METHODS FOR OPTIMAL RECURSIVE ESTIMATION OF NON-STATIONARY TIME SERIES, APPLICATIONS TO ATOMIC TIME AND FREQUENCY METROLOGY

Z.Y. Weng and J. Rutman
Laboratoire Primaire du Temps et des Fréquences
61, avenue de l'Observatoire, 75014 Paris, France
and
J. Uebersfeld
Département d'Electronique
Université Pierre et Marie Curie
Place Jussieu, 75005 Paris, France

ABSTRACT

In time and frequency metrology, the problems of characterization, prediction, approximation and modelization are of fundamental importance for theoretical and experimental studies. In this paper, an improved unified approach is proposed and developed, which is based on the optimal estimation theory and the digital recursive processing methods. For two different models of non-stationary time series, the digital recursive methods of optimal estimation are presented. By these unified methods one can synthesize some digital predictors, digital filters and digital differentiators. These digital estimators are used to characterize the frequency instabilities of atomic clocks, to predict the random variations of atomic time scales, and to smooth the time series data. For the modelization of the statistics of frequency and phase fluctuations some analytical procedures are proposed. Then the Markov models of atomic clock instabilities can be deduced. In order to emphasize the utility of the theory, application examples are given for some time comparison data between commercial cesium atomic clocks.

I. INTRODUCTION

In time and frequency metrology, there are some problems of fundamental importance for the theoretical and experimental studies of oscillators and atomic clocks: (1) modelization of the statistics of frequency and phase fluctuations; (2) characterization of frequency and phase instabilities; (3) prediction of random variations of atomic time scales; (4) approximation or smoothing of the frequency and time measurement data.

The commonly used methods of modelization and characterization are based on stationary models as the basic assumption. In general, one must assume the non-stationary models of variations of frequency and time. In this case, the method of structure functions is used for the characterization problem, the least-squares method is used for the approximation problem, and the fixed filter method or the ARIMA method are used for the prediction problem. But the solutions of problems provided by the above mentioned methods are not always completely satisfactory.

In this paper, improved unified methods are proposed and developed, which are based on the optimal estimation theory and the digital recursive processing methods. For the study of non-stationary fluctuations of frequency and phase in atomic clocks, two different approaches are used. The first approach relies on deterministic polynomial models with exponential weighting of data. The second approach utilizes non-stationary stochastic models with stationary increments. The optimal digital recursive methods for the estimation of non-stationary time series are developed for each of these two approaches. By the unified methods one can synthesize digital predictors, digital filters and digital differentiators. These digital recursive estimators are used for solving the problems of characterization, prediction and approximation.

From time comparison data between atomic clocks, one can use the digital differentiators for the characterization of frequency instability. The transfer functions of the differentiators are composed by two operations: pure differentiation and low-pass filtering. This method allows us to estimate the variance function and the power spectral density function. Therefore one can characterize frequency instabilities both in the time and in the Fourier frequency domains only from clock time comparison data.

From time comparison data between atomic clocks, one can use the digital recursive predictors for the prediction of random variations of atomic time scales. In the design of optimal predictors the additive measurement noise is taken into account, which is not negligible for the time comparisons between distant atomic clocks.

Time comparisons between atomic clocks via a satellite provide time series data. The conventional method for the smoothing of time series data is the classical least-squares method. But this method is not suited for the real-time data processing. One can use the digital recursive filters for the smoothing of time series by real-time data processing.

For the modelization of the statistics of frequency and phase fluctuations, some analysis of internal noises in atomic clocks will be given, and some theoretical Markov models of atomic clocks will be deduced. Then some analytical procedures of spectral approximation and model identification will be proposed. One can obtain the corresponding ARIMA models by the method of Z-transformation.

II. METHOD FOR OPTIMAL RECURSIVE ESTIMATION OF NON-STATIONARY TIME SERIES REPRESENTED BY DETERMINISTIC POLYNOMIAL MODELS

2.1 Problem Statement

In time and frequency metrology one often encounters problems where optimal estimators must be determined, which reproduce or transform random signals carrying useful information constituting non-stationary random processes. In this paragraph one will discuss the methods for determining optimal estimators in the case where besides a stationary random process the input data contain a mathematical expectation which can be represented in the form of a polynomial of finite order with respect to time.

When an optimal estimator is determined where the non-stationary part of the signal is present, a more complicated problem must be solved for the conditional minimum of the estimation error. Additional conditions arise from the fact that the mathematical expectation of the signal must be estimated with a given accuracy. At the same time, it is necessary to satisfy the condition of "exponentially fading memory" of the estimator, which reduces to the condition that the signal at the output of the estimator must be formed from the observed values of the input signal with an exponential weighting.

In the solution of the above mentioned problem, the classical Wiener filtering theory is not applicable for two reasons: (1) it assumes that the input signals are stationary random processes with zero mathematical expectations, (2) it assumes that the output signal is formed from an input signal observed over the semi-infinite time interval without weighting. Some authors have solved the filtering problem for non-stationary polynomial inputs with finite observation time. But the optimal filters resulting from their methods are difficult to realize in applications. [15]

In this paragraph the main attention is focussed on the solution of the problem of synthesis of optimal digital estimators for polynomial models with exponential weighting of data. This approach has not been considered until very recently. But it can yield some important results for practical applications.

2.2 Optimization with Exponential Weighting

One assumes that the time series data of measurements is represented by $y(iT)=g(iT)+n(iT)$ and $g(iT)=\sum_{j=0}^{r} g_j (iT)^j$ (2.1)
where $g(iT)$ is the deterministic polynomial component, $n(iT)$ is the stationary random component with autocorrelation function $R_n(lT)$ and power spectral density $S_n(z)$. The output of the estimator is designated by $x(iT)$, and the desired output by $x_d(iT)$. The impulse response of the real estimator is $k(iT)$, and the impulse response of the ideal estimator is $h(iT)$. Therefore one can obtain

$$x_d(lT)=T\sum_{i=-\infty}^{\infty} h(iT)g(lT-iT) \tag{2.2}$$

$$x(lT)=T\sum_{i=0}^{\infty} k(iT)\left[g(lT-iT)+n(lT-iT)\right] \tag{2.3}$$

The error of estimation is $\epsilon(lT)=x_d(lT)-x(lT)=\epsilon_g(lT)+\epsilon_n(lT)$ (2.4)

where

$$\epsilon_g(lT)=T\sum_{i=-\infty}^{\infty} g(lT-iT)h(iT)-T\sum_{i=0}^{\infty} g(lT-iT)k(iT) \tag{2.4a}$$

$$\epsilon_n(lT)=-T\sum_{i=0}^{\infty} n(lT-iT)k(iT) \tag{2.4b}$$

The variance of random error of the estimation is

$$\overline{\epsilon_n^2}=T\sum_{i_1=0}^{\infty} k(i_1 T)T\sum_{i_2=0}^{\infty} R_n(i_1 T-i_2 T)k(i_2 T) \tag{2.5}$$

One can determine the optimal impulse response function $k(iT)$ by the variational calculus with Lagrange multipliers. The minimum of the following expression will be determined.

$$J\{k\}=T\sum_{i_1=0}^{\infty} k(i_1 T)T\sum_{i_2=0}^{\infty} R_n(i_1 T-i_2 T)k(i_2 T)+pT\sum_{i_1=0}^{\infty} g(lT-i_1 T)\left[h(i_1 T)-k(i_1 T)\right] \exp(-ai_1 T) \tag{2.6}$$

The decomposition of $g(lT-i_1 T)$ to the Taylor series gives

$$g(lT-i_1 T)=g(lT)-i_1 Tg'(lT)+\frac{(i_1 T)^2}{2!} g^{(2)}(lT)+\cdots+(-1)^r \frac{(i_1 T)^r}{r!} g^{(r)}(lT) \tag{2.7}$$

Thus one can obtain

$$J\{k\}=T\sum_{i_1=0}^{\infty}k(i_1 T)\Big\{T\sum_{i_2=0}^{\infty}R_n(i_1 T-i_2 T)k(i_2 T)-\left[2p_0+2p_1 (i_1 T)+ +2p_2 (i_1 T)^2+\cdots+2p_r (i_1 T)^r\right] \exp(-ai_1 T)\Big\}+pT\sum_{i_1=0}^{\infty} g(lT-i_1 T)h(i_1 T)\exp(-ai_1 T) \tag{2.8}$$

where $2p_i=\frac{(-1)^i p}{i!} g^{(i)}(lT)$ for $i=0,1,2,\cdots,r$

By the rules of the calculus of variations, the optimal impulse response $k(iT)$, transforming the expression $J\{k\}$ into a minimum, is determined from

$$\frac{\partial}{\partial \Delta} J\{k+\Delta k\}\Big|_{\Delta=0}=0 \qquad (2.9)$$

where Δ is an arbitrary number. This formula is a necessary and sufficient condition for obtaining the minimum of $J\{k\}$. In this way, substituting $k(iT)+\Delta k(iT)$ for $k(iT)$ in (2.8), one obtains $J\{k+\Delta k\}=J\{k\}+2\Delta E_1+\Delta^2 E_2$ (2.10)

where

$$E_2=T\sum_{i_1=0}^{\infty} k(i_1 T)T\sum_{i_2=0}^{\infty} k(i_2 T)R_n(i_1 T-i_2 T) \qquad (2.10a)$$

$$E_1=T\sum_{i_1=0}^{\infty} k(i_1 T)\Big\{T\sum_{i_2=0}^{\infty}k(i_2 T)R_n(i_1 T-i_2 T)-\big[p_0+p_1 i_1 T+p_2 (i_1 T)^2+\cdots+p_r (i_1 T)^r\big]\exp(-ai_1T)\Big\} \qquad (2.10b)$$

Thus one finds that the condition for the minimum of $J\{k\}$ is determined by $E_1=0$, or equivalently, by the following equation :

$$T\sum_{i_2=0}^{\infty}R_n(i_1 T-i_2 T)k(i_2 T)=\big[p_0+p_1 i_1 T+p_2(i_1 T)^2+\cdots+p_r(i_1 T)^r\big]\exp(-ai_1 T) \qquad (2.11)$$

2.3 Solution of the Equation for the Optimal Impulse Response of the Digital Estimator

One will now solve the equation (2.11), which determines the impulse response of an optimal estimator. If the noise $n(iT)$ is a statistically independent time series, then the correlation function of $n(iT)$ can be expressed by $R_n(iT)=R_{n0}\,\delta(iT)$ (2.12)

where $\delta(iT)$ is the Kronecker function defined by the following conditions:

$\delta(iT)=1$ for $i=0$, and $\delta(iT)=0$ for $i\neq 0$ (2.12a)

In this case one can obtain

$$k(iT)=\big[A_0+A_1 iT+A_2(iT)^2+\cdots+A_r(iT)^r\big]\exp(-aiT) \qquad (2.13)$$

where $A_0=R_{n0}^{-1}\,p_0$, $A_1=R_{n0}^{-1}\,p_1$, $A_2=R_{n0}^{-1}\,p_2$, ..., $A_r=R_{n0}^{-1}\,p_r$ (2.13a)

By performing the Z-transformation one can find the transfer function

$$W(z)=T\sum_{i=0}^{\infty}k(iT)z^{-i}=T\sum_{i=0}^{\infty}\big[A_0+A_1 iT+A_2(iT)^2+\cdots+A_r(iT)^r\big]\exp(-aiT)z^{-i} \qquad (2.14)$$

In this formula $A_0, A_1, A_2, \cdots, A_r$ are the constants to be determined.

