J. Math. Soc. Japan
Vol. 47, No. 2, 1995

Application of the theory of KM_2O-Langevin equations to the non-linear prediction problem for the one-dimensional strictly stationary time series

Dedicated to Professor Kiyoshi Ito on his seventy-seven birthday

By Yasunori OKABE and Takashi OOTSUKA

(Received Feb. 5, 1993)
(Revised Aug. 16, 1993)

§ 1. Introduction.

We are inspired by Masani-Wiener's work ([4]) of the non-linear prediction problem of a one-dimensional discrete time strictly stationary process. The purpose of the present paper is to give computable algorithms for the non-linear predictor by applying the theory of KM_2O-Langevin equations.

We have already applied in [7] the theory of KM_2O-Langevin equations to the linear prediction problem for the multi-dimensional weakly stationary time series and given a refinement of Wiener-Masani's work in [13], [14] and [3] by obtaining computable algorithms for the linear predictor. The results in [7] play supplementary but useful roles in the present approach to the non-linear problem, as will be explained.

Let $\boldsymbol{X}=(X(n); n\in\mathbb{Z})$ be a real-valued strictly stationary time series on a probability space $(\Omega, \mathcal{B}, P)$ with mean zero. We shall impose the following two hypotheses which are the same as in [4]:

(H.1) *$\boldsymbol{X}$ is essentially bounded, i.e., there exists a positive constant $C>0$ such that $|X(n)(\omega)|\leqq C$ for any $n\in\mathbb{Z}$ and almost all $\omega\in\Omega$;*

(H.2) *For any distinct integers $n_1, n_2, \cdots, n_k$ $(k\in\mathbb{N})$ the spectrum of the distribution function of the k-dimensional random variable ${}^t(X(n_1), X(n_2), \cdots, X(n_k))$ has positive Lebesgue measure.*

The non-linear predictor $\hat{X}(\nu)$ of the future $X(\nu)$, $\nu>0$, on the basis of the present and past $X(l)$, $l\leqq 0$, is defined by

$$\hat{X}(\nu)=E(X(\nu)\,|\,\sigma(X(l); l\leqq 0)).$$

Masani and Wiener ([4]) have obtained a representation for the non-linear

This research was partially supported by Grant-in-Aid for Science Research No. 03452011, the Ministry of Education, Science and Culture, Japan.

predictor as follows:

$$E(X(\nu)|\sigma(X(l);l\leqq 0)) = \operatorname*{l.i.m.}_{n\to\infty} Q_n(X(0), X(-1), \cdots, X(-m_n)), \tag{1.1}$$

where, for each $n\in\mathbb{N}$, m_n is a nonnegative integer depending on n, and Q_n is a real polynomial in m_n+1 variables whose coefficients can be theoretically calculated in terms of the moments of the time series $\boldsymbol{X}$.

However, as Kallianpur has given some comments in [**12**], the representation (1.1) of the non-linear predictor lacks for computable algorithm which is fit for the application to applied science, because the determination of the coefficients of the polynomials Q_n involves the calculation of the determinants of matrices of different sizes, coming from their method of Schmidt's orthogonalization. On the other hand, Masani and Wiener have suggested in [**4**] that certain computable algorithm for the non-linear predictor may be obtained by means of the linear predictor for a suitably defined, infinite-dimensional, weakly stationary time series.

Following their suggestion, we shall derive an $\mathbb{R}^\infty$-valued weakly stationary time series $\mathcal{X}=(\mathcal{X}(n); n\in\mathbb{Z})$ and consider the d_q+1-dimensional subprocesses $\boldsymbol{X}^{(q)}=(X^{(q)}(n); n\in\mathbb{Z})$ generated by the first d_q+1-components of $\mathcal{X}$. We remark that $d_1=0$, d_q is increasing to ∞ as $q\to\infty$ and $\boldsymbol{X}^{(1)}=\boldsymbol{X}$. According to the theory of KM_2O-Langevin equations ([**5**], [**6**], [**9**]), for each $q\in\mathbb{N}$, the linear predictor for the d_q+1-dimensional subprocess $\boldsymbol{X}^{(q)}$ can be calculated from the KM_2O-Langevin data $\mathcal{LD}(\boldsymbol{X}^{(q)})$ which, corresponding to the fluctuation-dissipation theorem, is obtained from the computable algorithm in terms of the correlation function of $\boldsymbol{X}^{(q)}$. By obtaining a new algorithm computing the KM_2O-Langevin data $\mathcal{LD}(\boldsymbol{X}^{(q)})$ from the KM_2O-Langevin data $\mathcal{LD}(\boldsymbol{X}^{(q-1)})$ $(q=2, 3, \cdots)$, we can practically solve the non-linear prediction problem for the original time series $\boldsymbol{X}$, because the non-linear predictor for $\boldsymbol{X}$ can be obtained as the limit as $q\to\infty$ of the first component of the linear predictors for $\boldsymbol{X}^{(q)}$.

Now we shall explain the contents of this paper. In § 2, according to [**5**] and [**9**], we shall recall and rearrange the theory of KM_2O-Langevin equations for a d-dimensional weakly stationary time series $\boldsymbol{Z}=(Z(n); |n|\leqq N)$, where d, N are fixed natural numbers. In particular, we shall introduce the KM_2O-Langevin data $\mathcal{LD}(\boldsymbol{Z})$ associated with the time series $\boldsymbol{Z}$ which consists of the triplet of the forward and backward KM_2O-Langevin delay functions, the forward and backward KM_2O-Langevin partial correlation functions, and the forward and backward KM_2O-Langevin fluctuation functions. The KM_2O-Langevin data $\mathcal{LD}(\boldsymbol{Z})$, together with the forward and backward KM_2O-Langevin forces, will determine the forward and backward KM_2O-Langevin equations describing the time evolution of the time series $\boldsymbol{Z}$. We can obtain a concrete expression for the linear predictor for the time series $\boldsymbol{Z}$ in terms of the KM_2O-Langevin data

$\mathcal{L}\mathcal{D}(\boldsymbol{Z})$. Furthermore, associated with a d-dimensional weakly stationary time series $\boldsymbol{Z}=(Z(n)\,;\,n\in\mathbb{Z})$, we can construct the KM_2O-Langevin data $\mathcal{L}\mathcal{D}(\boldsymbol{Z})$.

§3 will develop the theory of the KM_2O-Langevin equations and obtain a new formula between the KM_2O-Langevin data $\mathcal{L}\mathcal{D}(\boldsymbol{Z})$ and the KM_2O-Langevin data $\mathcal{L}\mathcal{D}(\boldsymbol{Y})$, where the time series $\boldsymbol{Y}$ is a $d^{(1)}$-dimensional local and weakly stationary time series generated by the first $d^{(1)}$-components of the series $\boldsymbol{Z}$ $(1\leqq d^{(1)}<d)$.

In the last section, we shall return to the real-valued strictly stationary time series $\boldsymbol{X}=(X(n)\,;\,n\in\mathbb{Z})$ with mean zero satisfying conditions (H.1) and (H.2). By modifying the idea in Masani and Wiener ([**4**]), we shall derive an $\mathbb{R}^\infty$-valued weakly stationary time series $\boldsymbol{\mathcal{X}}=(\mathcal{X}(n)\,;\,n\in\mathbb{Z})$ and consider the d_q+1-dimensional subprocesses $\boldsymbol{X}^{(q)}=(X^{(q)}(n)\,;\,n\in\mathbb{Z})$ generated by the first d_q+1-components of $\boldsymbol{\mathcal{X}}$. We remark that the first components of $X^{(q)}(n)$ are equal to $X(n)$ $(q\in\mathbb{N},\ n\in\mathbb{Z})$ and the construction of the time series $\boldsymbol{X}^{(q)}$ with dimension d_q+1 is fit for the application to data analysis. Applying the results in §3 to these time series $\boldsymbol{X}^{(q)}$, we shall obtain an algorithm computing the KM_2O-Langevin data $\mathcal{L}\mathcal{D}(\boldsymbol{X}^{(q)})$ from the KM_2O-Langevin data $\mathcal{L}\mathcal{D}(\boldsymbol{X}^{(q-1)})$ $(q=2, 3, \cdots)$. Thus the non-linear prediction problem for the original real valued strictly stationary time series $\boldsymbol{X}$ can be practically solved as follows:

$$
\begin{aligned}
(1.2)\qquad & E(X(\nu)\,|\,\sigma(X(l)\,;\,l\leqq 0)) \\
&= \text{the first component of } \underset{N,q\to\infty}{\text{l.i.m.}} \sum_{k=0}^{N} Q_+(\boldsymbol{X}^{(q)})(N+\nu,\,N\,;\,N-k)X^{(q)}(-k),
\end{aligned}
$$

where, for each $q\in\mathbb{N}$, the $M(d_q+1\,;\,\mathbb{R})$-valued function $Q_+(\boldsymbol{X}^{(q)})(\cdot,\,*\,;\,\star)$ is called the forward prediction function associated with the time series $\boldsymbol{X}^{(q)}$ in the theory of the KM_2O-Langevin equations, which can be recursively calculated from the KM_2O-Langevin data $\mathcal{L}\mathcal{D}(\boldsymbol{X}^{(q)})$. By using the results in [**7**], furthermore, we can theoretically obtain an algorithm for the limit as $N\to\infty$ of the forward prediction functions $Q_+(\boldsymbol{X}^{(q)})(N+\nu,\,N\,;\,N-k)$ for any fixed $q,\,\nu\in\mathbb{N}$, $k\in\mathbb{N}^*$ $(\equiv\mathbb{N}\cup\{0\})$.

As the application of the theory of KM_2O-Langevin equations to data analysis, we are going to develop a new project of the stationary, causal and prediction analysis ([**9**], [8], [**10**]).

