1.6 White Noises and Pseudorandom Codes

■ Commonly used input signals for system identification

Reasonable selection of input signals for identification is one of the keys to obtain good identification results.

System identification requirements for input signals:

① Continuous excitation; ② Optimal input signal

Continuous excitation means that the spectrum of the input signal must be sufficient to cover the spectrum of a system.

The optimal input is to minimize a scalar function of the Fisher information matrix inverse:

$$J = \varphi(M^{-1})$$

where M is the Fisher information matrix: $M = E_{y|\theta}\left\{\left[\frac{\partial \log p\,(y|\theta)}{\partial\theta}\right]^T \left[\frac{\partial \log p\,(y|\theta)}{\partial\theta}\right]\right\}$

When $\varphi(.)$ is the determinant $\det(M^{-1})$, J is called the D-optimal criterion.

For the D-optimal criterion, Goodwin and Payne (1977) had the following conclusions:

① If the model structure is correct and the parameter estimate θ is an unbiased minimum variance estimate, then the accuracy of the parameter estimate θ depends on the input signal $u(k)$ via the Fisher information matrix M.

② If the system output data sequence is a Gaussian random sequence with independent and identical distribution, then the D-optimal input signal is a signal with an impulsive autocorrelation function.

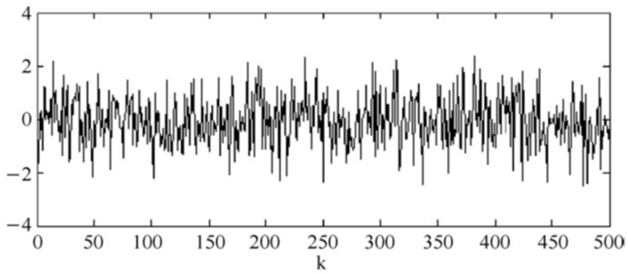
③ When N is large, the M-sequence of a white noise or pseudo-random signal can approximately meet this requirement.

■ White noise

The data used in system identification usually contains a noise.

From engineering practice, this noise can often be regarded as a stationary stochastic process with rational spectral density.

The white noise is the simplest stochastic process.



The mathematical description of a white noise is as follows:

For the stochastic process $\xi(t)$, its mean and autocorrelation function are

$$\mu_\xi(t) = 0$$

$$R_\xi(\tau) = \sigma^2\delta(\tau)$$

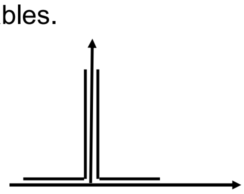
where $\delta(\tau)$ is the unit impulse function (also called Dirac function), namely,

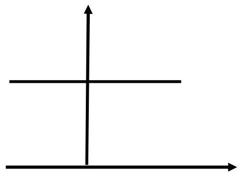
$$\boxed{\delta(t) = \begin{cases} \infty & t = 0 \\ 0 & t \neq 0 \end{cases}} \quad \text{subject to} \quad \int_{0^-}^{0^+} \delta(t)dt = 1$$

This stochastic process $\xi(t)$ is called a white noise.

A white noise is a stationary stochastic process with a mean value of 0 and a non-zero constant spectral density.

In other words, it is an idealized stochastic process composed of a series of uncorrelated random variables.





$R_w(\tau) = \sigma^2 \delta(\tau)$ $\quad$ $S_w(\omega) = \sigma^2 \quad -\infty < \omega < \infty$

Features

No memory, that is, the value at time t has nothing to do with the past value before time t, nor will it affect the future value after time t.

In another sense, stochastic signals at different time are not correlated with each other.

■ White noise sequence

A white noise sequence is a discrete form of a white noise process, which can be described as follows:

For a random sequence $\xi(k)$, its mean is zero and autocorrelation function is

$$R_\xi(k) = \sigma^2 \delta(k), \qquad k = 0, \pm 1, \pm 2, \ldots$$

where $\delta(k)$ is Kronecker function, *i.e.,*

$$\boxed{\delta(k) = \begin{cases} 1 & k = 0 \\ 0 & k \neq 0 \end{cases}}$$

The random sequence $\xi(k)$ is a white noise sequence.

Its spectral density is:

$$S_x(\omega) = \sum_{n=-\infty}^{\infty} R_x(n)e^{jn\omega} = \sigma^2$$

$$R_x(n) = E[x(k)x(k+n)] = \sigma^2\delta(n), \quad n = 0, \pm 1, \pm 2, \cdots$$

The concept of a scalar white noise sequence can be extended to the case of a vector.

A white noise sequence vector $\xi(k)$ is defined as

$$E\{\xi(k)\} = 0$$
$$Cov\ (\xi(k), \xi(k+l)) = E\{\xi(k)\xi^T(k+l)\} = R\delta(l)$$

where R is a positive definite constant matrix, $\delta(l)$ is a Kronecker function.

Self Assessment Question

Which one below is false?

A. Both a white noise sequence and a white noise process are white noises.

B. A random sequence ξ(k) with zero mean is a white noise sequence.

C. The spectral density of a white noise must be a constant.

D. White noise signals at different time are not correlated.

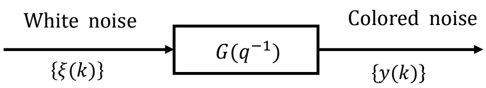
■ Colored noises

From the definition of a white nose, we can see that ideal white noise is only a theoretical abstraction, which cannot be realized physically. In reality, there is no such noise.