If the noise $n(iT)$ is a statistically correlated time series, the solution of equation (2.11) is more complicated. One supposes that the spectral density function of the noise $n(iT)$ can be represented by

$$S_n(z)=N(z)N(z^{-1})=\frac{A(z)}{B(z)}=\frac{a_K z^K+a_{K-1} z^{K-1}+\cdots+a_1 z+a_0+a_1 z^{-1}+\cdots+a_{K-1} z^{-K+1}+a_K z^{-K}}{b_m z^m+b_{m-1} z^{m-1}+\cdots+b_1 z+b_0+b_1 z^{-1}+\cdots+b_{m-1} z^{-m+1}+b_m z^{-m}} \qquad (2.15)$$

where $N(z)=\dfrac{E(z)}{C(z)}=\dfrac{e_0+e_1 z+e_2 z^2+\cdots+e_K z^K}{c_0+c_1 z+c_2 z^2+\cdots+c_m z^m}$

One supposes that the function $N(z)$ has neither zero nor pole outside the unit cycle in the Z plane. In this case, it can be demonstrated that the solution of equation (2.11) is the following :

$$k(iT)=\big[A_0+A_1 iT+A_2(iT)^2+\cdots+A_r(iT)^r\big]\exp(-aiT)+B_1 d_1^i+B_2 d_2^i+\cdots+B_{2K} d_{2K}^i \qquad (2.16)$$

In this formula $A_0, A_1, \cdots, A_r, B_1, B_2, \cdots, B_{2K}$ are the constants to be determined, and $d_1, d_2, \cdots, d_{2K}$ are the roots of the equation

$$A(z)=a_K z^K+a_{K-1} z^{K-1}+\cdots+a_1 z+a_0+a_1 z^{-1}+\cdots+a_{K-1} z^{-K+1}+a_K z^{-K}=0 \qquad (2.17)$$

By performing the Z-transformation one can find the transfer function

$$W(z)=T\sum_{i=0}^{\infty}k(iT)z^{-i}=T\sum_{i=0}^{\infty}\big[A_0+A_1 iT+A_2(iT)^2+\cdots+A_r(iT)^r\big]\exp(-aiT)z^{-i}+Tz\sum_{i=1}^{2K}\frac{B_i}{z-d_i} \qquad (2.18)$$

The optimal estimator must satisfy the following additional

condition: $\epsilon_g(lT)=T\sum_{i=-\infty}^{\infty} g(lT-iT)h(iT)-T\sum_{i=0}^{\infty} g(lT-iT)k(iT)=0$ (2.19)
Substituting the formula (2.7) into this condition one can obtain

$$\epsilon_g(iT)=\sum_{l=0}^{r}(d_l-m_l)\frac{(-1)^l}{l!}g^{(l)}(iT)=0 \tag{2.20}$$

where $d_l=T\sum_{i=-\infty}^{\infty}(iT)^l h(iT)$, $m_l=T\sum_{i=0}^{\infty}(iT)^l k(iT)$ (2.20a)
Therefore one must satisfy the following (r+1) conditions:

$$m_l=d_l \text{ or } T\sum_{i=0}^{\infty}(iT)^l k(iT)=T\sum_{i=-\infty}^{\infty}(iT)^l h(iT) \quad \text{for } l=0,1,2,\cdots,r \tag{2.21}$$

In the case of prediction for a time interval $t_o=l_oT$, one can obtain $h(iT)=\delta(iT+l_oT)$. Thus $d_l=T\sum_{i=-\infty}^{\infty}(iT)^l\delta(iT+l_oT)=(-l_oT)^l$ for $l=0,1,2,\cdots,r$
Therefore the (r+1) conditions are the following:

$$m_l=T\sum_{i=0}^{\infty}(iT)^l k(iT)=(-l_oT)^l \quad \text{for } l=0,1,2,\cdots,r \tag{2.22}$$

In the case of filtering one can obtain $h(iT)=\delta(iT)$, $d_l=T\sum_{i=-\infty}^{\infty}(iT)^l\delta(iT)$
Therefore the (r+1) conditions are the following:

$$m_o=1, \quad m_l=0 \quad \text{for } l=1,2,\cdots,r \tag{2.23}$$

In the case of estimation of the first derivative one can obtain the following (r+1) conditions: $m_o=0$, $m_1=-1$, and $m_l=0$ for $l=2,3,\cdots,r$ (2.24)
In the case of estimation of the second derivative one can obtain the following (r+1) conditions: $m_o=0$, $m_1=0$, $m_2=2$, and $m_l=0$ for $l=3,4,\cdots,r$ (2.25)

2.4 Synthesis of Some Digital Recursive Estimators

Using the general method, presented in the sections 2.2 and 2.3, one can synthesize some digital recursive estimators (predictors, filters and differentiators). The results of synthesis are presented in the following table, where the simplified notation $\theta=\exp(-aT)$ is used throughout.

Estimator type	Signal $g(iT)$	Noise $R_n(iT)$	Transfer function $W(z)$
filter	g_o	$R_{no}\delta(iT)$	$\dfrac{z(1-\theta)}{z-\theta}$
filter	$g_o+g_1 iT$	$R_{no}\delta(iT)$	$z\dfrac{(1-\theta^2)z+2\theta^2-2\theta}{(z-\theta)^2}$
differentiator(1)	$g_o+g_1 iT$	$R_{no}\delta(iT)$	$z\dfrac{(z-1)(1-\theta)^2}{T\,(z-\theta)^2}$
filter	$g_o+g_1 iT+g_2(iT)^2$	$R_{no}\delta(iT)$	$z\dfrac{(1-\theta^3)z^2-3\theta(1-\theta^2)z+3\theta^2(1-\theta)}{(z-\theta)^3}$
differentiator(2)	$g_o+g_1 iT+g_2(iT)^2$	$R_{no}\delta(iT)$	$z\dfrac{(z-1)^2(1-\theta)^3}{T^2\,(z-\theta)^3}$
differentiator(1)	$g_o+g_1 iT+g_2(iT)^2$	$R_{no}\delta(iT)$	$z(1-\theta)^2\dfrac{(z-1)[1.5(1+\theta)z-0.5(5\theta+1)]}{T\,(z-\theta)^3}$
predictor for $t_o=l_oT$	$g_o+g_1 iT$	$R_{no}\delta(iT)$	$z\dfrac{(1-\theta^2)(1+l_o\frac{1-\theta}{1+\theta})z-2\theta(1-\theta)\left[1+\frac{l_o}{2\theta}(1-\theta)\right]}{(z-\theta)^2}$

III. METHODS FOR OPTIMAL RECURSIVE ESTIMATION OF NON-STATIONARY TIME SERIES REPRESENTED BY STOCHASTIC MODELS WITH STATIONARY INCREMENTS

3.1 Problem Statement

The random variations of frequency and phase of oscillators and atomic clocks are non-stationary processes. In this paragraph one will develop the methods for optimal recursive estimation of non-stationary time series with stationary increments, represented by the ARIMA models or the Markov models. If the frequency fluctuations of atomic clocks are stationary and the frequency drift is negligible, one must consider the phase (time) fluctuations of atomic clocks as non-stationary processes with the stationary increments of first order. If the frequency fluctuations of atomic clocks are stationary, and the frequency drift has a constant value and is not negligible, one must consider the phase (time) fluctuations of atomic clocks as non-stationary processes with the stationary increments of second order. This stochastic approach is more complicated than the deterministic approach presented in the previous paragraph, because the statistics of random processes must be taken into account. In order to synthesize the optimal estimators one must have the knowledge of the signal and noise statistics. But from the viewpoint of the physical phenomena in the oscillators and atomic clocks, the stochastic approach is more reasonable.

In the solution of the above mentioned problem, the classical Wiener filtering theory is not directly applicable, because it assumes that the signal and the noise are stationary random processes. One will resolve the problem of synthesis of digital estimators by the method of variational calculus. This is a modification and extension of the Wiener's method to the optimization problem for the non-stationary sampled random signals with stationary increments. By this method one can determine the transfer functions of optimal recursive digital estimators. One can also deduce the algorithms for the realization of these digital estimators. In section 3.2 will be presented the direct method of optimal synthesis of digital estimators. In some cases, one can alternatively use the indirect method of synthesis. That is to say that one will synthesize the optimal continuous estimators at first. And then one can obtain the corresponding digital estimators by the transformation methods. The indirect method for synthesis of digital estimators will be presented in section 3.3. In section 3.4 will be presented the method for optimal synthesis of digital estimators for Markov models. For the non-stationary time series with stationary increments and for the case of steady-state optimization, using the general theory of Kalman filtering, one can obtain the time-invariant digital filters. And the transfer functions of these digital recursive filters can be deduced.

3.2 Direct Method for Optimal Synthesis of Digital Estimators

3.2.1 General Method for Optimal Synthesis of Digital Estimators

One assumes that the time series is represented by $y(iT)=u(iT)+n(iT)$ where T is the sampling period, $u(iT)$ is a non-stationary signal with stationary increment of m-th order, $n(iT)$ is a stationary noise. One can represent the problem of optimal estimation by the following diagram.

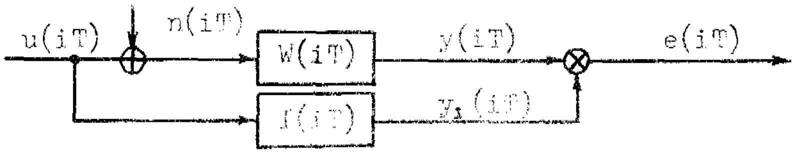


Fig. 3.1

where W(iT) is the impulse response of a real estimator, which is a linear constant system, I(iT) is the impulse response of the ideal estimator. One wishes to minimize the mean-square error

$$\sigma_e^2=\mathcal{E}[e^2(iT)]=\mathcal{E}\left[y_I^2(iT)-2y_I(iT)y(iT)+y^2(iT)\right] \tag{3.2.1}$$

By performing the Z-transformation one can obtain

$$Y_I(z)=U(z)I(z),\qquad Y(z)=[U(z)+N(z)]W(z)$$
$$E(z)=U(z)I(z)-[U(z)+N(z)]W(z)=U(z)[I(z)-W(z)]-W(z)N(z) \tag{3.2.2}$$

Because u(iT) is a non-stationary signal with stationary increment of m-th order, one obtains $U(z)=\dfrac{D(z)}{(1-z^{-1})^m}$ (3.2.3)

where D(z) is a stationary time series. Thus one can obtain

$$E(z)=\frac{D(z)}{(1-z^{-1})^m}[I(z)-W(z)]-W(z)N(z) \tag{3.2.4}$$

If one assumes no correlation between signal and noise, then $S_{dn}(z)=0$, $S_{nd}(z)=0$. The expected value of $E(z)E(z^{-1})$ becomes

$$\mathcal{E}[E(z)E(z^{-1})]=S_{ee}(z)=\frac{S_{dd}(z)}{(1-z^{-1})^m(1-z)^m}[I(z)-W(z)][I(z^{-1})-W(z^{-1})]+W(z)W(z^{-1})S_{nn}(z) \tag{3.2.5}$$

This expression gives the spetral density of the error in terms of the spectral density of the derivative of the signal and the spectral density of the noise. One uses the formula of inversion to obtain the mean-square error as

$$\sigma_e^2=\frac{1}{2\pi j}\oint_{|z|=1}S_{ee}(z)z^{-1}\,dz \tag{3.2.6}$$

One will derive the equation for the best W(z), which minimizes the mean-square error σ_e^2. By using the following short notations

$$W(z)=W_+,\quad W(z^{-1})=W_-,\quad I(z)=I_+,\quad I(z^{-1})=I_- \tag{3.2.7}$$

one obtains

$$\sigma_e^2=\frac{1}{2\pi j}\oint_{|z|=1}\left\{S_{dd}(z)(1-z^{-1})^{-m}(1-z)^{-m}(I_+-W_+)(I_--W_-)+S_{nn}(z)W_+W_-\right\}z^{-1}\,dz \tag{3.2.8}$$

By the rules of variational calculus, to determine the minimum of σ_e^2 one must give a variation $\Delta\eta(z)$ to the transfer function W(z) and find the quantity $\sigma_e^2\{W+\Delta\eta\}$. In this case the optimal transfer function is determined from $\frac{\partial}{\partial\Delta}\sigma_e^2\{W+\Delta\eta\}\big|_{\Delta=0}=0$. In this way one can obtain $\sigma_e^2\{W+\Delta\eta\}=$

$$\frac{1}{2\pi j}\oint_{|z|=1}\left\{\frac{S_{dd}(z)}{(1-z^{-1})^m(1-z)^m}(I_+-W_+-\Delta\eta_+)(I_--W_--\Delta\eta_-)+S_{nn}(z)(W_++\Delta\eta_+)(W_-+\Delta\eta_-)\right\}z^{-1}\,dz \tag{3.2.9}$$

$$\frac{\partial}{\partial\Delta}\sigma_e^2\{W+\Delta\eta\}\Big|_{\Delta=0}=\frac{1}{2\pi j}\oint_{|z|=1}\left\{\frac{S_{dd}(z)}{(1-z^{-1})^m(1-z)^m}\left[(I_+-W_+)(-\eta_-)+(-\eta_+)(I_--W_-)\right]+S_{nn}(z)(W_+\eta_-+W_-\eta_+)\right\}z^{-1}\,dz=0 \tag{3.2.10}$$

This may be written as the sum of two integrals :

$$\frac{1}{2\pi j}\oint_{|z|=1}\eta_-\left\{S_{dd}(z)(1-z^{-1})^{-m}(1-z)^{-m}(W_+-I_+)+S_{nn}(z)W_+\right\}z^{-1}\,dz+$$
$$\frac{1}{2\pi j}\oint_{|z|=1}\eta_+\left\{S_{dd}(z)(1-z^{-1})^{-m}(1-z)^{-m}(W_--I_-)+S_{nn}(z)W_-\right\}z^{-1}\,dz=0 \tag{3.2.11}$$

If, in the second integral, one makes the change of variable $z=-\zeta$

and uses the evenness property of $S_{dd}(z)$ and $S_{nn}(z)$, one can show that the two integrals are identical. Thus one obtains

$$\frac{1}{2\pi j}\oint_{|z|=1} \eta(z^{-1})\left\{S_{dd}(z)(1-z^{-1})^{-m}(1-z)^{-m}\left[W(z)-I(z)\right]+S_{nn}(z)W(z)\right\} z^{-1}\,dz=0 \qquad (3.2.12)$$

Now one defines the spectral factorization as following

$$S_{dd}(z)(1-z^{-1})^{-m}(1-z)^{-m}+S_{nn}(z)=\Delta(z)\,\Delta(z^{-1}) \qquad (3.2.13)$$

and requires that $\Delta(z)$ has poles and zeros inside the unit circle only, and that $\Delta(z^{-1})$ has poles and zeros outside the unit circle only.