The authors would like to thank the referee for valuable advices.

§2. The theory of KM_2O-Langevin equations.

We shall recall the theory of KM_2O-Langevin equations from [**5**], [**9**].

[2.1] Let d and N be any natural numbers. Let $\boldsymbol{Z}=(Z(n)\,;\,|n|\leqq N)$ be any d-dimensional real-valued local and weakly stationary time series on a probability space $(\Omega, \mathcal{B}, P)$ with covariance matrix function R^{Z}:

$$R^{Z}(n)=E(Z(n)^{t}Z(0)) \qquad (|n|\leqq N). \tag{2.1}$$

Then we define, for each $n\in\mathbb{N}$, $1\leqq n\leqq N$, two block Toeplitz matrices $T_{n}^{+}(\boldsymbol{Z})$, $T_{n}^{-}(\boldsymbol{Z})\in M(nd\,;\mathbb{R})$ by

$$T_{n}^{\pm}(\boldsymbol{Z})=\begin{pmatrix} R^{Z}(0) & R^{Z}(\pm 1) & \cdots & R^{Z}(\pm(n-1)) \\ R^{Z}(\mp 1) & R^{Z}(0) & \cdots & R^{Z}(\pm(n-2)) \\ \vdots & \vdots & \ddots & \vdots \\ R^{Z}(\mp(n-1)) & R^{Z}(\mp(n-2)) & \cdots & R^{Z}(0) \end{pmatrix}. \tag{2.2$_\pm$}$$

It is to be noted that

$$^{t}R^{Z}(n)=R^{Z}(-n) \qquad (|n|\leqq N), \tag{2.3}$$

$$T_{1}^{+}(\boldsymbol{Z})=T_{1}^{-}(\boldsymbol{Z})=R^{Z}(0). \tag{2.4}$$

In this subsection, we treat the case where the following condition holds:

$$T_{n}^{+}(\boldsymbol{Z}),\ T_{n}^{-}(\boldsymbol{Z})\in GL(nd\,;\mathbb{R}) \qquad (1\leqq n\leqq N). \tag{2.5}$$

We remark that condition (2.5) is equivalent to

$$\{Z_{j}(n)\,;1\leqq j\leqq d,\ |n|\leqq N\} \text{ is linearly independent in } L^{2}(\Omega, \mathcal{B}, P), \tag{2.6}$$

where $Z(n)={}^{t}(Z_{1}(n), \cdots, Z_{d}(n))$.

For any d-dimensional square-integrable stochastic process $\boldsymbol{Y}=(Y(n)\,;l\leqq n\leqq r)$ with a discrete time parameter space defined on the probability space $(\Omega, \mathcal{B}, P)$ $(l, r\in\mathbb{Z}, l<r)$, we define, for any $m, n\in\mathbb{Z}$, $l\leqq m\leqq n\leqq r$, a real closed subspace $\mathcal{L}_{m}^{n}(\boldsymbol{Y})$ of $L^{2}(\Omega, \mathcal{B}, P)$ by

$$\mathcal{L}_{m}^{n}(\boldsymbol{Y})=\text{the closed linear hull of } \{Y_{j}(k)\,;1\leqq j\leqq d,\ m\leqq k\leqq n\}. \tag{2.7}$$

According to the method of innovation, we introduce the d-dimensional forward (resp. backward) KM$_2$O-Langevin force $\boldsymbol{\nu}_{+}(\boldsymbol{Z})=(\nu_{+}(\boldsymbol{Z})(n)\,;0\leqq n\leqq N)$ (resp. $\boldsymbol{\nu}_{-}(\boldsymbol{Z})=(\nu_{-}(\boldsymbol{Z})(m)\,;-N\leqq m\leqq 0)$) as follows:

$$\nu_{+}(\boldsymbol{Z})(n)=Z(n)-P_{\mathcal{L}_{0}^{n-1}(\boldsymbol{Z})}Z(n) \qquad (0\leqq n\leqq N); \tag{2.8$_+$}$$

$$\nu_{-}(\boldsymbol{Z})(m)=Z(m)-P_{\mathcal{L}_{m+1}^{0}(\boldsymbol{Z})}Z(m) \qquad (-N\leqq m\leqq 0), \tag{2.8$_-$}$$

where $\mathcal{L}_{0}^{-1}(\boldsymbol{Z})=\mathcal{L}_{1}^{0}(\boldsymbol{Z})=\{0\}$.

For each $n\in\mathbb{N}^{*}$, $0\leqq n\leqq N$, let $V_{+}(\boldsymbol{Z})(n)$ (resp. $V_{-}(\boldsymbol{Z})(n)$) be the covariance matrix of $\nu_{+}(\boldsymbol{Z})(n)$ (resp. $\nu_{-}(\boldsymbol{Z})(-n)$). We call the function $V_{+}(\boldsymbol{Z})(\cdot)$ (resp. $V_{-}(\boldsymbol{Z})(\cdot)$) the forward (resp. backward) KM$_2$O-Langevin fluctuation function. The following causal relation holds among $\boldsymbol{Z}$, $\boldsymbol{\nu}_{+}(\boldsymbol{Z})$ and $\boldsymbol{\nu}_{-}(\boldsymbol{Z})$:

CAUSAL RELATION ([**5**], [**9**]).

(2.9) $$\nu_+(\boldsymbol{Z})(0) = \nu_-(\boldsymbol{Z})(0) = Z(0).$$

(2.10$_\pm$) $$E(\nu_\pm(\boldsymbol{Z})(\pm m)^t\nu_\pm(\boldsymbol{Z})(\pm n)) = \delta_{mn}V_\pm(\boldsymbol{Z})(n) \qquad (0 \leqq m,\ n \leqq N).$$

(2.11$_+$) $$\mathcal{L}_0^n(\boldsymbol{Z}) = \mathcal{L}_0^n(\boldsymbol{\nu}_+(\boldsymbol{Z})) \qquad (0 \leqq n \leqq N).$$

(2.11$_-$) $$\mathcal{L}_{-n}^0(\boldsymbol{Z}) = \mathcal{L}_{-n}^0(\boldsymbol{\nu}_-(\boldsymbol{Z})) \qquad (0 \leqq n \leqq N).$$

Let the system $\mathcal{LD}(\boldsymbol{Z})$ of elements in $M(d;\mathbb{R})$ be the *KM_2O-Langevin data* associated with the process $\boldsymbol{Z}$:

$$\mathcal{LD}(\boldsymbol{Z}) = \{\gamma_+(\boldsymbol{Z})(n,k),\ \gamma_-(\boldsymbol{Z})(n,k),\ \delta_+(\boldsymbol{Z})(m),\ \delta_-(\boldsymbol{Z})(m),\ V_+(\boldsymbol{Z})(l),\ V_-(\boldsymbol{Z})(l);$$
$$k,\ m,\ n \in \mathbb{N},\ 1 \leqq k < n \leqq N,\ 1 \leqq m \leqq N,\ l \in \mathbb{N}^*,\ 0 \leqq l \leqq N\}.$$

We know that $\boldsymbol{Z}$ satisfies the forward (resp. backward) KM_2O-Langevin equation (2.12$_+$) (resp. (2.12$_-$)):

KM_2O-LANGEVIN EQUATIONS ([**5**], [**9**]).

(2.12$_\pm$) $$Z(\pm n) = -\sum_{k=1}^{n-1}\gamma_\pm(\boldsymbol{Z})(n,k)Z(\pm k) - \delta_\pm(\boldsymbol{Z})(n)Z(0) + \nu_\pm(\boldsymbol{Z})(\pm n) \quad (1 \leqq n \leqq N).$$

In the sequal we adopt a convention to make the summation running the empty set 0. We call the function $\gamma_+(\boldsymbol{Z})(\cdot, *)$ (resp. $\gamma_-(\boldsymbol{Z})(\cdot, *)$) the forward (resp. backward) KM_2O-Langevin delay function associated with the process $\boldsymbol{Z}$. The function $\delta_+(\boldsymbol{Z})(\cdot)$ (resp. $\delta_-(\boldsymbol{Z})(\cdot)$) is said to be the forward (resp. backward) KM_2O-Langevin partial correlation function associated with the process $\boldsymbol{Z}$.

Concerning the relation between the Toeplitz matrices and the KM_2O-Langevin fluctuation functions, we can use the KM_2O-Langevin equations to see that

(2.13$_\pm$) $$\det T_n^\pm(\boldsymbol{Z}) = \prod_{k=0}^{n-1}\det V_\pm(\boldsymbol{Z})(k) \qquad (1 \leqq n \leqq N).$$

If follows from (2.5) and (2.13$_\pm$) that

(2.14) $$V_+(\boldsymbol{Z})(n),\ V_-(\boldsymbol{Z})(n) \in GL(d;\mathbb{R}) \qquad (0 \leqq n \leqq N).$$

The fluctuation-dissipation theorem (FDT) stated in §1 is the following:

FDT ([**2**], [**1**], [**11**], [**15**], [**5**], [**9**]). *For any $n,\ k \in \mathbb{N}$, $1 \leqq k < n \leqq N$,*

(2.15$_\pm$) $$\gamma_\pm(\boldsymbol{Z})(n,k) = \gamma_\pm(\boldsymbol{Z})(n-1,k-1) + \delta_\pm(\boldsymbol{Z})(n)\gamma_\mp(\boldsymbol{Z})(n-1,n-k-1);$$

(2.16$_\pm$) $$V_\pm(\boldsymbol{Z})(n) = (I - \delta_\pm(\boldsymbol{Z})(n)\delta_\mp(\boldsymbol{Z})(n))V_\pm(\boldsymbol{Z})(n-1);$$

(2.17) $$\delta_-(\boldsymbol{Z})(n)V_+(\boldsymbol{Z})(n-1) = V_-(\boldsymbol{Z})(n-1)^t\delta_+(\boldsymbol{Z})(n);$$

(2.18) $$\delta_-(\boldsymbol{Z})(n)V_+(\boldsymbol{Z})(n) = V_-(\boldsymbol{Z})(n)^t\delta_+(\boldsymbol{Z})(n),$$

where we put

(2.19) $\gamma_+(\boldsymbol{Z})(m, 0)=\delta_+(\boldsymbol{Z})(m)$ *and* $\gamma_-(\boldsymbol{Z})(m, 0)=\delta_-(\boldsymbol{Z})(m)$ $(1\leqq m\leqq N)$.

