)\nu_-(m+n-1,n-1), X(m+n)\rangle, \\ &\qquad\qquad\qquad\qquad \text{because of (UP4)}, \\ &= \langle \nu_+(m+1,n-1), X(m+n)\rangle \\ &\quad + \delta_+(m,n)\langle \nu_-(m+n-1,n-1), X(m+n)\rangle \\ &= \langle \nu_+(m+1,n-1), \nu_+(m+1,n-1)\rangle \\ &\quad + \delta_+(m,n)\langle \nu_-(m+n-1,n-1), \nu_+(m+1,n-1)\rangle, \\ &\qquad \text{(since } X(m+n) - \nu_+(m+1,n-1) \in \left(M_{m+1}^{m+n-1}\right)^d, \\ &= V_+(m+1,n-1) - \delta_+(m,n)\delta_-(m+n,n)V_+(m+1,n-1), \\ &\qquad\qquad\qquad \text{because of (UP1) combined with (13).} \end{aligned}$$

Similarly

$$
\begin{aligned}
&\langle \nu_+(m+1,n-1), \nu_-(m+n-1,n-1)\rangle \\
&\quad = \langle \nu_+(m+1,n-1), X(m)\rangle \\
&\quad = \Big\langle X\left((m+1)+(n-1)\right) \\
&\qquad\qquad + \sum_{k=0}^{n-2} \gamma_+(m+1,n-1,k)X(m+k+1), X(m)\Big\rangle \\
&\quad = R(m+n,m) + \sum_{k=0}^{n-2} \gamma_+(m+1,n-1,k)R(m+k+1,m),
\end{aligned}
$$

which — in view of (UP1) used with (12) — gives (10).

Now we move to the second statement of the theorem. Reading the formula (UP2) in terms of $X(m), X(m+1), \dots, X(m+n-1)$ in the "plus" case, and in terms of $X(m-n+1), X(m-n+2), \dots, X(m)$ in the "minus" case, we can rewrite it in the following equivalent form:

$$
\begin{aligned}
&-\sum_{k=1}^{n-1} \gamma_\pm(m,n,k)X(m\pm k) - \delta_\pm(m,n)X(m) \\
&\quad = \Big[-\sum_{k=0}^{n-2} \gamma_\pm(m\pm 1,n-1,k)X(m\pm k\pm 1) \\
&\qquad\quad - \delta_\pm(m,n)\sum_{k=0}^{n-2} \gamma_\mp(m\pm n\mp 1,n-1,k)X(m\pm n\mp k\mp 1)\Big] \\
&\qquad\qquad - \delta_\pm(m,n)X(m).
\end{aligned}
$$

Hence, we can construct the coefficients by recursion with respect to n. We start with $\delta_\pm(m,1)$. If $\gamma_\pm(m,n-1,k)$ are known, we can find a suitable $\delta_\pm(m,n)$ because of Corollary 1. Then the above formula allows us to define $\gamma_\pm(m,n,k)$ with $k>0$. □

Corollary 2 *Under the above assumptions*

$$\nu_\pm(m,n) = \nu_\pm(m\pm 1,n-1) + \delta_\pm(m,n)\nu_\mp(m\pm n\mp 1,n-1)$$

for all admissible m *and* n.

The next statement is the converse of Theorem 1.

Theorem 2 *If the matrices* $\gamma_\pm(m,n,k) \in \mathbf{R}^{d\times d}$ *are chosen so that they satisfy* (8), (10) *and* (11), *then the* KM_2O*-Langevin equations* (5) *hold true.*

Proof. We can use induction with respect to n. The case when $n = 1$, follows directly from (8), (6) and Corollary 1. Suppose now that n is such that (5) is satisfied with $\gamma_\pm(m, n-1, k)$. Then

$$\begin{aligned}
&\langle \nu_+(m+1, n-1), \nu_-(m+n-1, n-1)\rangle \\
&\quad = \langle \nu_+(m+1, n-1), X(m)\rangle \\
&\quad = R(m+n, m) + \sum_{k=0}^{n-2} \gamma_+(m+1, n-1, k) R(m+k+1, m), \\
&\qquad\qquad \text{because of the induction hypothesis,} \\
&\quad = -\delta_+(m,n) V_-(m+n-1, n-1), \ \text{ because of (8) and (10).}
\end{aligned}$$

Consequently (UP1) holds with

$$\begin{aligned}
S &= M_{m+1}^{m+n-1}, \\
\mathbf{x} &= X(m+n), \\
\mathbf{y} &= X(m), \\
\pi &= -\delta_+(m,n).
\end{aligned}$$

Therefore, in view of (UP2) and (11), the equation (5) is satisfied with $\gamma_+(m,n,k)$. Similarly, $\gamma_-(m,n,k)$ satisfy (5). □

References

[1] Benveniste A. and Chaure C., *AR and ARMA identification algorithms of Levinson type: an innovation approach*. IEEE Trans. Automat. Contr. Vol. AC-26, No.6, December 1981, 1243–1261.

[2] Durbin J., *The fitting of time series models*. Rev. Int. Inst. Stat. **28** (1960), 233–244.

[3] Karlsson E. and Hayes M.H., *Least Squares ARMA Modeling of Linear Time-Varying Systems: Lattice Filter Structures and Fast RLS Algorithms*. IEEE Trans. Acoust., Speech, and Signal Processing, Vol. ASSP-35, No.7, July 1987, 994–1014.

[4] Karlsson E., *Least Squares ARMA Modeling of Linear Time-Varying Systems — Lattice Filter Structures and Fast RLS Algorithms*. Ph.D. Thesis, Georgia Institute of Technology, Atlanta, 1987.