Using the notation $\Gamma(z)=S_{dd}(z)(1-z^{-1})^{-m}(1-z)^{-m}\,I(z)$ (3.2.14)

one obtains $$\frac{1}{2\pi j}\oint_{|z|=1} \eta(z^{-1})\Delta(z^{-1})\left[W(z)\,\Delta(z)-\frac{\Gamma(z)}{\Delta(z^{-1})}\right]z^{-1}\,dz=0 \qquad (3.2.15)$$

Expanding the term $\Gamma(z)/\Delta(z^{-1})$ in partial fractions, one obtains

$$\frac{\Gamma(z)}{\Delta(z^{-1})}=\left[\frac{\Gamma(z)}{\Delta(z^{-1})}\right]_{+}+\left[\frac{\Gamma(z)}{\Delta(z^{-1})}\right]_{-} \qquad (3.2.16)$$

where $\left[\frac{\Gamma(z)}{\Delta(z^{-1})}\right]_{+}$ has the poles inside the unit circle, and $\left[\frac{\Gamma(z)}{\Delta(z^{-1})}\right]_{-}$ has the poles outside the unit circle.

Thus $$\frac{1}{2\pi j}\oint_{|z|=1} \eta(z^{-1})\Delta(z^{-1})\left\{W(z)\Delta(z)-\left[\frac{\Gamma(z)}{\Delta(z^{-1})}\right]_{+}-\left[\frac{\Gamma(z)}{\Delta(z^{-1})}\right]_{-}\right\}z^{-1}\,dz=0 \qquad (3.2.17)$$

Note that $\eta(z^{-1})\Delta(z^{-1})\left[\Gamma(z)/\Delta(z^{-1})\right]_{-}$ has the poles outside the unit circle only, then one obtains $$\frac{1}{2\pi j}\oint_{|z|=1} \eta(z^{-1})\Delta(z^{-1})\left[\frac{\Gamma(z)}{\Delta(z^{-1})}\right]_{-}z^{-1}\,dz=0 \qquad (3.2.18)$$

Thus the requirement of realizability of the optimal estimator becomes

$$\frac{1}{2\pi j}\oint_{|z|=1} \eta(z^{-1})\,\Delta(z^{-1})\left\{W(z)\,\Delta(z)-\left[\Gamma(z)/\Delta(z^{-1})\right]_{+}\right\}z^{-1}\,dz=0 \qquad (3.2.19)$$

from which one obtains $$W(z)=\frac{1}{\Delta(z)}\left[\frac{\Gamma(z)}{\Delta(z^{-1})}\right]_{+}$$

3.2.2. Synthesis of some digital recursive estimators

Using the general method, presented in the section 3.2.1 , one can synthesize some digital recursive estimators for the non-stationary time series with stationary increments of m-th order. The results of synthesis of the digital recursive estimators can be represented in the following table.

$\frac{S_{dd}(z)}{(1-z^{-1})^{m}(1-z)^{m}}$	$S_{nn}(z)$	$I(z)$	Transfer function $W(z)$	Parameters values	ARIMA (pnq)
$\frac{d^2}{(1-z^{-1})(1-z)}$	c^2	1	$\frac{1-r_1}{z-r_1}$	$r_1=1+\frac{d^2}{2c^2}-\sqrt{\left(1+\frac{d^2}{2c^2}\right)^2-1}$	(010)
$\frac{d^2}{(1-z^{-1})^2(1-z)^2}$	c^2	1	$\frac{(2-r_1-r_2)z+r_1 r_2-1}{(z-r_1)(z-r_2)}$	$y_1=2+jd/c$, $y_2=2-jd/c$,	(020)
$\frac{d^2}{(1-z^{-1})^2(1-z)^2}$	c^2	$\frac{1-z^{-1}}{T}$	$z\,\frac{(z-1)(1-r_1)(1-r_2)}{T\,(z-r_1)(z-r_2)}$	$r_1=0.5y_1-\sqrt{0.25y_1^2-1}$ $r_2=0.5y_2-\sqrt{0.25y_2^2-1}$	(020)
$\frac{d^2}{(1-z^{-1})^3(1-z)^3}$	c^2	$\frac{(1-z^{-1})^2}{T^2}$	$z\frac{(z-1)^2(1-r_1)(1-r_2)(1-r_3)}{T^2\,(z-r_1)(z-r_2)(z-r_3)}$	r_1, r_2 and r_3 are the roots of equa-	(030)
$\frac{d^2}{(1-z^{-1})^3(1-z)^3}$	c^2	$\frac{1-z^{-1}}{T}$	$\frac{(z-1)(z+\theta')(1-r_1)(1-r_2)(1-r_3)}{T(1+\theta')(z-r_1)(z-r_2)(z-r_3)}$	tion $(z-1)^6-\frac{d^2}{c^2}z^3=0$ inside the unit	(030)
$\frac{d^2}{(1-z^{-1})^3(1-z)^3}$	c^2	1	$\frac{(z^2+\theta_1 z+\theta_0)(1-r_1)(1-r_2)(1-r_3)}{(1+\theta_1+\theta_0)(z-r_1)(z-r_2)(z-r_3)}$	cycle. θ', θ_0 and θ_1 are constants	(030)

3.3 Indirect Method for Optimal Synthesis of Digital Estimators

In some cases, the synthesis of optimal digital estimators is more difficult than the synthesis of corresponding optimal continuous estimators, because the algebraic procedure of spectral factorization is more cumbersome. One can alternatively use the indirect method of optimal synthesis. At first one will synthesize optimal continuous estimators. Then one can obtain the corresponding digital estimators by transformation methods.

3.3.1 General Method for Optimal Synthesis of Continuous Estimators

One defines $u(t)$ as the signal and $n(t)$ as the additive noise. One can represent the problem of optimal estimation by the following diagram.
$W(t)$ is the impulse response of a real estimator, $I(t)$ is the impulse response of an ideal estimator. The criterion of optimization is the mean-square error $\sigma_e^2=\mathcal{E}[e^2(t)]$

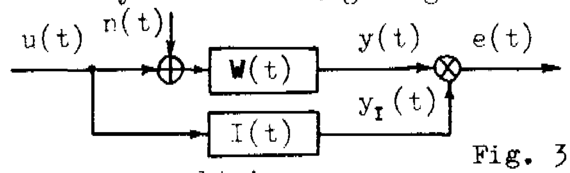


Fig. 3.2

By performing the Laplace transformation one can obtain

$$E(s)=U(s)[I(s)-W(s)]-W(s)N(s) \tag{3.3.1}$$

Because $u(t)$ is a non-stationary signal with a stationary increment of m-th order, one obtains $U(s)=D(s)/s^m$, where $D(s)$ is a stationary signal. One assumes that the spectral densities $S_{dn}(s)=S_{nd}(s)=0$, then the expected value of $E(s)E(-s)$ is

$$\mathcal{E}[E(s)E(-s)]=S_{ee}(s)=\frac{S_{dd}(s)}{s^m(-s)^m}[I(s)-W(s)][I(-s)-W(-s)]+W(s)W(-s)S_{nn}(s) \tag{3.3.2}$$

The mean-square error is $$\sigma_e^2=\frac{1}{2\pi j}\int_{-j\infty}^{j\infty} S_{ee}(s)ds \tag{3.3.3}$$

By the rules of variational calculus one can obtain

$$\frac{1}{2\pi j}\int_{-j\infty}^{j\infty}\eta(-s)\left\{S_{dd}(s)s^{-m}(-s)^{-m}[W(s)-I(s)]+S_{nn}(s)W(s)\right\}ds=0 \tag{3.3.4}$$

Now one defines spectral factorization as the following

$$\frac{S_{dd}(s)}{s^m(-s)^m}+S_{nn}(s)=\Delta(s)\Delta(-s) \tag{3.3.5}$$

and requires that $\Delta(s)$ has poles and zeros inside the left half-plane(LHP) only, and $\Delta(-s)$ has poles and zeros inside the right half-plane (RHP) only. Then one obtains $$\frac{1}{2\pi j}\int_{-j\infty}^{j\infty}\eta(-s)\Delta(-s)\left[W(s)\,\Delta(s)-\frac{\Gamma(s)}{\Delta(-s)}\right]ds=0 \tag{3.3.6}$$

where $\Gamma(s)=S_{dd}(s)s^{-m}(-s)^{-m}I(s)$. Expanding the term $\Gamma(s)/\Delta(-s)$ in partial fractions, one obtains $$\frac{\Gamma(s)}{\Delta(-s)}=\left[\frac{\Gamma(s)}{\Delta(-s)}\right]_+ +\left[\frac{\Gamma(s)}{\Delta(-s)}\right]_- \tag{3.3.7}$$

where $[\Gamma(s)/\Delta(-s)]_+$ has the poles inside the LHP only, and $[\Gamma(s)/\Delta(-s)]_-$ has the poles inside the RHP only. Noting that

$$\frac{1}{2\pi j}\int_{-j\infty}^{j\infty}\eta(-s)\,\Delta(-s)\left[\Gamma(s)/\Delta(-s)\right]_- ds=0 \tag{3.3.8}$$

one can obtain the transfer function of optimal continuous estimator

$$W(s)=\frac{1}{\Delta(s)}\left[\frac{\Gamma(s)}{\Delta(-s)}\right]_+ \tag{3.3.9}$$

3.3.2 Synthesis of Some Optimal Continuous Estimators

Using the general method, presented in the section 3.3.1, one can synthesize some optimal continuous estimators (predictors, filters and differentiators). The results of synthesis are presented in the following table.

ARIMA (pnq)	$\frac{S_{dd}(s)}{s^m(-s)^m}$	$S_{nn}(s)$	$I(s)$	$W(s)$	Parameter values
(110)	$\frac{d^2}{s(-s)(-s^2+\theta^2)}$	c^2	$e^{\tau s}$	$\frac{q_1 s+1}{(ab)^{-1}s^2+(a+b)(ab)^{-1}s+1}$	$a=\left[\frac{\theta^2}{2}+\left(\frac{\theta^4}{4}-\frac{d^2}{c^2}\right)^{1/2}\right]^{1/2}$, (*1)
(110)	$\frac{d^2}{s(-s)(-s^2+\theta^2)}$	c^2	$se^{\tau s}$	$\frac{ks}{(ab)^{-1}s^2+(a+b)(ab)^{-1}s+1}$	$b=\left[\frac{\theta^2}{2}-\left(\frac{\theta^4}{4}-\frac{d^2}{c^2}\right)^{1/2}\right]^{1/2}$, (*2)
(020)	$\frac{d^2}{(-s)^2 s^2}$	c^2	$e^{\tau s}$	$\frac{(\sqrt{2}T_c+\tau)s+1}{T_c^2 s^2+\sqrt{2}T_c s+1}$	$T_c=\sqrt{c/d}$
(020)	$\frac{d^2}{(-s)^2 s^2}$	c^2	s	$\frac{s}{T_c^2 s^2+\sqrt{2}T_c s+1}$	$T_c=\sqrt{c/d}$
(030)	$\frac{d^2}{(-s)^3 s^3}$	c^2	$e^{\tau s}$	$\frac{(2T_c^2+2T_c\tau+\frac{\tau^2}{2})s^2+(2T_c+\tau)s+1}{T_c^3 s^3+2T_c^2 s^2+2T_c s+1}$	$T_c=\sqrt[3]{c/d}$
(030)	$\frac{d^2}{(-s)^3 s^3}$	c^2	$se^{\tau s}$	$\frac{s[(2T_c+\tau)s+1]}{T_c^3 s^3+2T_c^2 s^2+2T_c s+1}$	$T_c=\sqrt[3]{c/d}$
(030)	$\frac{d^2}{(-s)^3 s^3}$	c^2	s^2	$\frac{s^2}{T_c^3 s^3+2T_c^2 s^2+2T_c s+1}$	$T_c=\sqrt[3]{c/d}$
(111)	$\frac{d^2(-s^2+\xi^2)}{s(-s)(-s^2+\theta^2)}$	c^2	$e^{\tau s}$	$\frac{q_1 s+1}{a_1^{-1}s^2+a_1 a_2^{-1}s+1}$	$a_1=\sqrt{2d\xi/c+\theta^2+d^2/c^2}$ $a_2=d\xi/c$, (*3)
(021)	$\frac{d^2(-s^2+\xi^2)}{(-s)^2 s^2}$	c^2	$e^{\tau s}$	$\frac{(a_1 a_2^{-1}+\tau)s+1}{a_2^{-1}s^2+a_1 a_2^{-1}s+1}$	$a_2=d\xi/c$ $a_1=\sqrt{2d\xi/c+d^2/c^2}$
(011)	$\frac{d^2(-s^2+\xi^2)}{(-s)\ s}$	c^2	1	$\frac{1}{T_c s+1}$	$T_c=\sqrt{1+\frac{d^2}{c^2}}\left(\frac{d}{c}\xi\right)^{-1}$

(*1) $q_1=\frac{1}{\theta}\left[1-\frac{ab\exp(-\theta\tau)}{(\theta+a)(\theta+b)}\right]$, (*2) $k=\frac{\exp(-\theta\tau)}{(ab)^{-1}\theta^2+(a+b)(ab)^{-1}\theta+1}$

(*3) $q_1=\frac{1}{\theta}+\frac{a_2(\theta^2-\xi^2)\exp(-\theta\tau)}{\xi^2\theta(\theta^2+a_1\theta+a_2)}$

3.3.3 Calculation of digital estimators by methods of transformation

One has obtained transfer functions of continuous estimators. By the method of Z-transformation or the state variables method one can determine the transfer functions of corresponding digital estimators, which are necessary for the realizations or for computer solutions. One will explain the state variables method only.