The relations ($2.16_\pm$) and (2.17) in FDT come from the following relation:

BURG'S RELATION ([**11**], [**15**], [**5**], [**9**]). *For any* $n\in\mathbb{N}$, $1\leqq n\leqq N$,

(2.20) $$\sum_{k=0}^{n-1}\gamma_+(\boldsymbol{Z})(n, k)R^Z(k+1)=\sum_{k=0}^{n-1}R^Z(k+1)^t\gamma_-(\boldsymbol{Z})(n, k).$$

FDT implies that both the KM_2O-Langevin delay and fluctuation functions can be recursively calculated from the KM_2O-Langevin partial correlation functions. On the other hand, the latter can be obtained from the correlation function R^Z by the following formulae:

KM_2O-LANGEVIN PARTIAL CORRELATION FUNCTIONS ([**2**], [**1**], [**11**], [**15**], [**5**], [**9**]). *For any* $n\in\mathbb{N}$, $1\leqq n\leqq N$,

($2.21_\pm$) $$\delta_\pm(\boldsymbol{Z})(n)=-(R^Z(\pm n)+\sum_{k=0}^{n-2}\gamma_\pm(\boldsymbol{Z})(n-1, k)R^Z(\pm(k+1)))V_\mp(\boldsymbol{Z})(n-1)^{-1}.$$

For any $m, n\in\mathbb{N}^*$, $0\leqq n\leqq m\leqq N$, we define $P_+(\boldsymbol{Z})(m, n)$, $P_-(\boldsymbol{Z})(m, n)$ and $e_+(\boldsymbol{Z})(m, n)$, $e_-(\boldsymbol{Z})(m, n)$ by

($2.22_\pm$) $$P_\pm(\boldsymbol{Z})(m, n)=E(Z(\pm m)^t\nu_\pm(\boldsymbol{Z})(\pm n))V_\pm(\boldsymbol{Z})(n)^{-1/2}$$

and

(2.23_+) $$e_+(\boldsymbol{Z})(m, n)=E((Z(m)-P_{\mathcal{L}_0^n(Z)}Z(m))^t(Z(m)-P_{\mathcal{L}_0^n(Z)}Z(m))),$$

(2.23_-) $$e_-(\boldsymbol{Z})(m, n)=E((Z(-m)-P_{\mathcal{L}_{-n}^0(Z)}Z(-m))^t(Z(-m)-P_{\mathcal{L}_{-n}^0(Z)}Z(-m))).$$

We call the function $P_+(\boldsymbol{Z})(\cdot, *)$ (resp. $P_-(\boldsymbol{Z})(\cdot, *)$) the forward (resp. backward) prediction function and the function $e_+(\boldsymbol{Z})(\cdot, *)$ (resp. $e_-(\boldsymbol{Z})(\cdot, *)$) the forward (resp. backward) prediction error function. Then we know

PREDICTION FORMULAE ([**5**], [**9**]). (i) *For any* $m, n\in\mathbb{N}^*$, $0\leqq n\leqq m\leqq N$,

(2.24_+) $$P_{\mathcal{L}_0^n(Z)}Z(m)=\sum_{k=0}^{n}P_+(\boldsymbol{Z})(m, k)V_+(\boldsymbol{Z})(k)^{-1/2}\nu_+(\boldsymbol{Z})(k);$$

(2.24_-) $$P_{\mathcal{L}_{-n}^0(Z)}Z(-m)=\sum_{k=0}^{n}P_-(\boldsymbol{Z})(m, k)V_-(\boldsymbol{Z})(k)^{-1/2}\nu_-(\boldsymbol{Z})(-k).$$

(ii) *For any* $m, n\in\mathbb{N}^*$, $0\leqq n<m\leqq N$,

(2.25_+) $$P_{\mathcal{L}_0^n(Z)}Z(m)=\sum_{k=0}^{n}Q_+(\boldsymbol{Z})(m, n; k)Z(k);$$

(2.25_-) $$P_{\mathcal{L}_{-n}^0(Z)}Z(-m)=\sum_{k=0}^{n}Q_-(\boldsymbol{Z})(m, n; k)Z(-k).$$

Here the $M(d; \mathbb{R})$-valued prediction functions $P_\pm(\boldsymbol{Z})(\cdot, *)$ and $Q_\pm(\boldsymbol{Z})(\cdot, *; \star)$

can be determined by the following algorithms:

PREDICTION ALGORITHMS ([**5**], [**9**]). (i) *For any* $m, k \in \mathbb{N}^*$, $0 \leqq k \leqq m \leqq N$,

$$(2.26_\pm) \qquad P_\pm(\boldsymbol{Z})(m, k) = \begin{cases} V_\pm(\boldsymbol{Z})(k)^{1/2} & \text{if } m=k \\ -\sum_{l=k}^{m-1} \gamma_\pm(\boldsymbol{Z})(m, l) P_\pm(\boldsymbol{Z})(l, k) & \text{if } m \geqq k+1. \end{cases}$$

(ii) *For any* $m, n, k \in \mathbb{N}^*$, $0 \leqq k \leqq n < m \leqq N$,

$$(2.27_\pm) \qquad Q_\pm(\boldsymbol{Z})(m, n; k) = -\sum_{l=n+1}^{m-1} \gamma_\pm(\boldsymbol{Z})(m, l) Q_\pm(\boldsymbol{Z})(l, n; k) - \gamma_\pm(\boldsymbol{Z})(m, k).$$

Finally the prediction error functions can be calculated by the following formulae:

PREDICTION ERROR FORMULAE ([**5**], [**9**]). (i) *For any* $m, n \in \mathbb{N}^*$, $0 \leqq n < m \leqq N$,

$$(2.28_\pm) \qquad e_\pm(\boldsymbol{Z})(m, n) = \sum_{k=n+1}^{m} P_\pm(\boldsymbol{Z})(m, k)\, {}^t P_\pm(\boldsymbol{Z})(m, k).$$

(ii) *In particular, for any* $n \in \mathbb{N}$, $1 \leqq n \leqq N$,

$$(2.29_\pm) \qquad e_\pm(\boldsymbol{Z})(n, n-1) = (I - \delta_\pm(\boldsymbol{Z})(n)\delta_\mp(\boldsymbol{Z})(n)) \cdots (I - \delta_\pm(\boldsymbol{Z})(1)\delta_\mp(\boldsymbol{Z})(1)) R^{\boldsymbol{Z}}(0).$$

[2.2] Let $\boldsymbol{Z} = (Z(n); n \in \mathbb{Z})$ be any d-dimensional real-valued weakly stationary time series on a probability space $(\Omega, \mathcal{B}, P)$ with covariance function $R^{\boldsymbol{Z}}$. In this subsection, we treat the case where the following condition holds:

$$(2.30) \qquad \{Z_j(n); 1 \leqq j \leqq d, n \in \mathbb{Z}\} \text{ is linearly independent in } L^2(\Omega, \mathcal{B}, P),$$

where $Z(n) = {}^t(Z_1(n), \cdots, Z_d(n))$.

By restricting the time parameter space, we have a d-dimensional real-valued local and weakly stationary time series $\boldsymbol{Z}_N = (Z(n); |n| \leqq N)$ $(N \in \mathbb{N})$. It then can be seen that the system $\{\mathcal{L}\mathcal{D}(\boldsymbol{Z}_N); N \in \mathbb{N}\}$ of the KM_2O-Langevin data $\mathcal{L}\mathcal{D}(\boldsymbol{Z}_N)$ $(N \in \mathbb{N})$ satisfies the following consistency condition:

$$\gamma_\pm(\boldsymbol{Z}_{N+1})(n, k) = \gamma_\pm(\boldsymbol{Z}_N)(n, k) \qquad (1 \leqq k < n \leqq N);$$
$$\delta_\pm(\boldsymbol{Z}_{N+1})(n) = \delta_\pm(\boldsymbol{Z}_N)(n) \qquad (1 \leqq n \leqq N);$$
$$V_\pm(\boldsymbol{Z}_{N+1})(n) = V_\pm(\boldsymbol{Z}_N)(n) \qquad (0 \leqq n \leqq N).$$

Therefore, we can construct a KM_2O-Langevin data $\mathcal{L}\mathcal{D}(\boldsymbol{Z})$ associated with the process $\boldsymbol{Z}$:

$$\mathcal{L}\mathcal{D}(\boldsymbol{Z}) = \{\gamma_\pm(\boldsymbol{Z})(n, k), \delta_\pm(\boldsymbol{Z})(m), V_\pm(\boldsymbol{Z})(l); k, m, n \in \mathbb{N}, k < n, l \in \mathbb{N}^*\}.$$

§ 3. A new formula for the KM$_2$O-Langevin data.