The noise contained in measured data in engineering practice is often colored noise.

The so-called colored noise (correlation noise) refers to the correlation between the noise at one time and the noise at another time in the noise sequence.

The "representation theorem" shows that the colored noise sequence can be regarded as the output of linear units driven by the white noise sequence.



Let $G(q^{-1})$ be a linear transfer function, also known as the shaping filter, which can be written as

$$G(q^{-1}) = \frac{C(q^{-1})}{D(q^{-1})}$$

where

$$C(q^{-1}) = 1 + c_1 q^{-1} + \ \dots + c_{n_c} q^{-n_c}$$

$$D(q^{-1}) = 1 + d_1 q^{-1} + \ \dots + d_{n_d} q^{-n_d}$$

and $C(q^{-1})$ and $D(q^{-1})$ are stable polynomials, that is, their roots are within the unit circle of the q-plane.

■ Simulation example

Let the colored noise sequence $e(k)$ be

$$e(k) = \frac{C(q^{-1})}{D(q^{-1})}\xi(k) = \frac{1 + 0.5q^{-1} + 0.2q^{-2}}{1 - 1.5q^{-1} + 0.7q^{-2} + 0.1q^{-3}}\xi(k)$$

where $\xi(k)$ is a white noise with variance of 1.

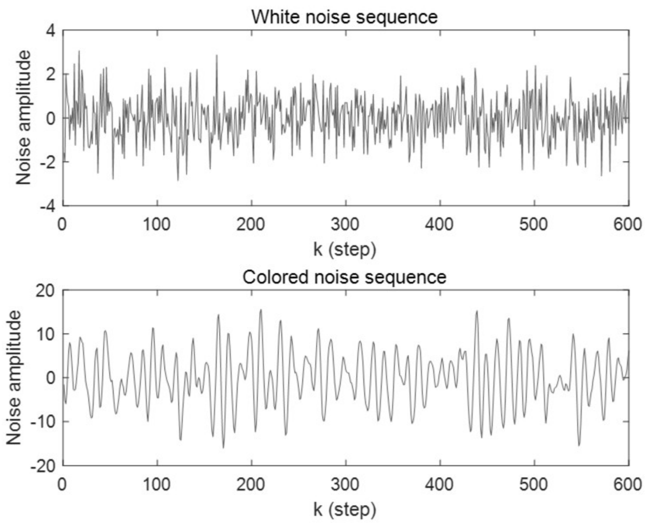
From the transfer function of the colored noise,

$$\begin{aligned} e(k) &= -\left(-1.5q^{-1} + 0.7q^{-2} + 0.1q^{-3}\right)e(k) + \left(1 + 0.5q^{-1} + 0.2q^{-2}\right)\xi(k) \\ &= 1.5e(k-1) - 0.7e(k-2) - 0.1e(k-3) + \xi(k) + 0.5\xi(k-1) + 0.2\,\xi(k-2) \end{aligned}$$

Assume the initials $e(-2), e(-1)$, $e(0), \xi(-1)$ and $\xi(0)$ are all zeros.

Given the white noise sequence $\xi(k)$, the corresponding colored noise sequence $e(k)$ can be calculated.

From a white noise to a colored noise



a white noise $\xi(k)$

$$e(k) = \frac{C(q^{-1})}{D(q^{-1})}\xi(k) = \frac{1 + 0.5q^{-1} + 0.2q^{-2}}{1 - 1.5q^{-1} + 0.7q^{-2} + 0.1q^{-3}}\xi(k)$$

a colored noise $e(k)$

♦ **Simulation codes**

example161.m

Self Assessment Question

Let the colored noise sequence $e(k)$ be

$$e(k) = 1.4e(k-1) - 0.5e(k-2) + 0.2\xi(k-2) + 0.3\xi(k-3)$$

where $\xi(k)$ is a white noise with variance of 1.

Display both $\xi(k)$ and $e(k)$ on the screen at the same time and calculate the mean and variance of the white noise $\xi(k)$.

saq161.m

■ M-sequence and inverse M-sequence

When establishing the mathematical model of a system using the experimental method, if the model structure is selected correctly, the accuracy of the model parameter identification will directly depend on the input signal.

The reasonable selection of the identification input signal is one of the keys to ensuring the ideal identification results.

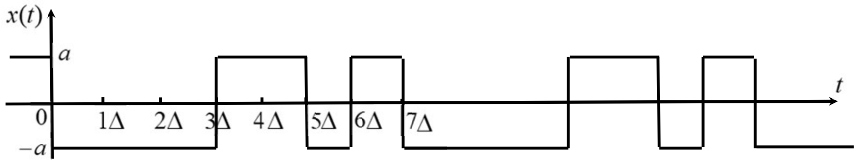
Theoretical analysis shows that selecting a white noise as the identification input signal can ensure a good identification effect, but this is not easy to achieve in engineering practice, because industrial equipment noises cannot act according to the definition of a white noise.

A linear shift register sequence (M-sequence for short) is introduced here.

M-sequence (Maximal length sequence)
— is a good identification input signal
— has the property of an approximate white noise
— can ensure good identification accuracy
— is also easy to implement in engineering

♦ **Pseudo-random binary sequence (PRBS)**

Pseudo-random binary sequence is the most commonly used and easiest to form a pseudo-random signal.