Given the transfer function of a continuous estimator $W(s)=\frac{Y(s)}{U(s)}=$

$$=\frac{b_q s^q+b_{q-1}s^{q-1}+\cdots+b_1 s+b_0}{a_p s^p+a_{p-1}s^{p-1}+\cdots+a_1 s+a_0} \quad \text{where } q\leq p-1 \tag{3.3.10}$$

one can find a differential equation of order p . Then one can represent the continuous estimator by a set of first order differential equations, called state variables equations.

$$\dot{X}(t)=FX(t)+GU(t)$$
$$Y(t)=CX(t)+DU(t) \tag{3.3.11}$$

where F,G,C and D are matrices, $\dot{X}(t)$, $X(t)$, $U(t)$ and $Y(t)$ are vectors. Then for the corresponding discret estimator, one can determine the state transition matrix ϕ and the input matrix Γ by the following formulas:

$$\phi=\exp(FT)\ ,\quad \Gamma=\int_0^T\phi(\rho)G\,d\rho \tag{3.3.12}$$

Thus for the discret estimator one can obtain the following state variables equations :

$$X_{k+1}=\phi X_k+\Gamma U_{k+1}$$
$$Y_k=C X_k+D U_k \tag{3.3.13}$$

These equations give the algorithms of digital estimator. They are natural and convenient for computer solutions of problems. In order to analyse the digital estimator, one can perform the Z-transformation of these equations. one obtains $X(z)=(zI-\Phi)^{-1}\Gamma zU(z)$

$$Y(z)=C\,X(z)+D\,U(z)=\left[C(zI-\Phi)^{-1}\Gamma z+D\right]U(z)$$

Therefore the transfer function of the digital estimator is

$$W(z)=Y(z)\left[U(z)\right]^{-1}=C(zI-\Phi)^{-1}\Gamma z+D \tag{3.3.14}$$

One will illustrate this method by two simple examples.

(1) The continuous estimator is $$W(s)=\frac{Y(s)}{U(s)}=\frac{1}{T_c s+1} \tag{3.3.15}$$

Thus $\dot{x}(t)=-\frac{1}{T_c}x(t)+\frac{1}{T_c}u(t)$, $y(t)=x(t)$.

One obtains $F=-1/T_c$, $G=1/T_c$, $\Phi=\exp(FT)=\exp(-T/T_c)$, $\Gamma=1-\exp(-T/T_c)$, $C=1$, $D=0$. The transfer function of the digital estimator is $$W(z)=C(zI-\Phi)^{-1}\Gamma z+D=z\frac{1-\exp(-T/T_c)}{z-\exp(-T/T_c)} \tag{3.3.16}$$

(2) The continuous estimator is $$W(s)=\frac{Y(s)}{U(s)}=\frac{\sqrt{2}\,T_c s+1}{T_c^2 s^2+\sqrt{2}\,T_c s+1} \tag{3.3.17}$$

One obtains the state equations
$$\begin{aligned}\dot{X}(t)&=F\,X(t)+G\,U(t)\\ Y(t)&=C\,X(t)+D\,U(t)\end{aligned} \tag{3.3.18}$$

where $X(t)=\begin{bmatrix}x_1(t)\\ x_2(t)\end{bmatrix}$, $F=\begin{bmatrix}0 & 1\\ -T_c^{-2} & -\sqrt{2}T_c^{-1}\end{bmatrix}$, $G=\begin{bmatrix}0\\ T_c^{-2}\end{bmatrix}$, $C(1\quad \sqrt{2}T_c)$, $D=0$

The state equations for digital estimator are
$$\begin{aligned}X_{K+1}&=\Phi\,X_K+\Gamma\,U_{K+1}\\ Y_K&=C\,X_K+D\,U_K\end{aligned} \tag{3.3.19}$$

where $X_K=\begin{bmatrix}x_K\\ \dot{x}_K\end{bmatrix}$, $X_{K+1}=\begin{bmatrix}x_{K+1}\\ \dot{x}_{K+1}\end{bmatrix}$, $\Phi=\begin{pmatrix}e^{-\alpha}(\cos\alpha+\sin\alpha) & \sqrt{2}T_c\,e^{-\alpha}\sin\alpha\\ -\sqrt{2}T_c^{-1}\,e^{-\alpha}\sin\alpha & e^{-\alpha}(\cos\alpha-\sin\alpha)\end{pmatrix}$

$\Gamma=\begin{pmatrix}1-e^{-\alpha}(\sin\alpha+\cos\alpha)\\ \sqrt{2}\,T_c^{-1}\,e^{-\alpha}\sin\alpha\end{pmatrix}$, $C=(1\quad \sqrt{2}T_c)$, $D=0$, $U_K=u_K$, $Y_K=y_K$, $\alpha=T/(\sqrt{2}T_c)$

The transfer function of the digital estimator is

$$W(z)=C(zI-\Phi)^{-1}\Gamma z+D=z\frac{\left[1+e^{-\alpha}(\sin\alpha-\cos\alpha)\right]z+e^{-2\alpha}-e^{-\alpha}(\sin\alpha+\cos\alpha)}{z^2-2ze^{-\alpha}\cos\alpha+e^{-2\alpha}} \tag{3.3.20}$$

Using this method, one can resolve more complicated problems.

3.4 Optimal Synthesis of Digital Estimators for Markov Models

For the optimal estimation of non-stationary time series, represented by Markov models, one can use the Kalman filtering theory. In general, one obtains digital estimators with time-varying parameters, which are difficult to realize in applications. For the cases of non-stationary time series with stationary increments, and for the cases of steady-state optimization, one can obtain digital estimators with time-invariant parameters, which are easy to realize in applications. One will at first summarize the basic results of Kalman filtering theory, and then develop some aspects of application to the estimation problems of non-stationary time series with stationary increments.

3.4.1 Basic Formulas of Optimal Filtering for the Case of White Noise [17]

One supposes that the state equations of the model are the following
$$\begin{aligned}X_K&=\Phi_{K,K-1}\,X_{K-1}+\Gamma_K W_K\\ Y_K&=H_K\,X_K+V_K\end{aligned} \tag{3.4.1}$$

where W_K and V_K are zero mean white noises, so that $EW_K=0$, $EV_K=0$, $\mathrm{Cov}(W_K,W_j)=Q_K\delta_{Kj}$, $\mathrm{Cov}(V_K,V_j)=R_K\delta_{Kj}$, $\mathrm{Cov}(W_K,V_j)=0$.

Then the linear optimal estimation $\hat{X}_K$ for the time series X_K can be determined by the following recursive formula :

$$\hat{X}_K = \phi_{K,K-1} \hat{X}_{K-1} + K_K (Y_K - H_K \phi_{K,K-1} \hat{X}_{K-1}) \quad (3.4.2)$$

The gain matrix is $K_K = P_{K/K-1} H_K^{\tau} (H_K P_{K/K-1} H_K^{\tau} + R_K)^{-1}$ (3.4.3)

The a priori variance is $P_{K/K-1} = \phi_{K,K-1} P_{K-1} \phi_{K,K-1}^{\tau} + \Gamma_{K-1} Q_{K-1} \Gamma_{K-1}^{\tau}$ (3.4.4)

The posterior variance is $P_K = (I - K_K H_K) P_{K/K-1} (I - K_K H_K)^{\tau} + K_K R_K K_K^{\tau}$ (3.4.5)

The estimation error is $\tilde{Y}_{K/K-1} = Y_K - H_K \phi_{K,K-1} \hat{X}_{K-1}$ (3.4.6)

In general, the matrices $\phi_{K,K-1}$, Γ_K, H_K, Q_K, R_K, K_K, $P_{K/K-1}$ and P_K depend on k, so that one obtains time-varying estimators. But for non-stationary time series with stationary increments, one can have the matrices ϕ, Γ, H, Q and R, which are independent of k. For ease of implementation, one will focus the main attention on the optimization for the steady state, but not for the transient state. In this case, one can obtain the matrices K, P^* and P, which are independent of k. Therefore, one can obtain suboptimal estimators with time-invariant parameters, expressed by the following equations :

$$X_K = \phi X_{K-1} + \Gamma W_K , \quad Y_K = H X_K + V_K \quad (3.4.1a)$$

$$EW_K = 0, \; EV_K = 0, \; Cov(W_K, W_j) = Q \delta_{j-K} , \; Cov(V_K, V_j) = R \delta_{j-K}, \; Cov(W_K, V_j) = 0$$

The estimator equation is $\hat{X}_K = \phi \hat{X}_{K-1} + K (Y_K - H \phi \hat{X}_{K-1})$ (3.4.2a)

The gain matrix is $K = P^* H^{\tau} (H P^* H^{\tau} + R)^{-1}$ (3.4.3a)

The a priori variance is $P^* = \phi P \phi^{\tau} + \Gamma Q \Gamma^{\tau}$ (3.4.4a)

The posterior variance is $P = (I - KH) P^* (I - KH)^{\tau} + KRK^{\tau}$ (3.4.5a)

The estimation error is $\tilde{Y}_{K/K-1} = Y_K - H \phi \hat{X}_{K-1}$ (3.4.6a)

By performing the Z-transformation, one can determine transfer functions of the time-invariant estimators. One obtains

$$X(z) = (zI - \phi)^{-1} \Gamma z W(z), \quad Y(z) = H (zI - \phi)^{-1} \Gamma z W(z) + V(z) \quad (3.4.7)$$

$$\hat{X}(z) = (zI - \phi)^{-1} K z \tilde{Y}(z), \quad \hat{Y}(z) = H \phi z^{-1} \hat{X}(z) = H \phi (zI - \phi)^{-1} K [Y(z) - \hat{Y}(z)] \quad (3.4.8)$$

Thus the transfer function of the digital estimator is

$$A(z) = \hat{Y}(z) [Y(z)]^{-1} = [I + H \phi (zI - \phi)^{-1} K]^{-1} H \phi (zI - \phi)^{-1} K \quad (3.4.9)$$

The open loop transfer function of the digital filter is

$$B(z) = H \phi (zI - \phi)^{-1} K = \hat{Y}(z) / \tilde{Y}(z) \quad (3.4.10)$$

3.4.2 Method for Solution of Optimal Filtering Problems for the Case of Coloured Perturbation

According to the definition, the ARIMA(p,n,q) time series corresponds to a continuous process, expressed by the following state variables equations :

$$\begin{aligned} \dot{X}(t) &= F\, X(t) + G\, L(t) \\ Y(t) &= H\, X(t) + V(t) \end{aligned} \quad (3.4.11)$$

where

$$F = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}, \quad G = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}, \quad X(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix}, \quad \dot{X}(t) = \begin{pmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \vdots \\ \dot{x}_n(t) \end{pmatrix}, \quad H = (1 \; 0 \; 0 \cdots 0)$$

L(t) is a coloured perturbation, V(t) is a white noise. One supposes that

$L(s)=M(s)W(s)$, where $M(s)=N(s)/D(s)$, $W(t)$ is a white perturbation. If all roots of the equation $D(s)=0$ are real and different, the state equations for the coloured perturbation $L(t)$ become

$$\begin{aligned}\dot{P}(t)&=A\,P(t)+B\,W(t)\\ L(t)&=C\,P(t)\end{aligned} \tag{3.4.12}$$

where

$$A=\begin{bmatrix}\theta_1 & & 0\\ & \theta_2 & \\ 0 & & \ddots\ \theta_p\end{bmatrix},\qquad B=\begin{bmatrix}1\\1\\ \vdots\\1\end{bmatrix},\qquad C=(c_1\ \ c_2\ \cdots\ c_p)$$

One can associate the state equations for $L(t)$ to the state equations of the model. One obtains

$$\begin{bmatrix}\dot{X}(t)\\ \dot{P}(t)\end{bmatrix}=\begin{bmatrix}F & GC\\ 0 & A\end{bmatrix}\begin{bmatrix}X(t)\\ P(t)\end{bmatrix}+\begin{bmatrix}0\\ B\end{bmatrix}W(t) \tag{3.4.13}$$

One supposes that $p_1(t)=x_{n+1}(t),\ \cdots,\ p_p(t)=x_{n+p}(t)$, $X^*(t)=\begin{bmatrix}X(t)\\ P(t)\end{bmatrix}$,

$F^*(t)=\begin{pmatrix}F & GC\\ 0 & A\end{pmatrix}$, $G^*=\begin{pmatrix}0\\ B\end{pmatrix}$

Then the extended state equations are the following :

$$\begin{aligned}\dot{X}^*(t)&=F^*X^*(t)+G^*W(t)\\ Y^*(t)&=H\,X^*(t)+V(t),\end{aligned}\quad \text{where } F^*=\begin{pmatrix}F & GC\\ 0 & A\end{pmatrix},\quad G^*=\begin{pmatrix}0\\ B\end{pmatrix} \tag{3.4.14}$$

By the methods of matrix calculus one can determine $\Phi^*=\exp(F^*T)$ and $\Gamma^*=\int_0^T\Phi^*(\psi)G^*d\psi$. Therefore the extended discrete state equations are the following :

$$\begin{aligned}X^*_{K+1}&=\Phi^*X^*_K+\Gamma^*W_{K+1}\\ Y^*_K&=H\,X^*_K+V_K\end{aligned} \tag{3.4.15}$$

Then one can use the basic formulas of section 3.4.1 for resolve the optimal filtering problems for the process model, expressed by the extended discrete state equations.