Let d, $d^{(1)}$, $d^{(2)}$, N be any natural numbers such that $d=d^{(1)}+d^{(2)}$ and let $\boldsymbol{Z}=(Z(n);\ |n|\leqq N)$ be any d-dimensional local and weakly stationary time series satisfying condition (2.6). We divide the components of $Z(n)$ into two blocks $Y(n)$ and $W(n)$, i.e.,

$$Z(n)=\begin{pmatrix} Y(n) \\ W(n) \end{pmatrix} \qquad (|n|\leqq N), \tag{3.1}$$

where $Y(n)={}^t(Z_1(n), \cdots, Z_{d^{(1)}}(n))$ and $W(n)={}^t(Z_{d^{(1)}+1}(n), \cdots, Z_{d^{(1)}+d^{(2)}}(n))$. It is to be noted that $\boldsymbol{Y}=(Y(n);\ |n|\leqq N)$ (resp. $\boldsymbol{W}=(W(n);\ |n|\leqq N)$ is a $d^{(1)}$-dimensional (resp. $d^{(2)}$-dimensional) weakly stationary time series satisfying condition (2.6).

In this section, we discuss how the KM$_2$O-Langevin data associated with $\boldsymbol{Z}$ is calculated by those associated with $\boldsymbol{Y}$ and $\boldsymbol{W}$. We define the mutual correlation function R^{YW} of $\boldsymbol{Y}$ and $\boldsymbol{W}$:

$$R^{YW}(n)=E(Y(n)\,{}^tW(0)) \qquad (|n|\leqq N). \tag{3.2}$$

Let $\mathcal{LD}(\boldsymbol{Z})$ (resp. $\mathcal{LD}(\boldsymbol{Y})$ and $\mathcal{LD}(\boldsymbol{W})$) be the KM$_2$O-Langevin data associated with $\boldsymbol{Z}$ (resp. $\boldsymbol{Y}$ and $\boldsymbol{W}$). We divide the components of matrices $\gamma_\pm(\boldsymbol{Z})(n, k)$ and $\delta_\pm(\boldsymbol{Z})(n)$ into four blocks $\gamma_\pm^{pq}(\boldsymbol{Z})(n, k)$, and $\delta_\pm^{pq}(\boldsymbol{Z})(n)$, for $p, q\in\mathbb{N}$, $1\leqq p$, $q\leqq 2$, i.e.,

$$\gamma_\pm(\boldsymbol{Z})(n, k)=\begin{pmatrix} \gamma_\pm^{11}(\boldsymbol{Z})(n, k) & \gamma_\pm^{12}(\boldsymbol{Z})(n, k) \\ \gamma_\pm^{21}(\boldsymbol{Z})(n, k) & \gamma_\pm^{22}(\boldsymbol{Z})(n, k) \end{pmatrix}$$

and

$$\delta_\pm(\boldsymbol{Z})(n)=\begin{pmatrix} \delta_\pm^{11}(\boldsymbol{Z})(n) & \delta_\pm^{12}(\boldsymbol{Z})(n) \\ \delta_\pm^{21}(\boldsymbol{Z})(n) & \delta_\pm^{22}(\boldsymbol{Z})(n) \end{pmatrix},$$

where $\gamma_\pm^{pq}(\boldsymbol{Z})(n, k)=((\gamma_\pm(\boldsymbol{Z})(n, k))_{ij})_{d^{(p-1)}+1\leqq i\leqq d^{(p-1)}+d^{(p)},\ d^{(q-1)}+1\leqq j\leqq d^{(q-1)}+d^{(q)}}$ with $d^{(0)}=0$ and $\delta_\pm^{pq}(\boldsymbol{Z})(n)=\gamma_\pm^{pq}(\boldsymbol{Z})(n, 0)$.

Furthermore, we divide the components of $\nu_\pm(\boldsymbol{Z})(n)$ into two blocks $\nu_\pm^1(\boldsymbol{Z})(n)$ and $\nu_\pm^2(\boldsymbol{Z})(n)$, i.e.,

$$\nu_\pm(\boldsymbol{Z})(n)=\begin{pmatrix} \nu_\pm^1(\boldsymbol{Z})(n) \\ \nu_\pm^2(\boldsymbol{Z})(n) \end{pmatrix},$$

where $\nu_\pm^1(\boldsymbol{Z})(n)={}^t(\nu_{\pm 1}(\boldsymbol{Z})(n), \cdots, \nu_{\pm d^{(1)}}(\boldsymbol{Z})(n))$ and $\nu_\pm^2(\boldsymbol{Z})(n)={}^t(\nu_{\pm(d^{(1)}+1)}(\boldsymbol{Z})(n), \cdots, \nu_{\pm(d^{(1)}+d^{(2)})}(\boldsymbol{Z})(n))$. Then, for any $n\in\mathbb{N}$, $1\leqq n\leqq N$, the KM$_2$O-Langevin equations (2.12$_\pm$) for $\boldsymbol{Z}$ are represented as follows:

(3.3$_\pm$) $$Z(\pm n) = -\sum_{k=1}^{n-1}\begin{pmatrix}\gamma_\pm^{11}(\boldsymbol{Z})(n,k) & \gamma_\pm^{12}(\boldsymbol{Z})(n,k)\\ \gamma_\pm^{21}(\boldsymbol{Z})(n,k) & \gamma_\pm^{22}(\boldsymbol{Z})(n,k)\end{pmatrix}\begin{pmatrix}Y(\pm k)\\ W(\pm k)\end{pmatrix} - \begin{pmatrix}\delta_\pm^{11}(\boldsymbol{Z})(n) & \delta_\pm^{12}(\boldsymbol{Z})(n)\\ \delta_\pm^{21}(\boldsymbol{Z})(n) & \delta_\pm^{22}(\boldsymbol{Z})(n)\end{pmatrix}\begin{pmatrix}Y(0)\\ W(0)\end{pmatrix} + \begin{pmatrix}\nu_\pm^{1}(\boldsymbol{Z})(\pm n)\\ \nu_\pm^{2}(\boldsymbol{Z})(\pm n)\end{pmatrix}.$$

By noting (3.1), we have

(3.4$_\pm$) $$Y(\pm n) = -\sum_{k=1}^{n-1}\gamma_\pm^{11}(\boldsymbol{Z})(n,k)Y(\pm k) - \sum_{k=1}^{n-1}\gamma_\pm^{12}(\boldsymbol{Z})(n,k)W(\pm k) - \delta_\pm^{11}(\boldsymbol{Z})(n)Y(0) - \delta_\pm^{12}(\boldsymbol{Z})(n)W(0) + \nu_\pm^{1}(\boldsymbol{Z})(\pm n);$$

(3.5$_\pm$) $$W(\pm n) = -\sum_{k=1}^{n-1}\gamma_\pm^{21}(\boldsymbol{Z})(n,k)Y(\pm k) - \sum_{k=1}^{n-1}\gamma_\pm^{22}(\boldsymbol{Z})(n,k)W(\pm k) - \delta_\pm^{21}(\boldsymbol{Z})(n)Y(0) - \delta_\pm^{22}(\boldsymbol{Z})(n)W(0) + \nu_\pm^{2}(\boldsymbol{Z})(\pm n).$$

We shall obtain other formulae, different from (2.21$_\pm$), by which the KM_2O-Langevin partial correlation functions $\delta_+(\boldsymbol{Z})(\cdot)$ and $\delta_-(\boldsymbol{Z})(\cdot)$ are recursively calculated from $\mathcal{LD}(\boldsymbol{Y})$, $\mathcal{LD}(\boldsymbol{W})$ and R^{YW} together with (2.15$_\pm$). For this purpose, we define $B_+(\boldsymbol{Y}|\boldsymbol{W})(l,k)$, $B_-(\boldsymbol{Y}|\boldsymbol{W})(l,k)$, $B_+(\boldsymbol{W}|\boldsymbol{Y})(l,k)$ and $B_-(\boldsymbol{W}|\boldsymbol{Y})(l,k)$ by

(3.6$_\pm$) $$B_\pm(\boldsymbol{Y}|\boldsymbol{W})(l,k) = R^{YW}(\pm l) + \sum_{j=0}^{k-2} R^{YW}(\pm(l-k+j+1))\,{}^t\gamma_\mp(\boldsymbol{W})(k-1,j)$$

and

(3.7$_\pm$) $$B_\pm(\boldsymbol{W}|\boldsymbol{Y})(l,k) = R^{WY}(\pm l) + \sum_{j=0}^{k-2} R^{WY}(\pm(l-k+j+1))\,{}^t\gamma_\mp(\boldsymbol{Y})(k-1,j)$$

for any $k, l \in \mathbb{N}^*$, $1 \leqq k \leqq N$, $0 \leqq l \leqq N$.

THEOREM 3.1. *For any $n \in \mathbb{N}$, $1 \leqq n \leqq N$,*

$$\delta_\pm(\boldsymbol{Z})(n) = \left\{\begin{pmatrix}\delta_\pm(\boldsymbol{Y})(n)V_\mp(\boldsymbol{Y})(n-1) & 0\\ 0 & \delta_\pm(\boldsymbol{W})(n)V_\mp(\boldsymbol{W})(n-1)\end{pmatrix} - \sum_{k=0}^{n-1}\gamma_\pm(\boldsymbol{Z})(n-1,k)\begin{pmatrix}0 & B_\pm(\boldsymbol{Y}|\boldsymbol{W})(k+1,n)\\ B_\pm(\boldsymbol{W}|\boldsymbol{Y})(k+1,n) & 0\end{pmatrix}\right\}V_\mp(\boldsymbol{Z})(n-1)^{-1},$$

where

(3.8) $$\gamma_+(\boldsymbol{Z})(j,j) = I \quad \textit{and} \quad \gamma_-(\boldsymbol{Z})(j,j) = I \quad (0 \leqq j \leqq N).$$

PROOF. We prove the plus part. We shall rewrite the first term F of the right-hand side of the plus part of (2.21$_\pm$) for any fixed $n \in \mathbb{N}$, $1 \leqq n \leqq N$:

$$F = -\Big(R^Z(\pm n) + \sum_{k=0}^{n-2}\gamma_\pm(\boldsymbol{Z})(n-1,k)R^Z(\pm(k+1))\Big).$$

We divide the components of matrix F into four blocks F^{pq} for $p, q \in \mathbb{N}$, $1 \leqq p$, $q \leqq 2$, i.e.,

$$F = \begin{pmatrix} F^{11} & F^{12} \\ F^{21} & F^{22} \end{pmatrix},$$

where $F^{pq} = ((F)_{ij})_{d(p-1)+1 \leq i \leq d(p-1)+d(p),\, d(q-1)+1 \leq j \leq d(q-1)+d(q)}$.