The signal has two levels, respectively $+\mathrm{a} \sim -\mathrm{a}$, which are fixed in the time interval Δ.

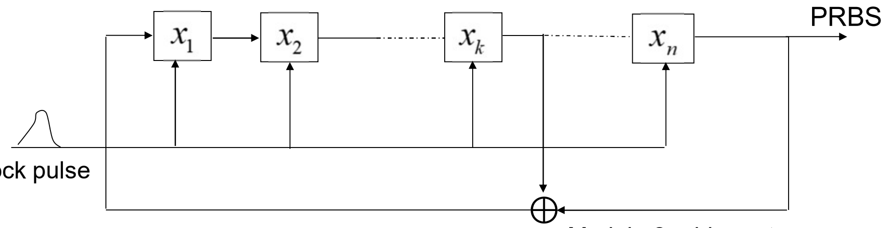
It is a periodic signal with a period of $T = N\Delta$.

In any period, there are $(N+1)/2$ intervals with one level $-\mathrm{a}$ and the other $(N-1)/2$ intervals with the other level $+\mathrm{a}$.

15

♦ **Generation of Pseudorandom Binary Sequences**

Pseudorandom binary sequences can usually be generated using a shift register with feedback



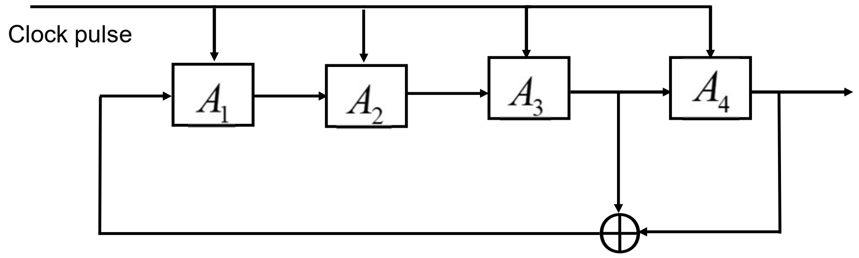
16

Modulo 2 adder gate (exclusive-OR gate):

$$1 \oplus 1 = 0, \quad 0 \oplus 0 = 0$$

$$1 \oplus 0 = 1, \quad 0 \oplus 1 = 1$$

Take the four-stage shift register as an example:

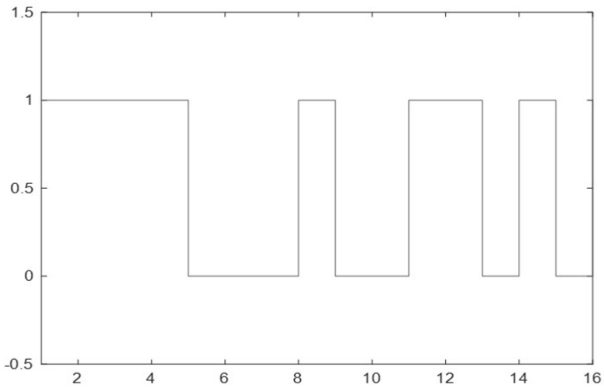


Time step	Output of modulo-2 adder	Output of A_1	Output of A_2	Output of A_3	Output of A_4
0	0	1	1	1	1
1	0	0	1	1	1
2	0	0	0	1	1
3	1	0	0	0	1
4	0	1	0	0	0
5	0	0	1	0	0
6	1	0	0	1	0
7	**1**	**1**	**0**	**0**	1
8	0	1	1	0	0
9	1	0	1	1	0
10	0	1	0	1	1
11	1	0	1	0	1
12	1	1	0	1	0
13	1	1	1	0	1
14	1	1	1	1	0
15	0	1	1	1	1

♦ **Simulation codes**

example162.m

Simulation result



19

The output of register A_4 is

111100010011010

If the modulo 2 adder gate uses the outputs of the second stage and the fourth stage, the output sequence is: 1111100.

In conclusion, for the shift registers with the same number of stages, the cycle lengths of the obtained sequences are different due to different feedback options.

The possible maximum length of PRBS that n-stage shift registers can form is

$$N = 2^n - 1$$

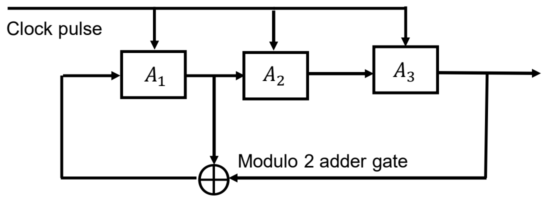
At this time, this sequence is called a binary maximum length sequence, or M-sequence for short.

20

Self Assessment Question

saq162.m

Take the three-stage shift register below to generate a PRBS



Complete the table below:

Time step	Output of modulo-2 adder	Output of A_1	Output of A_2	Output of A_3
0		1	0	1
⋮				
8				

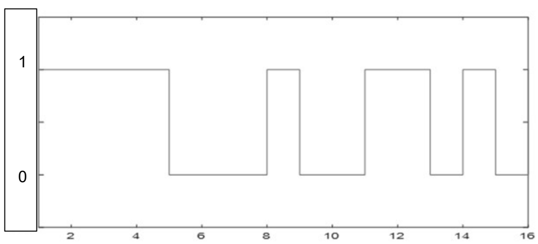
- **Properties of M-sequences**

The period length of the M-sequence generated by the n-stage shift registers is

$$N = 2^n - 1$$

In one cycle of the M sequence, the number of occurrences of the logic 1 state is $(N+1)/2$.