[5] Karlsson E., *A geometric perspective on the Schur algorithm*. In: Fourth Annual ASSP Workshop on Spectrum Estimation and Modeling, IEEE, Inc., 1988, 289–292.

[6] Lee D., Morf M. and Friedlander B., *Recursive Least Squares Ladder Estimation Algorithms*. IEEE Trans. Acoust., Speech, and Signal Processing, Vol. ASSP-29, No.3, June 1981, 627–641.

[7] Levinson N., *The Wiener rms error criterion in filter design and prediction*. J. Math. Phys. **25** (1947), 261–278.

[8] Matsuura M. and Okabe Y., *On a non-linear prediction problem for one-dimensional stochastic processes*. Japan. J. of Math. **27** (2001), 51–112.

[9] Matsuura M. and Okabe Y., *On the theory of KM_2O-Langevin equations for non-stationary and degenerate flows*. To appear in J. Math. Soc. Japan.

[10] Nakano Y., *On a causal analysis of economic time series*. Hokkaido Math. J. **24** (1995), 1–35.

[11] Okabe Y., *On a stochastic difference equation for the multi-dimensional weakly stationary process with discrete time*. In: Prospects of Algebraic Analysis, (Ed. by M. Kashiwara and T. Kawai), Academic Press, 1988, 601–645.

[12] Okabe Y. and Nakano Y., *The theory of KM_2O-Langevin equations and its applications to data analysis* (I): *Stationary analysis*. Hokkaido Math. J. **20** (1991), 45–90.

[13] Okabe Y., *Application of the theory of KM_2O-Langevin equations to the linear prediction problem for the multi-dimensional weakly stationary time series*. J. Math. Soc. Japan **45** (1993), 277–294.

[14] Okabe Y., *A new algorithm derived from the view-point of the fluctuation-dissipation principle in the theory of KM_2O-Langevin equations*. Hokkaido Math. J. **22** (1993), 199–209.

[15] Okabe Y. and Inoue A., *The theory of KM_2O-Langevin equations and its applications to data analysis* (II): *Causal analysis* (1). Nagoya Math. J. **134** (1994), 1–28.

[16] Okabe Y., *Langevin equations and causal analysis*. Amer. Math. Soc. Transl. **161** (1994), 19–50.

[17] Okabe Y. and Ootsuka T., *Application of the theory of KM_2O-Langevin equations to the non-linear prediction problem for the one-dimensional strictly stationary time series*. J. Math. Soc. Japan **47** (1995), 349–367.

[18] Okabe Y., *Nonlinear time series analysis based upon the fluctuation-dissiapation theorem*. Nonlinear Analysis, Theory, Methods and Applications, Vol. 30, No.4 (1997), 2249–2260.

[19] Okabe Y. and Yamane T., *The theory of KM_2O-Langevin equations and its applications to data analysis* (III): *Deterministic analysis*. Nagoya Math. J. **152** (1998), 175–201.

[20] Okabe Y., *On the theory of KM_2O-Langevin equations for stationary flows* (1): *characterization theorem*. J. Math. Soc. Japan **51** (1999), 817–841.

[21] Okabe Y., *On the theory of KM_2O-Langevin equations for stationary flows* (2): *construction theorem*. Acta Applicandae Mathematicae **63** (2000), 307–322.

[22] Okabe Y. and Matsuura M., *On the theory of KM_2O-Langevin equations for stationary flows* (3): *extension theorem*. Hokkaido Math. J. **29** (2000), 369–382.

[23] Okabe Y. and Kaneko A., *On a non-linear prediction analysis for multidimensional stochastic processes with its applications to data analysis*. Hokkaido Math. J. **29** (2000), 601–657.

[24] Rissanen J., *Algorithms for triangular decomposition of block Hankel and Toeplitz matrices with application to factoring positive matrix polynomials*. Math. Comp. **27** (1973), 147–154.

[25] Schur J., *Über Potenzereihen, die im Innern des Einheitskreises beschränkt sind*. J. reine angew. Math. **147** (1917), 205–232.

[26] Schur J., *Über Potenzereihen, die im Innern des Einheitskreises beschränkt sind*. (Fortsetzung), J. reine angew. Math. **148** (1918), 122–145.

[27] Shensa M.J., *Recursive Least Squares Lattice Algorithms – A Geometric Approach*. IEEE Trans. on Automat. Control, Vol. AC-26, No.3, June 1981, 695–702.

[28] Wiggins R.A. and Robinson E.A., *Recursive solution to the multichannel filtering problem*. J. Geophys. Res. **70** (1965), 1885–1891.

[29] Whittle P., *On the fitting of multivariate autoregressions, and the approximate canonical factorization of a spectral density matrix*. Biometrika **50** (1963), 129–134.

[30] Yagle A.E., *Fast algorithms for structured matrices in signal processing*. In: Handbook of Statistics, Vol. 10, Ed. by N.K. Bose and C.R. Rao, North-Holland, Amsterdam 1993, 933–972.

Maciej Klimek
Department of Mathematics
Uppsala University
P.O. Box 480, 751 06 Uppsala
Sweden
E-mail: Maciej.Klimek@math.uu.se

Erlendur Karlsson
Ericsson Radio Systems AB
Torshamnsgatan 23, 164 80 Stockholm
Sweden
E-mail: Erlendur.Karlsson@era.ericsson.se

Masaya Matsuura
Department of Mathematical Informatics
Graduate School of Information
Science and Technology
University of Tokyo
Bunkyo-ku, Tokyo 113-8656
Japan
E-mail: masaya@mist.i.u-tokyo.ac.jp

Yasunori Okabe
Department of Mathematical Informatics
Graduate School of Information
Science and Technology
University of Tokyo
Bunkyo-ku, Tokyo 113-8656
Japan
E-mail: okabe@mist.i.u-tokyo.ac.jp