At first we rewrite the (1, 1)-block F^{11} of F as follows:

$$F^{11} = -\Big(R^{Y}(n) + \sum_{k=0}^{n-2} \gamma_{+}^{11}(\boldsymbol{Z})(n-1, k) R^{Y}(k+1) + \sum_{k=0}^{n-2} \gamma_{+}^{12}(\boldsymbol{Z})(n-1, k) R^{WY}(k+1)\Big).$$

We shall rewrite the second term of the equation above; by using equation (2.12$_-$), we see from (2.10$_-$) and (2.11$_-$) that

$$\begin{aligned}
&\sum_{k=0}^{n-2} \gamma_{+}^{11}(\boldsymbol{Z})(n-1, k) R^{Y}(k+1) \\
&= \sum_{k=0}^{n-2} \gamma_{+}^{11}(\boldsymbol{Z})(n-1, k) E(Y(k-n+2)^{t}Y(-n+1)) \\
&= \sum_{k=0}^{n-2} \gamma_{+}^{11}(\boldsymbol{Z})(n-1, k) E\Big(Y(k-n+2)^{t}(-\sum_{j=0}^{n-2} \gamma_{-}(\boldsymbol{Y})(n-1, j) Y(-j))\Big) \\
&\quad + \sum_{k=0}^{n-2} \gamma_{+}^{11}(\boldsymbol{Z})(n-1, k) E(Y(k-n+2)^{t}\nu_{-}(\boldsymbol{Y})(-(n-1))) \\
&= -\sum_{k=0}^{n-2} \sum_{j=0}^{n-2} \gamma_{+}^{11}(\boldsymbol{Z})(n-1, k) R^{Y}(k-n+j+2)^{t}\gamma_{-}(\boldsymbol{Y})(n-1, j) \\
&= -\sum_{k=0}^{n-2} \sum_{j=0}^{n-2} \gamma_{+}^{11}(\boldsymbol{Z})(n-1, k) E(Y(k)^{t}Y(n-j-2))^{t}\gamma_{-}(\boldsymbol{Y})(n-1, j) \\
&= \sum_{j=0}^{n-2} E\Big(\Big(-\sum_{k=0}^{n-2} \gamma_{+}^{11}(\boldsymbol{Z})(n-1, k) Y(k)\Big)^{t}Y(n-j-2)\Big)^{t}\gamma_{-}(\boldsymbol{Y})(n-1, j).
\end{aligned}$$

On the other hand, by using equation (3.4$_+$), we see from (2.10$_+$) and (2.11$_+$) that

$$\begin{aligned}
&E\Big(\Big(-\sum_{k=0}^{n-2} \gamma_{+}^{11}(\boldsymbol{Z})(n-1, k) Y(k)\Big)^{t}Y(n-j-2)\Big) \\
&= E(Y(n-1)^{t}Y(n-j-2)) + E\Big(\Big(\sum_{k=0}^{n-2} \gamma_{+}^{12}(\boldsymbol{Z})(n-1, k) W(k)\Big)^{t}Y(n-j-2)\Big) \\
&\quad - E(\nu_{+}^{1}(\boldsymbol{Z})(n-1)^{t}Y(n-j-2)) \\
&= R^{Y}(j+1) + \sum_{k=0}^{n-2} \gamma_{+}^{12}(\boldsymbol{Z})(n-1, k) R^{WY}(k-n+j+2).
\end{aligned}$$

Further, by virtue of Burg's relation (2.20), we see

$$\sum_{k=0}^{n-2}\gamma_+^{11}(\boldsymbol{Z})(n-1,\ k)R^{\boldsymbol{Y}}(k+1)$$

$$=\sum_{k=0}^{n-2}\gamma_+(\boldsymbol{Y})(n-1,\ k)R^{\boldsymbol{Y}}(k+1)$$

$$+\sum_{j=0}^{n-2}\sum_{k=0}^{n-2}\gamma_+^{12}(\boldsymbol{Z})(n-1,\ k)R^{\boldsymbol{WY}}(k-n+j+2)^t\gamma_-(\boldsymbol{Y})(n-1,\ j).$$

According to the definition of $B_+(\boldsymbol{W}|\boldsymbol{Y})(\cdot,\ *)$, we see from (2.20_+) that

$$F^{11}=-\Big(R^{\boldsymbol{Y}}(n)+\sum_{k=0}^{n-2}\gamma_+(\boldsymbol{Y})(n-1,\ j)R^{\boldsymbol{Y}}(k+1)\Big)$$

$$-\sum_{k=0}^{n-2}\gamma_+^{12}(\boldsymbol{Z})(n-1,\ k)\Big(R^{\boldsymbol{WY}}(k+1)+\sum_{j=0}^{n-2}R^{\boldsymbol{WY}}(k-n+j+2)^t\gamma_-(\boldsymbol{Y})(n-1,\ j)\Big)$$

$$=\delta_+(\boldsymbol{Y})(n)V_-(\boldsymbol{Y})(n-1)-\sum_{k=0}^{n-2}\gamma_+^{12}(\boldsymbol{Z})(n-1,\ k)B_+(\boldsymbol{W}|\ \boldsymbol{Y})(k+1,\ n).$$

Therefore, according to (3.8), we get

(a) $$F^{11}=\delta_+(\boldsymbol{Y})(n)V_-(\boldsymbol{Y})(n-1)-\sum_{k=0}^{n-1}\gamma_+^{12}(\boldsymbol{Z})(n-1,\ k)B_+(\boldsymbol{W}|\ \boldsymbol{Y})(k+1,\ n).$$

Secondly, we rewrite the (2, 1)-block F^{21} of F as follows:

$$F^{21}=-\Big(R^{\boldsymbol{WY}}(n)+\sum_{k=0}^{n-2}\gamma_+^{21}(\boldsymbol{Z})(n-1,\ k)R^{\boldsymbol{Y}}(k+1)+\sum_{k=0}^{n-2}\gamma_+^{22}(\boldsymbol{Z})(n-1,\ k)R^{\boldsymbol{WY}}(k+1)\Big).$$

We shall rewrite the second term of the equation above; by using equation (2.12_-), we see from (2.10_-) and (2.11_-) that

$$\sum_{k=0}^{n-2}\gamma_+^{21}(\boldsymbol{Z})(n-1,\ k)R^{\boldsymbol{Y}}(k+1)$$

$$=\sum_{j=0}^{n-2}E\Big(\Big(-\sum_{k=0}^{n-2}\gamma_+^{21}(\boldsymbol{Z})(n-1,\ k)Y(k)\Big)^tY(n-j-2)\Big)^t\gamma_-(\boldsymbol{Y})(n-1,\ j).$$

On the other hand, by using equation (3.5_+), we have from (2.10_+) and (2.11_+) that

$$E\Big(\Big(-\sum_{k=0}^{n-2}\gamma_+^{21}(\boldsymbol{Z})(n-1,\ k)Y(k)\Big)^tY(n-j-2)\Big)$$

$$=R^{\boldsymbol{WY}}(j+1)+\sum_{k=0}^{n-2}\gamma_+^{22}(\boldsymbol{Z})(n-1,\ k)R^{\boldsymbol{WY}}(k-n+j+2).$$

Therefore, we obtain

$$F^{21}=-\Big(R^{\boldsymbol{WY}}(n)+\sum_{k=0}^{n-2}R^{\boldsymbol{WY}}(k+1)^t\gamma_-(\boldsymbol{Y})(n-1,\ j)\Big)$$

$$-\sum_{k=0}^{n-2}\gamma_+^{22}(\boldsymbol{Z})(n-1,\ k)\Big(R^{\boldsymbol{WY}}(k+1)+\sum_{j=0}^{n-2}R^{\boldsymbol{WY}}(k-n+j+2)^t\gamma_-(\boldsymbol{Y})(n-1,\ j)\Big).$$

According to the definition of $B_+(\boldsymbol{W}|\boldsymbol{Y})(\cdot, *)$ in (3.7$_+$) and (3.8), we get

(b) $$F^{21} = -\sum_{k=0}^{n-1} \gamma_+^{22}(\boldsymbol{Z})(n-1, k) B_+(\boldsymbol{W}|\boldsymbol{Y})(k+1, n).$$

Similarly, we can show

(c) $$F^{12} = -\sum_{k=0}^{n-1} \gamma_+^{11}(\boldsymbol{Z})(n-1, k) B_+(\boldsymbol{Y}|\boldsymbol{W})(k+1, n)$$

and

(d) $$F^{22} = \delta_+(\boldsymbol{W})(n) V_-(\boldsymbol{W})(n-1) - \sum_{k=0}^{n-1} \gamma_+^{21}(\boldsymbol{Z})(n-1, k) B_+(\boldsymbol{Y}|\boldsymbol{W})(k+1, n).$$

Thus we can conclude from (a), (b), (c) and (d) that the plus part holds. In the same way, the minus part is proved. (Q. E. D.)

As stated in §2, $V_+(\boldsymbol{Z})(\cdot)$ and $V_-(\boldsymbol{Z})(\cdot)$ are recursively calculated from $\delta_+(\boldsymbol{Z})(\cdot)$ and $\delta_-(\boldsymbol{Z})(\cdot)$ by (2.16$_\pm$). However, we can obtain other formulae for the KM$_2$O-Langevin fluctuation functions $V_\pm(\boldsymbol{Z})(\cdot)$, similar to Theorem 3.1.