The number of occurrences of the logic 0 state is $(N-1)/2$.



The segment of a sequence in which a state occurs consecutively is called the "run" length. An M-sequence generated by n-stage shift registers has (N+1)/2 segments of "runs" in one cycle.

Properties of a modulo 2 adder

Adding an M-sequence to the sequence delayed by r bits according to the principle of modulo 2 adder, the new sequence obtained is the original M-sequence delayed by q bits.

$$x \oplus D^r x = D^q x \quad r \geq 1,\ q \leq N-1$$

Example:

$$
\begin{array}{rccccccccccccccccl}
x: & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & \\
\oplus D^r x: & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & \quad (r=3) \\
\hline
D^q x: & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & \quad (q=4)
\end{array}
$$

♦ **Autocorrelation function of an M-sequence**

The autocorrelation function of an M-sequence is

$$R_x(\tau) = \begin{cases} a^2\left(1 - \dfrac{N+1}{N}\dfrac{|\tau|}{\Delta}\right) & -\Delta \leq \tau \leq \Delta \\ -a^2/N & \Delta < \tau \leq (N-1)\Delta \end{cases}$$

Proof:

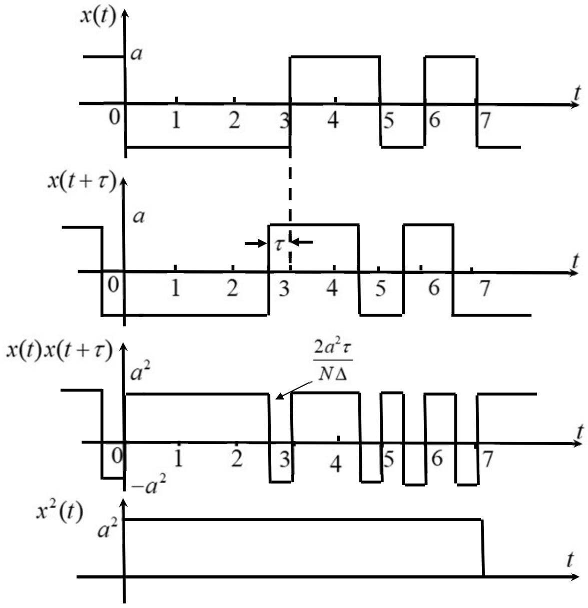
$$R_x(\tau) = \frac{1}{T}\int_0^T x(t)x(t+\tau)\,dt = \frac{1}{N\Delta}\int_0^{N\Delta} x(t)x(t+\tau)\,dt$$

According to the relationship between τ and Δ, the four cases are considered.

$\mathbf{1)}\ \boldsymbol{\tau = 0}$

$$R_x(0) = \frac{1}{T}\int_0^T x^2(t)\,dt = \frac{1}{N\Delta}\int_0^{N\Delta} x^2(t)\,dt = a^2$$

2) $-\Delta < \tau < \Delta$



$$R_x(\tau) = a^2 - \frac{(N+1)}{2}\frac{2a^2|\tau|}{N\Delta} = a^2\left(1 - \frac{N+1}{N}\frac{|\tau|}{\Delta}\right)$$

If $\tau = \Delta$, $\quad R_x(\Delta) = -a^2/N$

3） $\boldsymbol{\tau = \mu\Delta}$, $\quad \boldsymbol{\mu = 1, 2, \cdots N-1}$

$$R_x(\tau) = \frac{1}{N\Delta}\int_0^{N\Delta} x(t)x(t+\tau)\,dt = \frac{1}{N}\sum_{k=0}^{N-1} x(k\Delta)\,x(k\Delta+\mu\Delta)$$

There are α cases: $\quad x(k\Delta)x(k\Delta+\mu\Delta) = a^2$

There are β cases: $\quad x(k\Delta)x(k\Delta+\mu\Delta) = -a^2$

$$R_x(\tau) = \frac{a^2}{N}(\alpha - \beta)$$

According to the isomorphic relationship between $x(k\Delta)x(k\Delta+\mu\Delta)$ and $x(k\Delta) \oplus x(k\Delta+\mu\Delta)$, the multiplication of two binary sequences with magnitude a is equivalent to the new sequence with magnitude a^2 obtained after adding the two sequences by a modulo 2 adder.

$x(k\Delta)$	$x(k\Delta+\mu\Delta)$	$x(k\Delta) \oplus x(k\Delta+\mu\Delta)$	$x(k\Delta)x(k\Delta+\mu\Delta)$
a, (0)	a, (0)	a, (0)	a^2 , (0)
a, (0)	-a, (1)	-a, (1)	$-a^2$, (1)
-a, (1)	a, (0)	-a, (1)	$-a^2$, (1)
-a, (1)	-a ,(1)	a, (0)	a^2 , (0)

The new sequence is still an M-sequence (shift-add property) with $(N\text{-}1)/2$ logic 0 and $(N+1)/2$ logic 1 in one cycle.