3.4.3 Calculation of Some Digital Recursive Filters

Using the general methods presented in sections 3.4.1 and 3.4.2, One can design some digital recursive filters. The results are presented in the following table.

ARIMA (pnq)	$\dot{X}(t)=FX(t)+GW(t)$ $Y(t)=HX(t)+V(t)$	$X_K=\Phi X_{K-1}+\Gamma W_K$ $Y_K=H\,X_K+V_K$	K	Transfer function $A(z)$	Coeff. of ϵ_g
(010)	$F=0,\ G=1,\ H=1$	$\Phi=1,\ \Gamma=T,\ H=1$	a	$\dfrac{a}{z-1+a}$	$c_0=0$ $c_1=T/a$
(020)	$F=\begin{pmatrix}0&1\\0&0\end{pmatrix},\ G=\begin{pmatrix}0\\1\end{pmatrix},$ $H=(1\ \ 0).$	$\Phi=\begin{pmatrix}1&T\\0&1\end{pmatrix},$ $\Gamma=\begin{pmatrix}T^2/2\\T\end{pmatrix},$ $H=(1\ \ 0).$	a b/T	$\dfrac{(a+b)z-a}{z^2+(a+b-2)z+1-a}$ where $a=-b/2+\sqrt{2b}$	$c_0=0$ $c_1=0$ $c_2=2T^2/b$
(110)	$F=\begin{pmatrix}0&1\\0&\theta\end{pmatrix},\ G=\begin{pmatrix}0\\1\end{pmatrix},$ $H=(1\ \ 0).$	$\Phi=\begin{pmatrix}1&\theta^{-1}(e^{\theta T}-1)\\0&e^{\theta T}\end{pmatrix},$ $\Gamma=\begin{bmatrix}\theta^{-1}[\theta^{-1}(e^{\theta T}-1)-T]\\ \theta^{-1}(e^{\theta T}-1)\end{bmatrix}$ $H=(1\ \ 0)$	a b/T	$\dfrac{(a+\Gamma_1 b)z-a\Gamma_2}{z^2+(a+\Gamma_1 b-1-\Gamma_2)z+(1-a)\Gamma_2}$ where $\Gamma_1=(\theta T)^{-1}(e^{\theta T}-1)$, $\Gamma_2=\exp(\theta T)$.	$c_0=0$ $c_1=\dfrac{T(1-\Gamma_2)}{a(1-\Gamma_2)+\Gamma_1 b}$
(030)	$F=\begin{pmatrix}0&1&0\\0&0&1\\0&0&0\end{pmatrix},\ G=\begin{pmatrix}0\\0\\1\end{pmatrix},$ $H=(1\ 0\ 0).$	$\Phi=\begin{pmatrix}1&T&T^2/2\\0&1&T\\0&0&1\end{pmatrix},$ $\Gamma=\begin{pmatrix}T^3/6\\T^2/2\\T\end{pmatrix},$ $H=(1\ \ 0\ \ 0)$	a b/T $2c/T^2$	$\dfrac{q_2z^2+q_1z+q_0}{z^3+p_2z^2+p_1z+p_0}$ where $q_2=a+b+c$, $q_0=a$, $p_2=a+b+c-3$, $p_1=-2a-b+c+3$, $q_1=-2a-b+c$, $p_0=-1+a$.	$c_0=0$ $c_1=0$ $c_2=0$ $c_3=3T^3/c$

IV. APPLICATIONS OF DIGITAL RECURSIVE ESTIMATORS TO ATOMIC TIME AND FREQUENCY METROLOGY

In this paragraph, one will demonstrate the applications of digital recursive estimators to characterize the frequency instabilities of atomic clocks, to predict the random variations of atomic time scales, and to smooth time series data, obtained by the comparisons of atomic clocks.

4.1 Characterization of Frequency Instabilities of Atomic Clocks by Recursive Digital Differentiators

4.1.1 Estimation of Variance σ_y^2 by Using Recursive Digital Differentiators

For the characterization of frequency instabilities, phase comparison data of atomic clocks will be used. For the determination of the first derivative of phase fluctuations, one can use the digital differentiator

$$W_1(z) = z\frac{(z-1)}{T}\frac{(1-r_1)(1-r_2)}{(z-r_1)(z-r_2)} = \frac{z-1}{T}W_L(z), \text{ where } W_L(z)=\frac{z(1-r_1)(1-r_2)}{(z-r_1)(z-r_2)} \quad (4.1.1)$$

The corresponding continuous differentiator is

$$W_1(s)=\frac{s}{T_c^2 s^2+\sqrt{2}T_c s+1} = sW_L(s), \text{ where } W_L(s)=\frac{1}{T_c^2 s^2+\sqrt{2}\,T_c s+1} \quad (4.1.2)$$

For the determination of the second derivative of phase fluctuations, one can use the digital differentiator

$$W_2(z)=z\frac{(z-1)^2}{T^2}\frac{(1-r_1)(1-r_2)(1-r_3)}{(z-r_1)(z-r_2)(z-r_3)}=\frac{(z-1)^2}{T^2}W_L(z) \quad (4.1.3)$$

where $W_L(z)=\frac{z(1-r_1)(1-r_2)(1-r_3)}{(z-r_1)(z-r_2)(z-r_3)}$ is the transfer function of a digital low-pass filter. The corresponding continuous differentiator is

$$W_2(s)=\frac{s}{T_c^3 s^3+2T_c^2 s^2+2T_c s+1}=s^2 W_L(s), \text{ where } W_L(s)=\frac{1}{T_c^3 s^3+2T_c^2 s^2+2T_c s+1} \quad (4.1.4)$$

Therefore the transfer functions of optimal differentiators are composed by two operations: pure differentiation and low-pass filtering. The corresponding low-pass filters are the Butterworth filters, which are the approximations to the ideal low-pass filter within the pass-band $\omega_c=1/T_c$.
From the time series y_i obtained at the output of the digital differentiator $W_1(z)$, one can determine the mean value and the variance by the following formulas:

$$m_y=\frac{1}{N}\sum_{i=1}^{N} y_i, \qquad \sigma_y^2=\frac{1}{N-1}\sum_{i=1}^{N}(y_i-m_y)^2 \quad (4.1.5)$$

For each value of the parameter T_c one can obtain a value of the variance σ_y^2. By changing the parameter T_c, one can determine the curve of variance function $\sigma_y^2(T_c)$. When the parameter T_c increases, the variance σ_y^2 decreases. From the curve $\sigma_y^2(T_c)$ one can obtain the curve $\sigma_y^2(\omega_c)$ by a simple computation. The curves $\sigma_y(T_c)$ give the frequency instability characterization in the time domain. (see Appendices, Fig. A1).

From the time series d_i obtained at the output of the digital differentiator $W_2(z)$, one can determine the mean value and the variance by the following formulas:

$$m_d=\frac{1}{N}\sum_{i=1}^{N} d_i, \quad \sigma_d^2=\frac{1}{N-1}\sum_{i=1}^{N}(d_i-m_d)^2 \quad (4.1.6)$$

The values of m_d and σ_d^2 give the frequency drift characterization of oscillators.

4.1.2 Filtering Method for the Power Spectral Density Determination

One supposes that there is a curve of the power spectral density $S_y(\omega)$. The spectral function $F_y(\omega)$ can be determined by $F_y(\omega)=\int_0^{\omega} S_y(\omega)\,d\omega$

And the variance is $\sigma_y^2=\frac{1}{\pi}\int_0^{\infty} S_y(\omega)\,d\omega=\frac{1}{\pi}F_y(\omega)\Big|_{\omega=\infty}$ (4.1.7)

One can obtain also $S_y(\omega)=\dfrac{dF_y(\omega)}{d\omega}$ (4.1.8)

Therefore $F_y(\omega)$ is the primitive function of $S_y(\omega)$, and $S_y(\omega)$ is the first derivative of $F_y(\omega)$. The value of $F_y(\omega)$ for $\omega=\infty$ is the total variance σ_y^2. The value of $F_y(\omega)$ for $\omega=\omega_i$ is a partial variance.

$F_y(\omega_i)=\int_0^{\omega_i} S_y(\omega)\,d\omega=\pi\,\sigma_y^2(\omega_i)$. From the curve $\sigma_y^2(T_c)$ one can determine the curve $\sigma_y^2(\omega_c)$. Then one can compute the spectral function $F_y(\omega_c)=\pi\,\sigma_y^2(\omega_c)$ for different values of ω_c. The power spectral density $S_y(\omega)$ can be obtained by the differentiation of the curve $F_y(\omega_c)$:

$$S_y(\omega_i)=\frac{\Delta F_y(\omega_i)}{\Delta\omega_i}=\frac{F_y(\omega_i)-F_y(\omega_{i-1})}{\omega_i-\omega_{i-1}} \qquad (4.1.9)$$

4.1.3 Relations to the Conventional Characterization Methods

It is of interest to compare the proposed characterization method with the conventional methods, especially with the Allan-variance method and the three samples variance method. In this section, some comparisons between these methods are given.

(1) Comparison with the Allan-variance method

According to the definition of the Allan variance one obtains (see [8])

$$\sigma_y^2(2,\tau,\tau)=\frac{1}{2}\langle(\bar{y}_2-\bar{y}_1)^2\rangle\ .\ \text{Thus}\ \ \sigma_y(\tau)=\frac{1}{\sqrt{2}}(\bar{y}_2-\bar{y}_1) \qquad (4.1.10)$$

where $\bar{y}_1=\frac{1}{\tau}\int_{t_1}^{t_1+\tau}y(\theta)\,d\theta$, $\bar{y}_2=\frac{1}{\tau}\int_{t_2}^{t_2+\tau}y(\theta)\,d\theta$, and $t_2=t_1+\tau$

Hence one obtains the following transfer function

$$H(s)=\frac{1}{\sqrt{2}\tau s}(1-e^{-\tau s})^2,\quad |H(j\omega)|^2=\frac{2\sin^4\pi f\tau}{(\pi f\tau)^2},\quad \text{where } s=j\omega=j2\pi f \qquad (4.1.11)$$

According to the Padé approximation one can obtain (see [16])

$$e^{-\tau s}=\frac{1-s\tau/2}{1+s\tau/2},\quad \text{so that}\quad 1-e^{-\tau s}=\frac{\tau s}{\tau s/2+1} \qquad (4.1.12)$$

Thus the approximate transfer function of the Allan variance method is

$$H(s)=\frac{1}{\sqrt{2}\tau s}(1-e^{-\tau s})^2\approx\frac{1}{2}\tau s H_L(s),\quad \text{where } H_L(s)=\frac{1}{\tau^2s^2/4+\tau s+1} \qquad (4.1.13)$$

The equivalent pass-band of the low-pass filter $H_L(s)$ is $f_e^*=1/(4\tau)$

For the optimal continuous differentiator

$$W_1(s)=\frac{s}{T_c^2s^2+\sqrt{2}\,T_c s+1}=sW_L(s),\quad \text{where}\quad W_L(s)=\frac{1}{T_c^2s^2+\sqrt{2}T_c s+1} \qquad (4.1.14)$$

the equivalent pass-band of the low-pass filter $W_L(s)$ is $f_e=(4\sqrt{2}\,T_c)^{-1}$