THEOREM 3.2. *For any $n \in \mathbf{N}$, $0 \leqq n \leqq N$,*

$$V_\pm(\boldsymbol{Z})(n) = \begin{pmatrix} V_\pm(\boldsymbol{Y})(n) & 0 \\ 0 & V_\pm(\boldsymbol{W})(n) \end{pmatrix}$$
$$+ \sum_{k=0}^{n} \gamma_\pm(\boldsymbol{Z})(n, n-k) \begin{pmatrix} 0 & B_\mp(\boldsymbol{Y}|\boldsymbol{W})(k, n+1) \\ B_\mp(\boldsymbol{W}|\boldsymbol{Y})(k, n+1) & 0 \end{pmatrix}.$$

PROOF. We divide the components of matrices $V_\pm(\boldsymbol{Z})(n)$ into four blocks $V_\pm^{pq}(\boldsymbol{Z})(n)$ for $p, q \in \mathbf{N}$, $1 \leqq p, q \leqq 2$, i.e.,

$$V_\pm(\boldsymbol{Z})(n) = \begin{pmatrix} V_\pm^{11}(\boldsymbol{Z})(n) & V_\pm^{12}(\boldsymbol{Z})(n) \\ V_\pm^{21}(\boldsymbol{Z})(n) & V_\pm^{22}(\boldsymbol{Z})(n) \end{pmatrix},$$

where $V_\pm^{pq}(\boldsymbol{Z})(n) = ((V_\pm(\boldsymbol{Z})(n))_{ij})_{d(p-1)+1 \leq i \leq d(p-1)+d(p),\, d(q-1)+1 \leq j \leq d(q-1)+d(q)}$.

We prove only the plus part, because the minus part is proved in the same way. By using equation (3.4$_+$) for $\boldsymbol{Z}$, it follows from (2.10$_+$) and (2.11$_+$) that

$$\begin{aligned} V_+^{11}(\boldsymbol{Z})(n) &= E(\nu_+^1(\boldsymbol{Z})(n)\,{}^tY(n)) + E\Big(\nu_+^1(\boldsymbol{Z})(n)\,{}^t\Big(\sum_{k=0}^{n-1} \gamma_+^{11}(\boldsymbol{Z})(n, k) Y(k)\Big)\Big) \\ &\quad + E\Big(\nu_+^1(\boldsymbol{Z})(n)\,{}^t\Big(\sum_{k=0}^{n-1} \gamma_+^{12}(\boldsymbol{Z})(n, k) W(k)\Big)\Big) \\ &= E(\nu_+^1(\boldsymbol{Z})(n)\,{}^tY(n)). \end{aligned}$$

Further, by using equation (2.12$_+$) for $\boldsymbol{Y}$ and noting (2.10$_+$) and (2.11$_+$) that

$$V_+^{11}(\boldsymbol{Z})(n) = E\Big(\nu_+^1(\boldsymbol{Z})(n)^t\Big(-\sum_{k=0}^{n-1}\gamma_+(\boldsymbol{Y})(n,\ k)Y(k)\Big)\Big)+E(\nu_+^1(\boldsymbol{Z})(n)^t\nu_+(\boldsymbol{Y})(n))$$
$$= E(\nu_+^1(\boldsymbol{Z})(n)^t\nu_+(\boldsymbol{Y})(n)).$$

By using equation (3.4_+) for $\boldsymbol{Z}$, we see that

$$V_+^{11}(\boldsymbol{Z})(n) = E(Y(n)^t\nu_+(\boldsymbol{Y})(n))+E\Big(\Big(\sum_{k=0}^{n-1}\gamma_+^{11}(\boldsymbol{Z})(n,\ k)Y(k)\Big)^t\nu_+(\boldsymbol{Y})(n)\Big)$$
$$+E\Big(\Big(\sum_{k=0}^{n-1}\gamma_+^{12}(\boldsymbol{Z})(n,\ k)W(k)\Big)^t\nu_+(\boldsymbol{Y})(n)\Big)$$
$$= V_+(\boldsymbol{Y})(n)+\sum_{k=0}^{n-1}\gamma_+^{12}(\boldsymbol{Z})(n,\ k)E(W(k)^t\nu_+(\boldsymbol{Y})(n)).$$

On the other hand, by using equation (2.12_+) for $\boldsymbol{Y}$,

$$V_+^{11}(\boldsymbol{Z})(n) = V_+(\boldsymbol{Y})(n)+\sum_{l=1}^{n}\gamma_+^{12}(\boldsymbol{Z})(n,\ n-l)E(W(n-l)^t\nu_+(\boldsymbol{Y})(n))$$
$$= V_+(\boldsymbol{Y})(n)+\sum_{l=1}^{n}\gamma_+^{12}(\boldsymbol{Z})(n,\ n-l)E(W(n-l)^tY(n))$$
$$+\sum_{l=1}^{n}\gamma_+^{12}(\boldsymbol{Z})(n,\ n-l)E\Big(W(n-l)^t\Big(\sum_{j=0}^{n-1}\gamma_+(\boldsymbol{Y})(n,\ j)Y(j)\Big)\Big)$$
$$+\sum_{l=1}^{n}\gamma_+^{12}(\boldsymbol{Z})(n,\ n-l)E\Big(W(n-l)^t\Big(\sum_{j=0}^{n-1}\gamma_+(\boldsymbol{Y})(n,\ j)Y(j)\Big)\Big)$$
$$= V_+(\boldsymbol{Y})(n)+\sum_{l=1}^{n}\gamma_+^{12}(\boldsymbol{Z})(n,\ n-l)R^{WY}(-l)$$
$$+\sum_{l=1}^{n}\gamma_+^{12}(\boldsymbol{Z})(n,\ n-l)\sum_{j=0}^{n-1}R^{WY}(-(l-n+j))^t\gamma_+(\boldsymbol{Y})(n,\ j).$$

Therefore, according to the definition of $B_-(\boldsymbol{W}|\,\boldsymbol{Y})(\cdot,\ *)$ in (3.7_-) and (3.8),

$$\text{(e)}\qquad V_+^{11}(\boldsymbol{Z})(n) = V_+(\boldsymbol{Y})(n)+\sum_{k=0}^{n}\gamma_+^{12}(\boldsymbol{Z})(n,\ n-l)B_-(\boldsymbol{W}|\,\boldsymbol{Y})(k,\ n+1).$$

In the same way as in $V_+^{11}(\boldsymbol{Z})(n)$, it follows from ($3.4_+$), ($3.5_+$), ($2.10_+$), ($2.11_+$) and ($2.12_+$) that

$$V_+^{21}(\boldsymbol{Z})(n) = E(\nu_+^2(\boldsymbol{Z})(n)^tY(n))$$
$$= E(\nu_+^2(\boldsymbol{Z})(n)^t\nu_+(\boldsymbol{Y})(n))$$
$$= E(W(n)^t\nu_+(\boldsymbol{Y})(n))+\sum_{k=0}^{n-1}\gamma_+^{22}(\boldsymbol{Z})(n,\ k)E(W(k)^t\nu_+(\boldsymbol{Y})(n))$$
$$= R^{WY}(0)+\sum_{l=0}^{n-1}R^{WY}(n-l)^t\gamma_+(\boldsymbol{Y})(n,\ l)+\sum_{l=1}^{n}\gamma_+^{22}(\boldsymbol{Z})(n,\ n-l)R^{WY}(-l)$$
$$+\sum_{l=1}^{n}\gamma_+^{22}(\boldsymbol{Z})(n,\ n-l)\sum_{j=0}^{n-1}R^{WY}(-(l-n+j))^t\gamma_+(\boldsymbol{Y})(n,\ j).$$

Therefore, according to the definition of $B_-(\boldsymbol{W}|\boldsymbol{Y})(\cdot,*)$ in (3.7_-) and (3.8),

(f) $$V_+^{21}(\boldsymbol{Z})(n)=\sum_{k=0}^{n}\gamma_+^{22}(\boldsymbol{Z})(n,\,n-k)B_-(\boldsymbol{W}|\boldsymbol{Y})(k,\,n+1).$$

Similarly, we obtain

(g) $$V_+^{12}(\boldsymbol{Z})(n)=\sum_{k=0}^{n}\gamma_+^{11}(\boldsymbol{Z})(n,\,n-k)B_-(\boldsymbol{Y}|\boldsymbol{W})(k,\,n+1)$$

and

(h) $$V_+^{22}(\boldsymbol{Z})(n)=V_+(\boldsymbol{Z})(n)+\sum_{k=0}^{n}\gamma_+^{21}(\boldsymbol{Z})(n,\,n-k)B_-(\boldsymbol{Y}|\boldsymbol{W})(k,\,n+1).$$

Thus we can conclude from (e), (f), (g) and (h) that the plus part holds. (Q. E. D.)

§ 4. The non-linear prediction problem.

Let $\boldsymbol{X}=(X(n);n\in\mathbb{Z})$ be a one-dimensional strictly stationary time series on a probability space $(\Omega,\mathcal{B},P)$ with mean zero. Moreover we impose the same hypotheses as in Masani-Wiener [4]:

(H.1) *$\boldsymbol{X}$ is essentially bounded;*

(H.2) *for any distinct integers $(n_1,\cdots,n_k)$ the spectrum of the distribution function of the k-dimensional random variable ${}^t(X(n_1),\cdots,X(n_k))$ has positive Lebesgue measure.*

For any subset $\mathcal{A}$ of $L^2(\Omega,\mathcal{B},P)$, we denote by $[\mathcal{A}]$ the closed subspace of $L^2(\Omega,\mathcal{B},P)$, generated by all elements of $\mathcal{A}$.