27

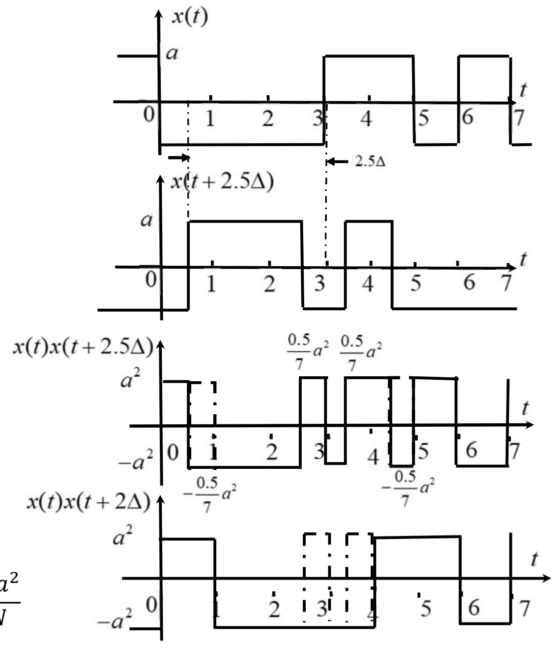
$$\alpha = \frac{N-1}{2},\ \beta = \frac{N+1}{2}$$

4) $\boldsymbol{\mu\Delta < \tau < (\mu\Delta + \Delta)}$

μ is an integer, $\mu \neq 0, \mu \neq N-2$

Assume $\mu = 2, \tau = 2.5\Delta$

$$R_x(\tau) = \frac{-a^2}{N}$$



28

1) $\tau = 0$

$$R_x(0) = a^2$$

2) $-\Delta < \tau < \Delta$

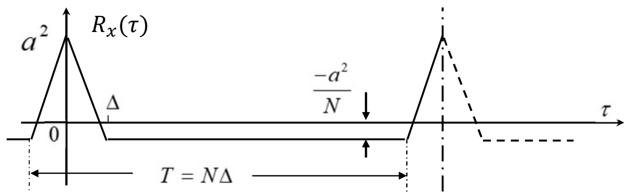
$$R_x(\tau) = a^2\left(1 - \frac{N+1}{N}\frac{|\tau|}{\Delta}\right)$$

3) $\tau = \mu\Delta \quad \mu = 1,2,\cdots N-1$

$$R_x(\tau) = \frac{-a^2}{N}$$

4) $\mu\Delta < \tau < (\mu+1)\Delta$

$$R_x(\tau) = \frac{-a^2}{N}$$



$R_x(\tau)$ is a periodic even function.

Self Assessment Question

For a PRBS with $\Delta = 1, a = 1$ and $N = 15$, draw the cure of its autocorrelation function $R_x(\tau)$.

♦ **Spectral density of an M-sequence**

Knowing the spectral density of an M-sequence, one can directly estimate what M-sequence needs to be selected according to the frequency band of an identification system.

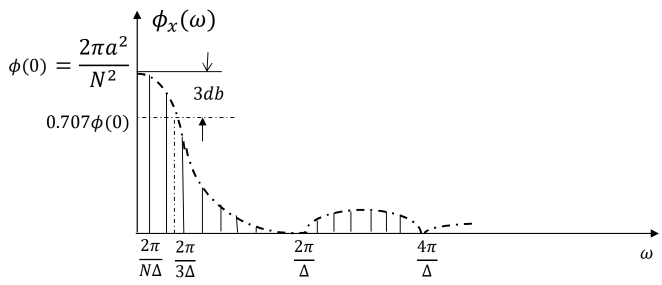
$$\phi_x(\omega) = \frac{1}{2}\int_{-\infty}^{+\infty} R_x(\tau)e^{j\tau\omega}d\tau$$

$$= \frac{2\pi a^2}{N^2}\delta(\omega) + \frac{2\pi a^2(N+1)}{N^2}\sum_{n=-\infty, n\neq 0}^{\infty}\left[\frac{\sin(n\omega_0\,\Delta/2)}{n\omega_0\,\Delta/2}\right]^2\delta(\omega - n\omega_0)$$

where $\omega_0 = \dfrac{2\pi}{T} = \dfrac{2\pi}{N\Delta}$

1) For $n = N, 2N, \cdots,\quad \phi_x(\omega) = 0$

The M-sequence does not contain integer multiple frequency components of the fundamental frequency ω_0,