Therefore the condition of equivalence is $\tau=\sqrt{2}T_c$ or $T_c=\tau/\sqrt{2}$ (4.1.15)

(2) Comparison with the three samples variance method

According to the definition of the three samples variance method [8] one obtains

$$\sigma_y^2(3,\tau,\tau)=\frac{1}{9}(2\bar{y}_2-\bar{y}_1-\bar{y}_3)^2\ .\ \text{Thus}\ \ \sigma_y(\tau)=\frac{1}{3}(2\bar{y}_2-\bar{y}_1-\bar{y}_3) \qquad (4.1.16)$$

The transfer function of this method is

$$H(s)=\frac{1}{3\tau s}(1-e^{-\tau s})^3,\quad |H(j2\pi f)|^2=\frac{16\sin^6\pi f\tau}{9(\pi f\tau)^2} \qquad (4.1.17)$$

Using the Padé approximation $e^{-\tau s}=\dfrac{1-s\tau/2}{1+s\tau/2}$ and $1-e^{-\tau s}=\dfrac{s}{\tau s/2+1}$ (4.1.18)

one can obtain the approximate transfer function of this method

$$H(s)=\frac{1}{3\tau s}(1-e^{-\tau s})^3\approx\frac{1}{3}\tau^2 s^2 H_L(s),\ \text{where}\quad H_L(s)=\frac{1}{(\tau s/2+1)^3} \qquad (4.1.19)$$

The equivalent pass-band of the low-pass filter $H_L(s)$ is $f_e^*=3/(16\tau)$

For the optimal continuous differentiator

$$W(s)=\frac{s^2}{T_c^3 s^3+2T_c^2 s^2+2T_c s+1}=s^2 W_L(s),\ \text{where}\ W_L(s)=\frac{1}{T_c^3 s^3+2T_c^2 s^2+2T_c s+1} \qquad (4.1.20)$$

the equivalent pass-band of the low-pass filter $W_L(s)$ is $f_e=(6T_c)^{-1}$

Therefore the condition of equivalence is $\tau=\frac{9}{8}T_c$ or $T_c=\frac{8}{9}\tau$ (4.1.21)

4.2 Prediction of Random Variations of Atomic Time Scales by Digital Recursive predictors

4.2.1 Statement of the Atomic Time Scale Prediction Problem

The time scale prediction problem is one of practical importance in such areas as utilization of portable clock data, control of time and frequency at remote autonomous stations, and atomic time scale formation with extrapolation. Several prediction methods have been proposed. They fall into two general classes: fixed polynomial filter methods and autoregressive integrated moving average (ARIMA) methods. By these methods, some results of prediction of atomic time scales have been obtained.[9] [11].

But there are three main problems in prediction, which have not been resolved. (1) The fixed polynomial filter method by the least-squares is not a recursive method. Thus it is not applicable for real-time data processing. (2) According to the ARIMA prediction method of Box and Jenkins, the additive measurement noise is not taken into account. (3) By these two methods one cannot make the prediction of derivatives of time series data. Therefore they are not applicable to the frequency prediction problem.

To resolve the above mentioned problems, one will use digital recursive predictors for atomic time and frequency prediction. Hence the additive measurement noise can be taken into account, and one can predict time and frequency variations simultaneously.

4.2.2 Realization of Digital Recursive Polynomial Predictors with Exponential Weighting of Data

The transfer function of polynomial predictor of second degree is

$$W(z)=\frac{z^{\ell_0}X(z)}{Y(z)}=\frac{(1-\theta^2)[1+l_0(1-\theta)(1+\theta)^{-1}]-2\theta(1-\theta)[1+l_0(1-\theta)(2\theta)^{-1}]z^{-1}}{1-2\theta z^{-1}+\theta^2 z^{-2}} \qquad (4.2.1)$$

Thus the algorithm of this predictor is the following :

$$x_{n+\ell_0}=2\theta x_{n+\ell_0-1}-\theta^2 x_{n+\ell_0-2}+(1-\theta^2)(1+l_0\frac{1-\theta}{1+\theta})y_n-2\theta(1-\theta)(1+l_0\frac{1-\theta}{2\theta})y_{n-1} \qquad (4.2.2)$$

where the prediction time is $t_0=l_0 T$, T is the sampling period.

The transfer function of polynomial predictor of third degree is

$$W(z)=\frac{z^{\ell_0}X(z)}{Y(z)}=\frac{A_0T+\theta(-2A_0T+A_1T^2+A_2T^3)z^{-1}+\theta^2(A_0T-A_1T^2+A_2T^3)z^{-2}}{1-3\theta z^{-1}+3\theta^2 z^{-2}-\theta^3 z^{-3}} \qquad (4.2.3)$$

where $A_0T=(1-\theta^3)\left[1+\frac{3l_0(1-\theta^2)}{2(1+\theta+\theta^2)}+\frac{l_0^2(1-\theta)^2}{2(1+\theta+\theta^2)}\right]$, $A_1T^2=-\frac{3}{2}(1-\theta)^2(1+\theta)\Big[1+$
$+l_0\frac{(1-\theta)(9\theta+1)}{6\theta^2}+l_0^2\frac{(1-\theta)^2(3\theta+1)}{6\theta^2(1+\theta)}\Big]$, $A_2T^3=\frac{(1-\theta)^3}{2}\left[1+l_0\frac{(3\theta+1)(1-\theta)}{2\theta^2}+l_0^2\frac{(1-\theta)^2}{2\theta^2}\right]$

The algorithm of this predictor is the following :

$$x_{n+\ell_0}=3\theta x_{n+\ell_0-1}-3\theta^2 x_{n+\ell_0-2}+\theta^3 x_{n+\ell_0-3}+(A_0T)y_n+\theta(-2A_0T+A_1T^2+A_2T^3)y_{n-1}+\theta^2(A_0T-A_1T^2+$$

$+A_2 T^3)y_{n-2}$ (4.2.4)

4.2.3 Realization of Digital Recursive Predictors Based on Stochastic Models with Stationary Increments

In the paragraph III the synthesis of optimal predictors for several models has been done by the indirect method. At first, the transfer functions for continuous predictors have been obtained. Then one must use the Z-transformation method or the state variables method in order to obtain the corresponding digital recursive predictors.

For the model ARIMA(0,2,0), the optimal continuous predictor is

$$W(s)=\frac{(\sqrt{2}\,T_c+\tau)s+1}{T_c^2 s^2+\sqrt{2}\,T_c s+1} \quad \text{where } T_c=\sqrt{c/d} \tag{4.2.5}$$

For the model ARIMA(0,3,0), the optimal continuous predictor is

$$W(s)=\frac{(2T_c^2+2T_c\tau+\tau^2/2)s+(2T_c+\tau)s+1}{T_c^3 s^3+2T_c^2 s^2+2T_c s+1} \quad \text{where } T_c=\sqrt[3]{c/d} \tag{4.2.6}$$

For the model ARIMA(0,3,0), the optimal continuous predictor of first derivative is

$$W(s)=\frac{s[(2T_c+\tau)s+1]}{T_c^3 s^3+2T_c^2 s^2+2T_c s+1} \quad \text{where } T_c=\sqrt[3]{c/d} \tag{4.2.7}$$

In order to demonstrate the digital realization method of continuous predictors, one will transform a continuous predictor of second degree by using the state variables method. For a given transfer function (4.2.5) one can determine the following state equations:

$$\dot{X}(t)=FX(t)+GU(t)$$
$$Y(t)=CX(t)+DU(t) \tag{4.2.8}$$

where $X(t)=\begin{pmatrix}x_1(t)\\x_2(t)\end{pmatrix}$, $\dot{X}(t)=\begin{pmatrix}\dot{x}_1(t)\\\dot{x}_2(t)\end{pmatrix}$, $F=\begin{bmatrix}0 & 1\\-T_c^{-2} & -\sqrt{2}T_c^{-1}\end{bmatrix}$, $G=\begin{bmatrix}0\\T_c^{-2}\end{bmatrix}$, $C=\begin{bmatrix}1 & (\sqrt{2}T_c+\tau)\end{bmatrix}$, $D=0$. One designates $V_K=\dot{X}_K T$, $V_{K+1}=\dot{X}_{K+1}T$ and $\alpha=T/(\sqrt{2}T_c)$.

The state variables equations for the digital predictor are the following :

$$X_{K+1}=\Phi X_K+\Gamma U_{K+1}$$
$$Y_K=C X_K+D U_K \tag{4.2.9}$$

where $X_K=\begin{bmatrix}x_K\\v_K\end{bmatrix}$, $X_{K+1}=\begin{bmatrix}x_{K+1}\\v_{K+1}\end{bmatrix}$, $\Phi=\begin{bmatrix}e^{-\alpha}(\cos\alpha+\sin\alpha) & \alpha^{-1}e^{-\alpha}\sin\alpha\\-2\alpha e^{-\alpha}\sin\alpha & e^{-\alpha}(\cos\alpha-\sin\alpha)\end{bmatrix}$,

$\Gamma=\begin{bmatrix}1-e^{-\alpha}(\sin\alpha+\cos\alpha)\\2\alpha e^{-\alpha}\sin\alpha\end{bmatrix}$, $C=\begin{bmatrix}1 & (\alpha^{-1}+\tau/T)\end{bmatrix}$, $D=0$.

4.3 Approximation of Non-stationary Time Series Data by Digital Recursive Filters

4.3.1 Statement of the Approximation Problem of Non-stationary Time Series Data

In atomic time and frequency metrology, especially for the time comparison between distant atomic clocks by satellite, one often has necessity to solve the approximation or smoothing problem of time series data. Conventionally, one can use the least-squares method of approximation by algebraic polynomials. In order to simplify the algorithm of computation, one can use the least-squares method of approximation by orthogonal polynomials (Legendre polynomials). In this case, one can save the operation of matrix inversion in the determination of coefficients of polynomials. But these least-squares methods are essentially batch-wise processing methods. They are not well suited for real-time digital processing of time series data.

For the processing of non-stationary time series of long duration and for real-time processing, one must use digital recursive filtering methods. When one compares two distant atomic clocks by a satellite, there are no knowledges of the statistics of the process. In this case, one can use digital recursive polynomial filtering with exponential weighting for the data approximation or smoothing.

4.3.2 Time Series Approximation by Digital Recursive Polynomial Filtering with Exponential Weighting of Data

Using the results, obtained in paragraph II, one can solve the problem of non-stationary time series approximation by digital recursive polynomial filtering. One will demonstrate this application by two simple examples.
(1) For the smoothing of stationary time series data, one can use the digital recursive polynomial filter of first degree. The transfer function of this filter is

$$W(z)=\frac{X(z)}{Y(z)} = z\frac{1-\theta}{z-\theta} = \frac{1-\theta}{1-\theta z^{-1}} \qquad (4.3.1)$$

Thus the algorithm of realization of this filter is the following :

$$x_n = \theta x_{n-1}+(1-\theta)y_n \quad \text{or} \quad x_n = x_{n-1}+(1-\theta)\, e_n \ , \text{ where } \ e_n = y_n - x_{n-1} \qquad (4.3.2)$$

(2) If the time series data contain linear deterministic component, one must use the digital recursive polynomial filter of second degree. The transfer function of this filter is

$$W(z)=\frac{X(z)}{Y(z)} = z\frac{(1-\theta^2)z+2\theta(\theta-1)}{(z-\theta)^2} = \frac{(1-\theta^2)+2\theta(\theta-1)z^{-1}}{1-2\theta z^{-1}+\theta^2 z^{-2}} \qquad (4.3.3)$$

The algorithm of realization of this digital filter is

$$x_n = 2\theta x_{n-1} - \theta^2 x_{n-2} + (1-\theta^2)y_n + 2\theta(\theta-1)y_{n-1} \qquad (4.3.4)$$

One can also use another algorithm for realization of this digital filter, expressed by the following matrix equation :

$$\begin{pmatrix} x_{n+1} \\ v_{n+1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_n \\ v_n \end{pmatrix} + \begin{pmatrix} 1-\theta^2 \\ (1-\theta)^2 \end{pmatrix} e_{n+1} \ , \text{ where } \ e_{n+1} = y_{n+1} - x_n - v_n \qquad (4.3.5)$$

Using this method, one can solve more complicated problems by the same procedure.

V. NON-STATIONARY TIME SERIES MODELIZATION BY DIGITAL RECURSIVE METHODS

5.1 Problem Statement

The problem of atomic time and frequency modelization has a fundamental importance for theoretical and experimental studies. The conventional method for modelization of the statistics of frequency fluctuations is based on stationary models. The commonly used model of frequency instabilities is the power-law spectral density model, expressed by the following formulas:

$$S_y(f) = \sum_{\alpha=-2}^{2} h_\alpha f^\alpha \quad \text{for } 0 \leq f \leq f_h \quad \text{and} \quad S_y(f)=0 \quad \text{for} \quad f > f_h \qquad (5.1.1)$$

Based on this model, the phase fluctuations can be expressed by the following model:

$$S_x(f) = (4\pi^2 f^2)^{-1}\, S_y(f) \qquad (5.1.2)$$

The second widely used model is the deterministic polynomial model. When there is a linear frequency drift, expressed by a first-order polynomial model, the phase drift can be modeled by a second order polynomial.
The third proposed method for modelization of frequency and phase fluctuations is the ARIMA stochastic models. These models can be used for the modelization of non-stationary time series with stationary increments.