To obtain the non-linear predictor $\hat{X}(\nu)=E(X(\nu)|\sigma(X(l);l\leqq 0))$ is reduced to getting a projection of $X(\nu)$ $(\nu\in\mathbb{N})$ as follows:

LEMMA 4.1 (Masami-Wiener [4]).

(i) $$E(X(\nu)|\sigma(X(l);l\leqq 0))=P_{\mathcal{M}_{-\infty}^0}X(\nu)\qquad(\nu\in\mathbb{N}),$$

where

$$\mathcal{M}_{-\infty}^0=\left[1,\,\prod_{k=0}^{m}X(n_k)^{p_k};\,m\in\mathbb{N}^*,\,p_k\in\mathbb{N},\,n_k\in\mathbb{Z}\ (0\leqq k\leqq m),\,n_0<\cdots<n_m\leqq 0\right].$$

(ii) $$\left\{1,\,\prod_{k=0}^{m}X(n_k)^{p_k};\,m\in\mathbb{N}^*,\,p_k\in\mathbb{N},\,n_k\in\mathbb{Z}\ (0\leqq k\leqq m),\,n_0<\cdots<n_m\leqq 0\right\}$$

is linearly independent in $L^2(\Omega,\mathcal{B},P)$.

We shall obtain certain computable algorithm for $\hat{X}(\nu)$. For that purpose, we shall show the following lemma.

LEMMA 4.2.

$$E(X(\nu)\,|\,\sigma(X(l)\,;\,l\leqq 0)) = P_{\mathcal{K}_{-\infty}^{0}}X(\nu) \qquad (\nu\in\mathbb{N}),$$

where

$$\mathcal{K}_{-\infty}^{0} = \Big[\prod_{k=0}^{m} X(n-k)^{p_k} - E\Big(\prod_{k=0}^{m} X(n-k)^{p_k}\Big);\ m\in\mathbb{N}^*,\ n\leqq 0,$$

$$p_0\in\mathbb{N},\ p_k\in\mathbb{N}^*\ (1\leqq k\leqq m)\Big].$$

PROOF. By Lemma 4.1(i), what we need to prove is that $P_{\mathcal{M}_{-\infty}^{0}}X(\nu)= P_{\mathcal{K}_{-\infty}^{0}}X(\nu)$ for any $\nu\in\mathbb{N}$. For any $m\in\mathbb{N}^*$, $n\leqq 0$, $p_0\in\mathbb{N}$, $p_k\in\mathbb{N}^*$ $(1\leqq k\leqq m)$, there exist $M\in\mathbb{N}^*$, $q_l\in\mathbb{N}$, $n_l\in\mathbb{Z}$ $(0\leqq l\leqq M)$, $n_0<\cdots<n_M\leqq 0$ such that

$$\prod_{k=0}^{m} X(n-k)^{p_k} = \prod_{l=0}^{M} X(n_l)^{q_l},$$

it can be seen that

$$\mathcal{M}_{-\infty}^{0}\ominus\mathcal{K}_{-\infty}^{0} = [1].$$

Therefore, we see that $P_{\mathcal{M}_{-\infty}^{0}\ominus\mathcal{K}_{-\infty}^{0}}X(\nu)=P_{[1]}X(\nu)=E(X(\nu))=0$. Thus, it follows that Lemma 4.2 holds. (Q. E. D.)

For the purpose of parametrizing the infinite-dimensional subspace $\mathcal{K}_{-\infty}^{0}$, we define a subset Λ of $\{0, 1, 2, \cdots\}^{\mathbb{N}^*}$ by

$$\Lambda = \{\boldsymbol{p}=(p_0, p_1, p_2, \cdots)\in\{0, 1, 2, \cdots\}^{\mathbb{N}^*};\ p_0\geqq 1 \text{ and there exists } m\in\mathbb{N}^* \text{ such that } p_m\neq 0,\ p_k=0\ (k\geqq m+1)\}.$$

For any $\boldsymbol{p}\in\Lambda$, a one-dimensional strictly stationary time series $\boldsymbol{\varphi}_{\boldsymbol{p}}=(\varphi_{\boldsymbol{p}}(n)\,;\, n\in\mathbb{Z})$ is introduced by

$$\varphi_{\boldsymbol{p}}(n) = \prod_{k=0}^{\infty} X(n-k)^{p_k}$$

and we set

$$G = \{\boldsymbol{\varphi}_{\boldsymbol{p}}\,;\ \boldsymbol{p}\in\Lambda\}.$$

We shall order the elements of G to arrange them in a sequence $\{\boldsymbol{\varphi}_j\,;\, j\in\mathbb{N}^*\}$. For each $q\in\mathbb{N}$, we define a subset Λ_q of Λ and a subset $G^{(q)}$ of G by

$$\Lambda_q = \{\boldsymbol{p}=(p_0, p_1, \cdots)\in\Lambda\,;\ q=\sum_{k=0}^{\infty}(k+1)\cdot p_k\} \quad\text{and}\quad G^{(q)} = \{\boldsymbol{\varphi}_{\boldsymbol{p}}\,;\ \boldsymbol{p}\in\Lambda_q\}.$$

Then we have the disjoint union

$$G = \bigcup_{q\in\mathbb{N}} G^{(q)}.$$

Now we shall order the elements of G. For any $\boldsymbol{\varphi}_{\boldsymbol{p}}\in G^{(q)}$ and $\boldsymbol{\varphi}_{\boldsymbol{p}'}\in G^{(q')}$, we say that $\boldsymbol{\varphi}_{\boldsymbol{p}}$ precedes $\boldsymbol{\varphi}_{\boldsymbol{p}'}$ if and only if $q<q'$ or $q=q'$ and in addition, there

exists $k_0 \in \mathbb{N}^*$ such that $p_k = p'_k$ $(0 \leqq k \leqq k_0 - 1)$ and $p_{k_0} > p'_{k_0}$. Then we have

$$G = \{\boldsymbol{\varphi}_j\,;\, j \in \mathbb{N}^*\}$$

and

$$G^{(q)} = \{\boldsymbol{\varphi}_{d_{q-1}+1}, \boldsymbol{\varphi}_{d_{q-1}+2}, \cdots, \boldsymbol{\varphi}_{d_q}\},$$

where

$$d_q = \text{the number of } \left\{\bigcup_{r=1}^{q} G^{(r)}\right\} - 1$$

and

$$\begin{aligned}&(\varphi_{d_{q-1}+1}(n), \varphi_{d_{q-1}+2}(n), \cdots, \varphi_{d_q}(n))\\ &\quad = (X(n)^q, X(n)^{q-2}X(n-1), \cdots, X(n)X(n-q+2)).\end{aligned}$$

For example,

$$(d_1, d_2, d_3, d_4) = (0, 1, 3, 6)$$

and

$$\begin{aligned}&(\varphi_0(n), \varphi_1(n), \varphi_2(n), \varphi_3(n), \varphi_4(n), \varphi_5(n), \varphi_6(n))\\ &\quad = (X(n), X(n)^2, X(n)^3, X(n)X(n-1), X(n)^4, X(n)^2X(n-1), X(n)X(n-2)).\end{aligned}$$

By using the system $G = \{\boldsymbol{\varphi}_j\,;\, j \in \mathbb{N}^*\}$, we define $\boldsymbol{X}^{(q)} = (X^{(q)}(n)\,;\, n \in \mathbb{Z})$ and $\boldsymbol{Y}^{(q)} = (Y^{(q)}(n)\,;\, n \in \mathbb{Z})$ by

$$X^{(q)}(n) = \begin{pmatrix} \varphi_0(n) - E(\varphi_0(n)) \\ \varphi_1(n) - E(\varphi_1(n)) \\ \vdots \\ \varphi_{d_q}(n) - E(\varphi_{d_q}(n)) \end{pmatrix}$$

and

$$Y^{(q)}(n) = \begin{pmatrix} \varphi_{d_{q-1}+1}(n) - E(\varphi_{d_{q-1}+1}(n)) \\ \varphi_{d_{q-1}+2}(n) - E(\varphi_{d_{q-1}+2}(n)) \\ \vdots \\ \varphi_{d_q}(n) - E(\varphi_{d_q}(n)) \end{pmatrix}.$$

Then, by virtue of Lemma 4.1(ii), we have the following lemma.

LEMMA 4.3.

(i) *For any* $q \in \mathbb{N}$, $\boldsymbol{X}^{(q)}$ *is a* d_q+1-*dimensional weakly stationary time series satisfying condition* (2.30).

(ii) $\boldsymbol{X}^{(1)} = \boldsymbol{X}$.

(iii) $X^{(q)}(n) = \begin{pmatrix} X^{(q-1)}(n) \\ Y^{(q)}(n) \end{pmatrix}$ $(q = 2, 3, \cdots)$.

(iv) $\left[\bigcup_{N=0}^{\infty} \bigcup_{q=1}^{\infty} \mathcal{L}_{-N}^{0}(\boldsymbol{X}^{(q)})\right] = \mathcal{K}_{-\infty}^{0}$.

We shall show how the non-linear predictor of $\boldsymbol{X}$ is expressed by using the linear predictor of $X^{(q)}$.