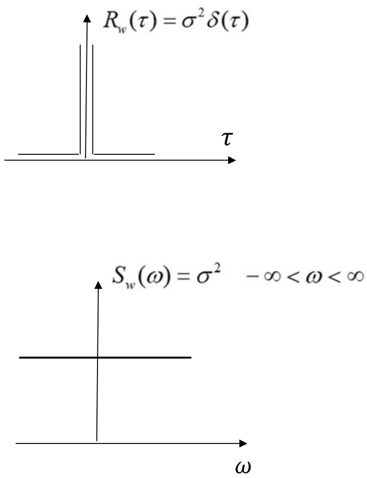


2) For $\omega = \dfrac{2\pi}{3\Delta}$,

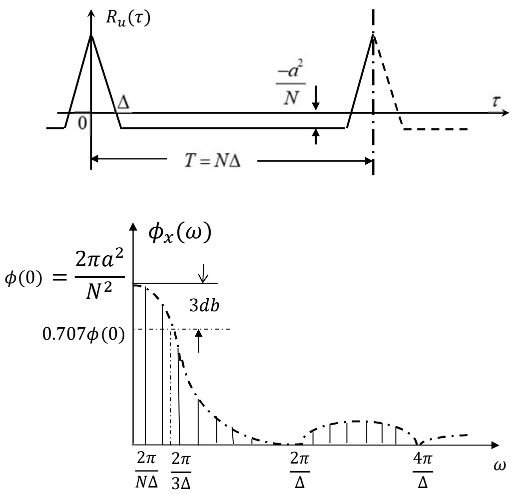
$$\left(\frac{\sin(\,\omega\Delta/2)}{\omega\Delta/2}\right)^2 = 0.707$$

⟹ $\phi_x(\omega)$ with a bandwidth of $\omega_0 = \dfrac{2\pi}{N\Delta}$ ~ $\omega = \dfrac{2\pi}{3\Delta}$

♦ **White noise and M-sequence**



white noise



M-sequence

♦ **Selection of M-sequence parameters in practical applications**

1) Choose a, according to the allowable signal-to-noise ratio of a system being identified

2) Choose Δ ⟹ $\Delta < \dfrac{2\pi}{3\omega_{\max}}$ where $\omega_{\max}$ is the maximal working frequency.

3) Choose N ⟹ $N = (1.2{\sim}1.5)T_s/\Delta$ where T_s is the settling time.

♦ **Using the M-sequence as an input to identify the impulse response**

$$R_{xy}(\tau) = \int_0^{\infty} g(\sigma)\, R_x(\tau - \sigma) d\sigma$$

⇓ M sequence is a periodic signal

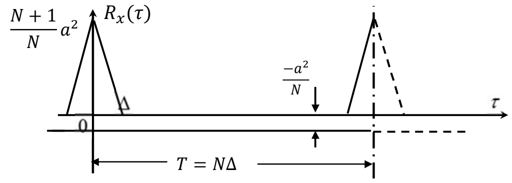
$$R_{xy}(\tau) = \int_0^{T} g(\sigma)\, R_x(\tau - \sigma) d\sigma$$

Note: The details of obtaining the above formula will be given in Section 3.3.)

$$R_x(\tau) = R_x^{(1)}(\tau) + R_x^{(2)}(\tau)$$

where $R_x^{(1)}(\tau) = \begin{cases} a^2(1+\frac{1}{N})(1-\frac{|\tau|}{\Delta}) & -\Delta \le \tau \le \Delta \\ 0 & \Delta < \tau \le (N-1)\Delta \end{cases}$

$R_x^{(2)}(\tau) = -a^2/N$



$R_x(\tau)$ ⇓ approximation

$$R_x(\tau) \cong \frac{N+1}{N} a^2 \Delta \delta(\tau) - \frac{a^2}{N}$$

$$\Rightarrow \quad R_{xy}(\tau) \cong \int_0^T \left[\frac{N+1}{N} a^2 \Delta \delta(\tau - \sigma) - \frac{a^2}{N}\right] g(\sigma) d\sigma$$

$$= \frac{N+1}{N} a^2 \Delta g(\tau) - \frac{a^2}{N} \int_0^T g(\sigma) d\sigma$$

If the period of the M-sequence is set to be greater than the transition time of the identified system, the impulse response of the identified system will basically decay to zero after the time is greater than the period of the M-sequence.

Therefore, for a stable system, $\frac{a^2}{N}\int_0^T g(\sigma)d\sigma$ can be regarded as a bounded constant C.

Let $S = \frac{N+1}{N} a^2 \Delta$, then

$$g(\tau) = (R_{xy}(\tau) + C)/S$$

When N is large, C is small and negligible.

Note:

$g(0) = 2(R_{xy}(0) + C)/S$ (because it starts from time zero)

Calculation of the cross-correlation function: $R_{xy}(\tau) = \frac{1}{T}\int_0^T x(t)y(t+\tau)dt$

If $x(t)$ and $y(t)$ are approximated by steps with step Δ, then

$$R_{xy}(\tau) \cong \frac{1}{N}\sum_{i=0}^{N-1} x(i\Delta)y(i\Delta+\tau)$$

Since the M-sequence can only take the value $\pm a$, $R_{xy}(\tau) \cong \frac{a}{N}\sum_{i=0}^{N-1}[\mathrm{sign}(x(i\Delta)]y(i\Delta+\tau)$

To improve the accuracy of $R_{xy}(\tau)$, several cycles of the M-sequence can be measured.

For example, if the M-sequence of r cycles is continuously measured as an input, then

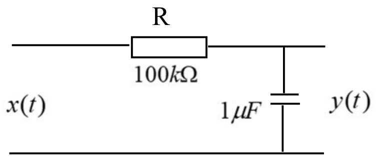
$$R_{xy}(\tau) = \frac{a}{rN}\sum_{i=0}^{rN-1}[\mathrm{sign}(x(i\Delta)]y(i\Delta+\tau)$$

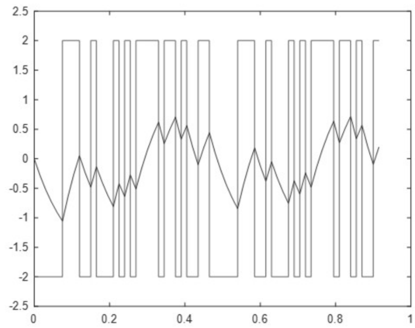
Self Assessment Question

What conditions should the parameters $(\mathrm{N}, \mathrm{a}, \Delta)$ of an M-sequence satisfy so that

$g(\tau) = R_{xy}(\tau)$?