But there are some problems in modelization of atomic clocks, which have not been resolved completely.(1) The power-law spectral density models are not completely suited for modelization of non-stationary frequency and phase fluctuations. And the procedures of modelization are not natural and convenient for computer simulations. (2) The deterministic polynomial models can not reflect the statistics of frequency and phase random fluctuations. (3) The ARIMA models have been constructed by empirical methods. The physical origins of different ARIMA models have not been explained and derived. The ARIMA representations of models can be further modified in order to get the Markov representations, which are more natural and convenient for computer solutions and simulations.

In the section 5.2 one will analyze the internal noises of atomic clocks in order to interpret the physical origins of different ARIMA models. Then some theoretical Markov models can be deduced. In the section 5.3, analytical procedures of spectral approximation and model identification will be proposed.

5.2 Analysis of Internal Noises and Deduction of Theoretical Models for Atomic Clocks

5.2.1 Analysis of Internal Noises of Atomic Clocks

Consider the system block diagram of a cesium atomic clock, shown in the following figure.

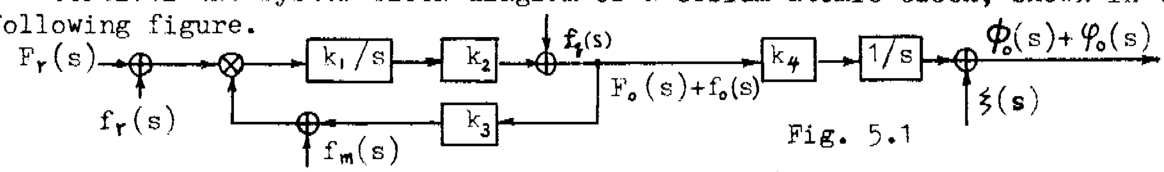


Fig. 5.1

One supposes that the noise of the oscillator $f_q(t)$, the noise of the frequency multiplier $f_m(t)$, and the noise of the atomic reference $f_r(t)$ are uncorrelated. Then the random variations of the frequency of the atomic clock is determined by the following formula :

$$f_o(s)=f_r(s)\frac{k_1k_2}{s+k_1k_2k_3}+f_q(s)\frac{s}{s+k_1k_2k_3}-f_m(s)\frac{k_1k_2}{s+k_1k_2k_3} \quad (5.2.1)$$

The random variations of the phase of the atomic clock is determined by the following formula :

$$\varphi_o(s)=f_o(s)\frac{k_4}{s}+\xi(s) \quad (5.2.2)$$

Using these general formulas, one can analyze frequency and phase fluctuations for different cases of noises $f_r(s)$, $f_q(s)$ and $f_m(s)$. And then one can obtain different theoretical models for atomic clocks. The results of some theoretical analysis can be represented in the following table.

Noise type	$f_r(s)$	$f_q(s)$	$\xi(s)$	$f_o(s)$	$\varphi_o(s)$	ARIMA (p,n,q)
White noises	r	b	0	$\frac{b(s+rb^{-1}k_1k_2)}{s+k_1k_2k_3}$	$\frac{bk_4(s+rb^{-1}k_1k_2)}{s(s+k_1k_2k_3)}$	(1,1,1)
Random walks	r/s	b/s	0	$\frac{b(s+rb^{-1}k_1k_2)}{s(s+k_1k_2k_3)}$	$\frac{bk_4(s+rb^{-1}k_1k_2)}{s^2(s+k_1k_2k_3)}$	(1,2,1)
Freq. drift	0	d/s^2	0	$\frac{d}{s(s+k_1k_2k_3)}$	$\frac{dk_4}{s^2(s+k_1k_2k_3)}$	(1,2,0)
W.N.+ F.D.	r	$b+\frac{d}{s^2}$	0	$\frac{bs^2+rk_1k_2s+d}{s(s+k_1k_2k_3)}$	$k_4\frac{bs^2+rk_1k_2s+d}{s^2(s+k_1k_2k_3)}$	(1,2,2)

5.2.2 Deduction of Markov Models for Atomic Clocks

All formulas, which express the theoretical models of atomic clocks, can be represented by Markov models. Using state variables method, one can transform the formulas of $f_o(s)$ and $\varphi_o(s)$ to matrix equations, which have the Gauss-Markov properties. This method of representation is more convenient for computer simulation and computer solution of problems.

From the formulas $\varphi_o(s)$, expressed by the Laplace transformation, one can obtain the continuous state equations

$$\begin{aligned}\dot{X}(t) &= F\,X(t)+G\,B(t)\\ \varphi_o(t) &= C\,X(t)+D\,V(t)\end{aligned} \qquad (5.2.3)$$

where $B(t)$ is a unity white noise, which generates the stochastic model, $V(t)$ is the noise of observation or additive measurement noise.
Performing the transformation, one can obtain the discrete state equations

$$\begin{aligned}X_{K+1} &= \Phi X_K + \Gamma B_{K+1}\\ \varphi_{oK} &= C\,X_K + D\,V_K\end{aligned} \quad \text{where } \Phi=\exp(FT)\ ,\ \Gamma=\int_0^T \Phi(p)G\,dp \qquad (5.2.4)$$

Performing the Z-transformation, one obtains

$$X(z)=(zI-\Phi)^{-1}\Gamma\, zB(z),\quad \varphi_o(z)=CX(z)+DV(z)=C(zI-\Phi)^{-1}\Gamma\, zB(z)+DV(z) \qquad (5.2.5)$$

Therefore the transfer function of the discrete model is

$$H(z)=\varphi_o(z)/B(z)=C(zI-\Phi)^{-1}\Gamma\, z \qquad (5.2.6)$$

As an example, one will transform a simple atomic clock model to the corresponding Markov model. One supposes that

$$\varphi_o(s)=\frac{bk_4 s+rk_1k_2k_4}{s(s+k_1k_2k_3)} \qquad (5.2.7)$$

One can obtain the continuous state equations

$$\begin{aligned}\dot{X}(t)&=FX(t)+GB(t)\\ \varphi_o(t)&=CX(t)+DV(t)\end{aligned} \quad \text{where } F=\begin{bmatrix}0 & 1\\ 0 & -\alpha\end{bmatrix},\ G=\begin{bmatrix}0\\ 1\end{bmatrix},\ C=(rk_1k_2k_4\ \ bk_4),\ D=0 \qquad (5.2.8)$$

The discrete state equations are

$$\begin{aligned}X_{K+1}&=\Phi X_K+\Gamma B_{K+1}\\ \varphi_{oK}&=C\,X_K+DV_K\end{aligned}$$

$$\text{where } \Phi=\begin{bmatrix}1 & \Omega\\ 0 & \theta\end{bmatrix},\ \Gamma=\begin{bmatrix}\alpha^{-1}(T-\Omega)\\ \Omega\end{bmatrix},\ C=(rk_1k_2k_4\ \ bk_4),\ D=0 \qquad (5.2.9)$$

In the formulas (5.2.8) and (5.2.9) the coefficients are the following :

$$\alpha=k_1k_2k_3\ ,\quad \theta=\exp(-\alpha T)\ ,\quad \Omega=\alpha^{-1}\left[1-\exp(-\alpha T)\right] \qquad (5.2.10)$$

5.3 Analytical Methods of Spectral Approximation and Model Identification

There are two methods for the identification of models of time series data. The time domain method is based on the curves of autocorrelation functions. The frequency domain method is based on the curves of power spectral densities. The commonly used time domain method of identification of general ARIMA models is quite complicated, except for simple autoregressive models. Therefore, it is of interest to study the frequency domain method for identification of models. For a non-stationary time series with stationary increments of n-th order, one must at first take the n-th order differences in order to obtain stationary time series. Then one determines the curves of autocorrelation function and the curves of power spectral density. One will identify the models from the power spectral density curves $S(\omega)$.

5.3.1 Interpolation Method of Spectral Approximation

For a given curve $S(\omega)$ one will make the approximation by the following rational function

$$S^*(\omega)=\frac{B_0+B_1\omega^2+B_2\omega^4+\cdots+B_m\omega^{2m}}{1+A_1\omega^2+A_2\omega^4+\cdots+A_n\omega^{2n}} \qquad (5.3.1)$$

In order to determine the (n+m) coefficients, one supposes that $\omega=\omega_i$ and requires that $S^*(\omega_i)=S(\omega_i)$ for $i=1,2,3,\cdots,(n+m)$. (5.3.2)

Then one obtains

$$(1+A_1\omega_i^2+A_2\omega_i^4+\cdots+A_n\omega_i^{2n})S(\omega_i)-(B_0+B_1\omega_i^2+B_2\omega_i^4+\quad+B_m\omega_i^{2m})=0 \tag{5.3.3}$$

for $i=1,2,3,\cdots,(n+m)$, where $B_0=S(\omega)\big|_{\omega=0}=S_0$

From these (n+m) linear equations one can determine the coefficients $A_1,A_2,A_3,\cdots,A_n$ and $B_1,B_2,B_3,\cdots,B_m$.

For example, one supposes that $S^*(\omega)=\dfrac{B_0+B_1\omega^2}{1+A_1\omega^2+A_2\omega^4}$, where $B_0=S_0$ (5.3.4)

One can determine the coefficients A_1,A_2 and B_1 from the following equations

$$(1+A_1\omega_i^2+A_2\omega_i^4)\ S_i-(S_0+B_1\omega_i^2)=0 \quad \text{for } i=1,2,3. \tag{5.3.5}$$

5.3.2 Least-squares Method of Spectral Approximation

For a given curve $S(\omega)$, one will make the approximation by two steps. The first step is the least-squares approximation by a polynomial as

$$S(\omega)\approx\sum_{i=0}^{n} a_i\omega^{2i}=a_0+a_1\omega^2+a_2\omega^4+\cdots+a_n\omega^{2n} \tag{5.3.6}$$

The second step is the utilization of the Padé approximation method, which yields

$$a_0+a_1\omega^2+a_2\omega^4+\cdots+a_n\omega^{2n}\approx\frac{p_0+p_1\omega^2+p_2\omega^4+\cdots+p_L\omega^{2L}}{1+q_1\omega^2+q_2\omega^4+\cdots+q_M\omega^{2M}} \tag{5.3.7}$$

At the first step one obtains the error of approximation

$d(\omega)=S(\omega)-\sum_{i=0}^{n} a_i\omega^{2i}$. For the minimization of the integral

$D=\int_0^{\Omega}\left[S(\omega)-\sum_{i=0}^{n}a_i\omega^{2i}\right]^2 d\omega$, one obtains the necessary condition $\dfrac{\partial D}{\partial a_j}=0$

for $j=0,1,2,\cdots,n$. Thus one obtains (n+1)equations in order to determine the coefficients $a_0,a_1,a_2,\cdots,a_n$. These equations are the following :

$$\sum_{i=0}^{n} a_i\left[\int_0^{\Omega}\omega^{2(i+j)}d\omega\right]=\int_0^{\Omega}\omega^{2j}S(\omega)d\omega \quad \text{for } j=0,1,2,\cdots,n \tag{5.3.8}$$

For example, one can obtain the following approximation :

$$S(\omega)\approx a_0+a_1\omega^2+a_2\omega^4+a_3\omega^6\approx\frac{B_0+B_1\omega^2}{1+A_1\omega^2+A_2\omega^4}$$

$$\text{where } B_0=a_0,\ B_1=a_1+\frac{a_0a_1a_2-a_0^2a_3}{a_0a_2-a_1^2},\ A_1=\frac{a_0a_3-a_1a_2}{a_0a_2-a_1^2},\ A_2=\frac{a_1a_3-a_2^2}{a_0a_2-a_1^2} \tag{5.3.9}$$

5.3.3 Determination of Parameters of the Model by Spectral Factorization

In order to determine the parameters of the model, one must perform the spectral factorization according to the following formula :

$$S^*(\omega)=\frac{B_0+B_1\omega^2+B_2\omega^4+\cdots+B_m\omega^{2m}}{1+A_1\omega^2+A_2\omega^4+\cdots+A_n\omega^{2n}}=B_0H(j\omega)\ H(-j\omega) \quad \text{where } j\omega=s \tag{5.3.10}$$

For example, if $S^*(\omega)=\dfrac{B_0+B_1\omega^2}{1+A_1\omega^2+A_2\omega^4}$, one can obtain

$$H(s)=\frac{\sqrt{B_0^{-1}B_1}\ s+1}{\sqrt{A_2}\,s^2+(A_1+2\sqrt{A_2})^{1/2}s+1} \tag{5.3.11}$$

Then, by the Z-transformation method or by the state variables method, one can determine the corresponding discrete model ARMA (2,1), expressed by the following formula : $H(z)=\dfrac{b_1z+b_0}{a_2z^2+a_1z+a_0}$ (5.3.12)

VI. CONCLUSIONS

In time and frequency metrology, the conventional used methods are essentially non-recursive methods. They have been developed separately according to the envisaged problems, either for the characterization, or for the prediction, or for the approximation, or for the modelization. The interrelations between these problems and corresponding methods have not been clarified. One finds very little common points of these problems and these used methods.