THEOREM 4.1. *For any* $\nu>0$,

$$E(X(\nu)\mid\sigma(X(l)\,;\,l\leqq 0))$$
$$=\textit{the first component of } \underset{N,q\to\infty}{\text{l.i.m.}}\Big(\sum_{k=0}^{N}Q_+(X^{(q)})(N+\nu,N;N-k)X^{(q)}(-k)\Big).$$

PROOF. By Lemmas 4.2 and 4.3(iv), we have

$$E(X(\nu)\mid\sigma(X(l)\,;\,l\leqq 0))=\underset{N,q\to\infty}{\text{l.i.m.}}\,P_{\mathcal{L}_{-N}^{0}(X^{(q)})}X(\nu)$$
$$=\text{the first component of }\underset{N,q\to\infty}{\text{l.i.m.}}P_{\mathcal{L}_{-N}^{0}(X^{(q)})}X^{(q)}(\nu).$$

By applying the prediction formula (2.25_+) to the time series $X^{(q)}$, we have

$$\begin{aligned}P_{\mathcal{L}_{-N}^{0}(X^{(q)})}X^{(q)}(\nu)&=U(-N)P_{\mathcal{L}_{0}^{N}(X^{(q)})}X^{(q)}(N+\nu)\\&=U(-N)\Big(\sum_{k=0}^{N}Q_+(X^{(q)})(N+\nu,N;k)X^{(q)}(k)\Big)\\&=\sum_{k=0}^{N}Q_+(X^{(q)})(N+\nu,N;k)X^{(q)}(k-N)\\&=\sum_{k=0}^{N}Q_+(X^{(q)})(N+\nu,N;N-k)X^{(q)}(-k),\end{aligned}$$

where $U(-N)$ is a unitary operator from $\mathcal{L}_0^N(X^{(q)})$ to $\mathcal{L}_{-N}^0(X^{(q)})$ such that $U(-N)X^{(q)}(n)=X^{(q)}(n-N)$ $(0\leqq n\leqq N)$. Therefore, we get Theorem 4.1. (Q.E.D.)

We shall explain the structure of algorithm computing the coefficients $Q_+(X^{(q)})(\cdot,*;\star)$ $(q\in\mathbf{N})$ in Theorem 4.1. Let $\mathcal{LD}(X^{(q)})$ (resp. $\mathcal{LD}(X^{(q-1)})$ and $\mathcal{LD}(Y^{(q)})$) be the KM$_2$O-Langevin data associated with $X^{(q)}$ (resp. $X^{(q-1)}$ and $Y^{(q)}$). By (2.27_+),

$$Q_\pm(X^{(q)})(m,n\,;\,k)=-\sum_{l=n+1}^{m-1}\gamma_\pm(X^{(q)})(m,l)Q_\pm(X^{(q)})(l,n\,;\,k)-\gamma_\pm(X^{(q)})(m,k),\tag{4.1}$$

which implies that, for each fixed $q\in\mathbf{N}$, $Q_\pm(X^{(q)})(\cdot,*;\star)$ can be calculated from $\mathcal{LD}(X^{(q)})$. By virtue of FDT, $\mathcal{LD}(X^{(q)})$ can be recursively calculated from the KM$_2$O-Langevin partial correlation functions $\delta_\pm(X^{(q)})(\cdot)$. By applying Theorem 3.1 to the time series $X^{(q)}$, we obtain an algorithm computing $\delta_\pm(X^{(q)})(\cdot)$ in Theorem 4.2. The crux is that the $\delta_\pm(X^{(q)})(\cdot)$ can be calculated from $\mathcal{LD}(X^{(q-1)})$, $\mathcal{LD}(Y^{(q)})$ and $R^{X^{(q-1)}Y^{(q)}}$ $(q=2,3,\cdots)$.

THEOREM 4.2. *For any* $n,q\in\mathbf{N}$, $2\leqq q$,

$$\delta_\pm(X^{(q)})(n)=\left\{\begin{pmatrix}\delta_\pm(X^{(q-1)})(n)V_\mp(X^{(q-1)})(n-1) & 0\\ 0 & \delta_\pm(Y^{(q)})(n)V_\mp(Y^{(q)})(n-1)\end{pmatrix}\right.$$

$$-\sum_{k=0}^{n-1}\gamma_\pm(X^{(q)})(n-1,\ k)\cdot$$

$$\left.\cdot\begin{pmatrix}0 & B_\pm(X^{(q-1)}\,|\,Y^{(q)})(k+1,\ n)\\ B_\pm(Y^{(q)}\,|\,X^{(q-1)})(k+1,\ n) & 0\end{pmatrix}\right\}V_\mp(X^{(q)})(n-1)^{-1},$$

where

$$\gamma_+(X^{(q)})(j,\ j)=I \quad and \quad \gamma_-(X^{(q)})(j,\ j)=I \quad (j\in\mathbb{N}^*).$$

Finally we shall make a comment concerning the global behavior of the prediction functions $Q_\pm(X^{(q)})(N+\nu,\ N;\ N-k)$ as $N\to\infty$ in order to complete the representation for the non-linear predictor in Theorem 4.1. For that purpose, we need the following stronger condition (H.3) than (H.2), besides (H.1):

(H.3) *For each $q\in\mathbb{N}$, the weakly stationary process $X^{(q)}$ has the spectral density matrix function $\Delta(X^{(q)})(\theta)$ defined on $[-\pi,\ \pi)$ such that*

$$\log(\det(\Delta(X^{(q)})))\in L^1(-\pi,\ \pi). \tag{4.2}$$

By Theorems 4.2, 5.1 and 5.2 in [7], we find that, for each $q\in\mathbb{N}$, the following limits exist:

$$V_\pm(X^{(q)})\equiv\lim_{n\to\infty}V_\pm(X^{(q)})(n); \tag{4.3$_\pm$}$$

$$\gamma_\pm(X^{(q)})(k)\equiv\lim_{n\to\infty}\gamma_\pm(X^{(q)})(n,\ n-k) \qquad (k\in\mathbb{N}^*); \tag{4.4$_\pm$}$$

$$P_\pm(X^{(q)})(k)\equiv\lim_{n\to\infty}P_\pm(X^{(q)})(n,\ n-k) \qquad (k\in\mathbb{N}^*). \tag{4.5$_\pm$}$$

Moreover they satisfy the following recursive relations: for any $k\in\mathbb{N}$,

$$\begin{cases}P_\pm(X^{(q)})(0)=V_\pm(X^{(q)})^{1/2}\\ P_\pm(X^{(q)})(k)=-\sum_{l=0}^{k-1}\gamma_\pm(X^{(q)})(k-l)P_\pm(X^{(q)})(l).\end{cases} \tag{4.6$_\pm$}$$

By virtue of Theorem 6.5 in [7], we can theoretically obtain the algorithms for the limits as $N\to\infty$ of the prediction functions $Q_\pm(X^{(q)})(N+\nu,\ N;\ N-k)$ for any $q,\ \nu\in\mathbb{N}$, $k\in\mathbb{N}^*$: the limits

$$Q_\pm(X^{(q)})(\nu,\ k)\equiv\lim_{N\to\infty}Q_\pm(X^{(q)})(N+\nu,\ N;\ N-k) \tag{4.7$_\pm$}$$

exist and they satisfy the following recursive relations:

$$Q_\pm(X^{(q)})(\nu,\ k)=-\sum_{l=1}^{\nu-1}\gamma_\pm(X^{(q)})(\nu-l)Q_\pm(X^{(q)})(l,\ k)-\gamma_\pm(X^{(q)})(\nu+k). \tag{4.8$_\pm$}$$

References

[1] J. Durbin, The fitting of time series models, Rev. Int. Stat., **28** (1960), 233-244.
[2] N. Levinson, The Wiener RMS error criterion in filter design and prediction, J. Math. Phys., **25** (1947), 261-278.
[3] P. Masani, The prediction theory of multivariate stochastic processes, III, Unbounded spectral densities, Acta Math., **104** (1960), 141-162.
[4] P. Masani and N. Wiener, Non-linear prediction, Probability and Statistics, The Harald Cramér Volume, (ed. U. Grenander), John Wiley, 1959, pp. 190-212.
[5] Y. Okabe, On a stochastic difference equation for the multi-dimensional weakly stationary process with discrete time, Prospect of Algebraic Analysis, (ed. M. Kashiwara and T. Kawai), Academic Press, Tokyo, 1988, pp. 601-645.
[6] Y. Okabe, Langevin equation and causal analysis, Sûgaku, **43** (1991), 322-346 (in Japanese).
[7] Y. Okabe, 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.
[8] Y. Okabe and A. Inoue, The theory of KM_2O-Langevin equations and its applications to data analysis (II): Causal analysis (I), Nagoya Math. J., **134** (1994), 1-28.
[9] Y. Okabe and Y. Nakano, The theory of KM_2O-Langevin equations and its applications to data analysis (I): Stationay analysis, Hokkaido Math. J., **20** (1991), 45-90.
[10] Y. Okabe and O. Ootsuka, The theory of KM_2O-Langevin equations and its applications to data analysis (III): Prediction analysis, in preparation.
[11] P. Whittle, On the fitting of multivariate autoregressions, and the approximate canonical factorization of a spectral density matrix, Biomerika, **50** (1963), 129-134.
[12] N. Wiener, Collected Works, Vol. **3**, The MIT Press, 1981.
[13] N. Wiener and P. Masani, The prediction theory of multivariate stochastic processes, I, The regularity condition, Acta Math., **98** (1957), 111-150.
[14] N. Wiener and P. Masani, The prediction theory of multivariate stochastic processes, II, The linear predictor, Acta Math., **99** (1958), 93-137.
[15] R. A. Wiggins and E. A. Robinson, Recursive solution to the multichannel fitting problem, J. Geophys. Res., **70** (1965), 1885-1891.

Yasunori OKABE
Department of Mathematics
Faculty of Science
Hokkaido University
Sapporo 060
Japan

Present Address
Department of Mathematical Engineering
and Information Physics
Faculty of Engineering
University of Tokyo
Tokyo 113
Japan

Takashi OOTSUKA
Department of Mathematics
Faculty of Science
Hokkaido University
Sapporo 060
Japan

Present Address
High School of
Abashiri Minamigaoka
Abashiri 093
Japan