Example: Consider the circuit below.



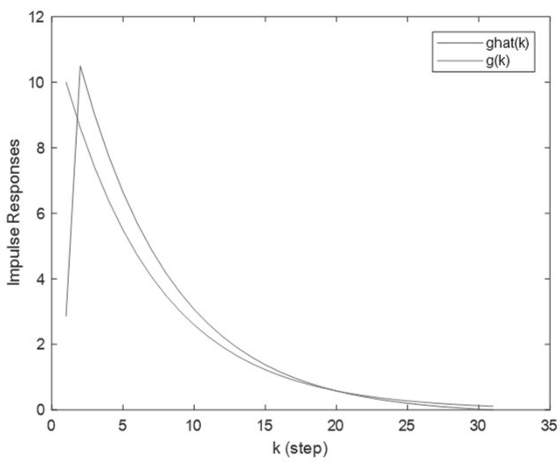


Impulse response: $g(t) = \frac{1}{\tau}e^{-t/\tau} = 10e^{-10t}$

A 5-stage shift register is used to generate an M-sequence as an input signal, and the impulse response of the system is identified.

The related parameters of the M-sequence are set as: $a = 2, \Delta = 15ms$

M-sequence and corresponding system response



$$S = \frac{N+1}{N}a^2\Delta$$

$$\hat{g}(i\Delta) = (R_{xy}(i\Delta) + C)/S$$

$$\hat{g}(0) = 2(R_{xy}(0) + C)/S$$

$$R_{xy}(\tau) \cong \frac{a}{N}\sum_{i=0}^{N-1}[sign(x(i\Delta)]y(i\Delta + \tau)$$

Flowchart of the system identification

Initialize $R, C, n, a, del, Total$

↓

Generate M-sequence Out=genPRBS(n,a,del,total) as a system input

↓

Produce system output y

↓

Calculate system input and output correlation function Rxy

↓

Calculate the system impulse estimate ghat and the system real impulse output g

♦ **Simulation codes**

example163.m

Self Assessment Question

Generate an M-sequence as an input signal using a 5-stage shift register. The related parameters of the M-sequence are $a = 1, \Delta = 30ms$.

Plot both the identified and true impulse response of the system below and also the M-sequence.

$$G(s) = \frac{100}{s^2 + 10s + 100}$$

saq163.m

Wiener-Schinchin relation

The spectral density $S_x(\omega)$ of a stochastic process $x(t)$ and autocorrelation function $R_x(\tau)$ form a Fourier transform pair

$$S_x(\omega) = \int_{-\infty}^{\infty} R_x(\tau) e^{-j} \quad d\tau$$

$$R_x(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_x(\omega) e^{j\omega\tau} \, d\omega$$

Define the cross-spectral density $S_{xy}(\omega)$ as the Fourier transform of the cross-correlation function $R_{xy}(\tau)$

$$S_{xy}(\omega) = \int_{-\infty}^{\infty} R_{xy}(\tau) e^{-j\omega\tau} \, d\tau$$

$$R_{xy}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{xy}(\omega) e^{j\omega\tau} \, d\omega$$

43

♦ **Generic M-sequences**

There is an infinite binary sequence $x_1, x_2, \dots, x_p, x_{p+1}, \dots.$, and exist the following relationships among the elements:

$$x_i = a_1 x_{i-1} \oplus a_2 x_{i-2} \oplus \cdots \oplus a_p x_{i-p}$$

where $i = p+1, p+2, \dots,$
the coefficients $a_1, a_2, \dots, a_{p-1}$ are 0 or 1,
the coefficient a_p is always 1,
$\oplus$ is an exclusive operator.

As long as the coefficients $a_1, a_2, \dots, a_{p-1}$ are properly selected, the sequence can cycle with the longest cycle of $2^p - 1$ bits.

This binary sequence with the longest cycle period is called M-sequence.

How to select coefficients $a_1, a_2, \dots, a_{p-1}$ to ensure the generation of M-sequences and the properties of M-sequences, refer to [*].

[*] 方崇智，萧德云，过程辨识，北京:清华大学出版社，1988.

44

♦ **Inverse M-sequence**

The spectrum analysis shows that the M-sequence contains DC components, which will cause "net disturbance" to the identification system, which is usually not expected.

The inverse M-sequence will overcome this shortcoming and is a more ideal pseudorandom code sequence than an M-sequence.

Algorithm of the inverse M-sequence

Let M(k) be an M-sequence with a period of N_P bit and an element value of 0 or 1, and S(k) be a square wave sequence with a period of 2 bit and an element value of 0 or 1 in turn.

Carry out the XOR operation on the two sequences by bit.

The resultant composite sequence is an inverse M-sequence with a period of $2N_P$ bit and an element value of 0 or 1, *i.e.,*

$$IM(k) = M(k) \oplus S(k)$$

Transform the logical values 0 and 1 of the inverse M-sequence into -1 and 1, respectively.