In this paper, an attempt to unify these problems and the corresponding methods is made. The principal ideas are based on the optimal estimation theory and the digital recursive processing methods. Also the mathematical methods of statistics and linear systems theory have been extensively used. Several methods for optimal recursive estimation of non-stationary time series have been used and developed. Two different models of non-stationarities have been supposed: the deterministic polynomial models and the stochastic models with stationary increments. For these two types of models, one has synthesized optimal digital recursive estimators (predictors, filters and differentiators). One has applied these estimators to the atomic time and frequency metrology. It is clarified that the problems of characterization, prediction, approximation and modelization are particular cases of the general problem of optimal estimation of the states and the parameters. Thus, one can resolve these problems by the unified theory and methods.This new approach allows us to establish the fundamental problems of time and frequency metrology on a sound and rigorous mathematical basis regardless of the user's applications. It can also provide much insight into more complicated problems. Then the interrelations between the proposed methods and the conventional methods can be easily derived.

From the viewpoint of the study of mathematical and statistical methods, the main contributions of this work are the following :

(1) Utilization of the exponential weighting method to resolve the problem of optimal estimation for the deterministic polynomial models. This method allows us to avoid the difficulties and drawbacks in the realization of the finite memory polynomial filters.

(2) Optimal synthesis of digital recursive estimators for the stochastic models of non-stationary time series with stationary increments. This is an extension of the classical Wiener filtering theory to the non-stationary and discrete cases.

(3) Proposition of the design method of suboptimal digital filters according to the general theory of Kalman filtering. For the non-stationary time series with stationary increments and for the case of steady-state optimization, one has obtained some time-invariant digital filters. The transfer functions of these digital filters have been deduced.

(4) Synthesis of different optimal digital recursive estimators(predictors, filters and differentiators) by the unified methods. Though digital filters are widely used in many other domains, the digital predictors and digital differentiators are more particular and are especially important for time and frequency metrology.

From the viewpoint of the applications of statistical methods to atomic time and frequency metrology, the main contributions of this work are the

following :
(1) Application of digital recursive differentiators to characterize the frequency and phase instabilities of atomic clocks. This method allows us to estimate the variance and power spectral density function of frequency instability. Therefore one can characterize frequency instabilities both in the time and in the Fourier frequency domains.
(2) Application of digital recursive predictors to predict the random variations of atomic time scales. Two types of digital predictors can be used : the optimal predictors for deterministic polynomial models with exponential weighting of data, and the optimal predictors for stochastic non-stationary models with stationary increments.
(3) Application of digital recursive filters to smooth the time series data, obtained from time comparisons of distant atomic clocks via a satellite. This method is better than the least-squares method for real-time data processing.
(4) For the modelization of the statistics of frequency and phase fluctuations, one has deduced some theoretical models from the structure of atomic clocks. Then,two analytical procedures of spectral approximation and the spectral factorization method are proposed, which allows us to identify the parameters of stochastic models of atomic clocks.

We have developed several methods for optimal estimation of non-stationary time series. We have also applied these methods to atomic time and frequency metrology. Several new concepts and definitions have been proposed. For the latter, we have pointed out their specific advantages in applications. At present, the existing models and methods for time and frequency metrology are well documented. They provide a good background for actual applications. However, future researches will certainly include the development of more sophisticated approaches, that may possibly improve the methods of time and frequency metrology. In this regard, it seems that the optimal estimation theory can play a key role. This is a very promising approach, because this theory has a rigorous mathematical basis, and its concepts are quite general to cover the domain of time and frequency metrology. Many practical problems can be deduced from this theory as the particular cases. Using the optimal estimation theory, we have tried to solve some fundamental problems of time and frequency metrology. However, the problems of time and frequency metrology are quite diversified and very profound. This rich domain is still widely open for further researches.

REFERENCES

[1] Special Issue on Frequency Stability, Proc.IEEE, vol.54, pp. 101-338, Feb. 1966.

[2] L.S. Cutler and C.L. Searle, "Some aspects of the theory and measurement of frequency fluctuations in frequency standards," in [1], pp. 136-154.

[3] D.W. Allan, "Statistics of atomic frequency standards," in [1], pp. 221-230.

[4] J.A. Barnes, et al. "Characterization of frequency stability", IEEE Trans. Instrum. Meas., vol. IM-20, pp. 105-120, May 1971.

[5] G.M.R. Winkler, "A brief review of frequency stability measures," in Proc. 8th Annu. PTTI Applications and Planning Meeting, Nov.30-Dec.2, 1976,pp. 489-527.

[6] P. Kartaschoff, "Frequency and Time", Academic Press, London, New York, San Francisco, 1978.

[7] W.C. Lindsey, C.M. Chie and W.E. Leavitt, "Interpretation and application of oscillator instability measures using structure functions," in Proc. 8th Annu. PTTI Applications and Planning Meeting, Nov.30-Dec.2, 1976, pp.535-560.

[8] J. Rutman, "Characterization of phase and frequency instabilities in precision frequency sources : fifteen years of progress," Proc.IEEE, vol.66, pp. 1048-1075, Sept. 1978.

[9] D.B. Percival, "Prediction error analysis of atomic frequency standards," in Proc. 31st Annu. Frequency Control Symp.,(Atlantic City,NJ), USA, June 1-3, 1977, pp. 319-326.

[10] J. Uebersfeld, "Bruit de scintillation, stabilité ultime," Seminaire sur les étalons de fréquence, leur caractérisation et leur utilisation. LPMO Besançon France, Du 23 au 25 Mars 1981, pp. U1-U10.

[11] G.E.P. Box and G.M. Jenkins, "Time Series Analysis, Forecasting and Control," San Francisco, CA, Holden-Day, 1976.

[12] Special Issue on System Identification and Time Series Analysis, IEEE Trans. on Automatic Control, vol. AC-19, № 6, Dec. 1974.

[13] R.J. Bennett, "Spatial Time Series : Analysis-Forecasting-Control," Pion Limited, London,1979.

[14] H.J. Blinchikoff and A.L. Zverev, "Filtering in the Time and Frequency Domains." Wiley- Interscience, 1976.

[15] P.D. Krut'ko, "Statistical Dynamics of Sampled- Data Systems" ILIFFE Books LTD, London, 1969.

[16] G.A. Baker, "Essentials of Padé Approximation", Academic Press, New York, 1975.

[17] A.H. Jazwinski, "Stochastic Processes and Filtering Theory" Academic Press, New York and London, 1970.

[18] J.A. Cadzow, "Discrete- Time Systems " Prentice- Hall, Inc., New Jersey , 1973.

[19] R.A. Gabel and R.A. Roberts, "Signals and Linear Systems " John Wiley and Sons, Inc. New York, Toronto, 1980.

[20] J.S. Bendat and A.G. Piersol, "Random Data : Analysis and Measurement Procedures". Wiley-Interscience, New York, 1971.

[21] M. Kunt, "Traitement Numérique des Signaux." Editions Georgi, Lausanne, 1980.

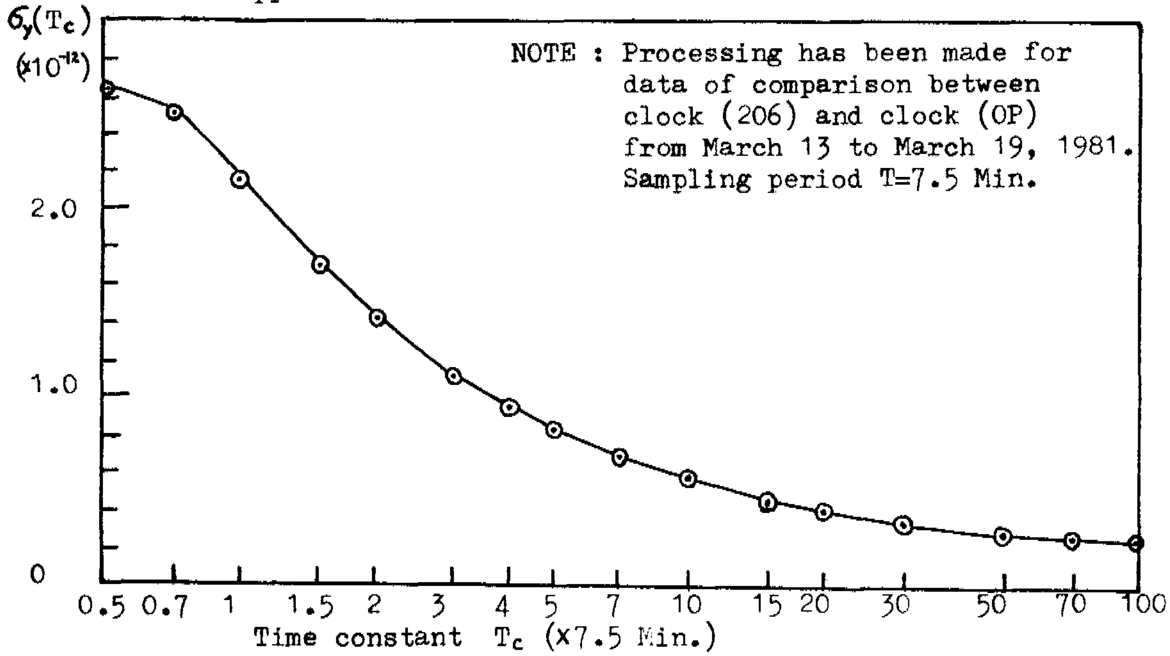


Fig. A1 - Variance function $\sigma_y(T_c)$

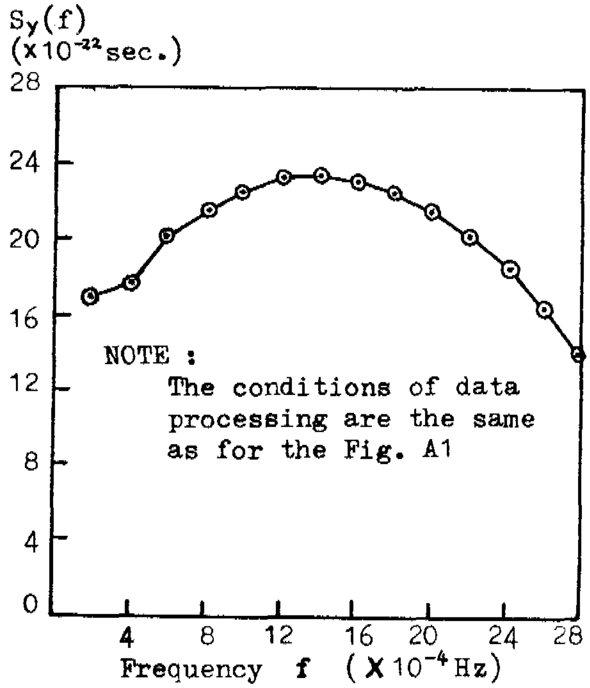
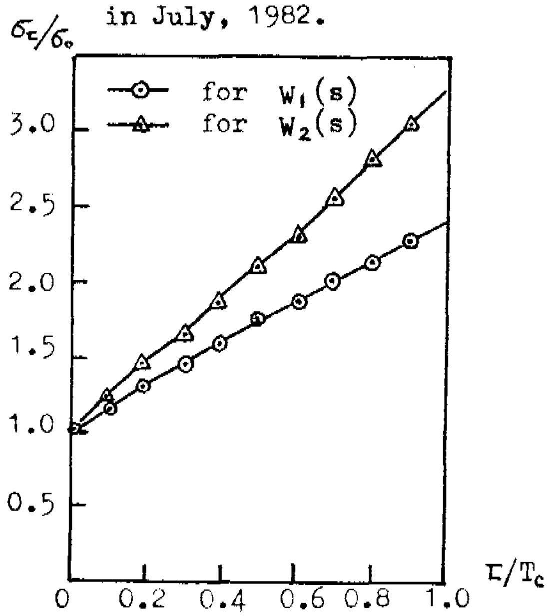


Fig.A2- Power spectral density $S_y(f)$

NOTE: Processing has been made for data of comparison between clock (195) and clock (OP) in July, 1982.



T=30 Min. T/T_c=0.1, σ_0=3.2 ns

Fig.A3 - Variance of time prediction
$W_1(s)$ is determined by (4.2.5)
$W_2(s)$ is determined by (4.2.6)

QUESTIONS AND ANSWERS

None for Paper #25.