Features of the inverse M-sequence

— The mean value of the inverse M-sequence is 0.

— Although the inverse M-sequence is only the result of the simple combination of an M-sequence and a square wave sequence, its properties are better than the M-sequence.

— It is more widely used in the field of identification.

■ Simulation example

Let the M-sequence be generated by the following 4-bit shift register:

$$x_i = x_{i-3} \oplus x_{i-4}$$

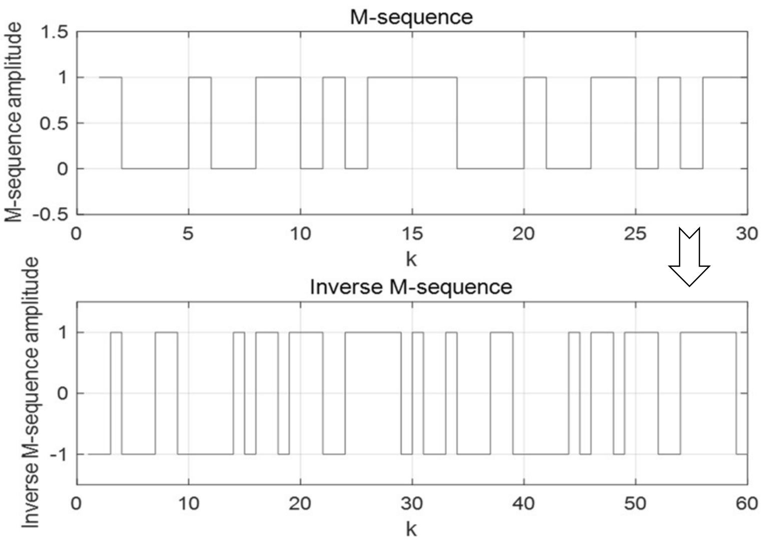
Assuming that the initial value of shift register $x(-3) = 0, x(-2) = x(-1) = x(0) = 1,$ and square wave $S(1)$=1, then

$$x_1 = x_{-2} \oplus x_{-3} = 0 \oplus 1 = 1, \qquad IM(1) = x_1 \oplus S(1) = 1 \oplus 1 = 0$$
$$x_2 = x_{-1} \oplus x_{-2} = 1 \oplus 1 = 0, \qquad IM(2) = x_2 \oplus S(2) = 0 \oplus 0 = 0$$

The M sequence and inverse M sequence can be obtained by iterative calculation.
The cycle of M-sequence is 15 bits, and the cycle of inverse M-sequence is 30 bits.

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
x_i	1	0	0	0	1	0	0	1	1	0	1	0	1	1	1	1	0	0	0	1	0	0	1	1	0	1	0	1	1	1
S	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0
IM	0	0	1	0	0	0	1	1	0	0	0	0	0	1	0	1	1	0	1	1	1	0	0	1	1	1	1	1	0	1
$\overline{IM}$	−1	−1	1	−1	−1	−1	1	1	−1	−1	−1	−1	−1	1	−1	1	1	−1	1	1	1	−1	−1	1	1	1	1	1	−1	1

From a M-sequence to an inverse M-sequence



M-sequence $M(k)$

$IM(k) = M(k) \oplus S(k)$

Inverse M-sequence $IM(k)$

♦ **Simulation codes**

example164.m

Self Assessment Question

Let the M-sequence be generated by the following 5-shift register with initial value of [1,0,1,1,1]:

$$x_i = x_{i-1} \oplus x_{i-3} \oplus x_{i-5}$$

and a square wave with S(1)=0. Generate an inverse M-sequence.

saq164.m

1.7 Experiment Design of System Identification

■ Input signal design

When designing the physical experiment of dynamic system identification, there are several problems that need to be clarified in advance:

- Which signals are the input of the system and which signals are the output of the system;
- Which signals need to be collected by sensors and which signals need to be stimulated;
- When to measure. Generally, a fixed sampling period T is used.

The selection of excitation signal u may have a great impact on the identification results.

The duration of the experiment depends on how many sets of input-output signals are needed, that is, N.

Among those factors, the most critical and difficult to determine is the choice of the input signal:

- For different systems, users have different degrees of freedom for selecting input signals;
- For process control, if in a production mode, users may not be able to specify input signals arbitrarily, because they may interfere with normal production;
- For the economic system or ecosystem, it may not be possible to make any change to the input of the system, and their operation is not based on human will;
- On the experimental device in the laboratory, there are relatively large degrees of freedom, and the system input can be designed arbitrarily, as long as it meets the requirements of maximum power.

♦ **Input signals of identification experiments**

For the input signal selection of an open-loop system identification experiment, there are three basic criteria:

- The asymptotic characteristic of identification is only reflected in the frequency spectrum of the input signal, not the waveform of the input signal;
- The amplitude of an input signal must be limited to a certain range, *i.e.*,

$$u_{min} \leq u(t) \leq u_{ma}$$

where u_{max} is the upper bound and u_{min} is the lower bound.

- Periodic signals may have certain advantages in identification.

♦ **Amplitude factor**

For system identification, the stronger the power of the input signal is, the better. However, in practical applications, the amplitude of the input signal is limited.

An amplitude factor is introduced

$$C_r^2 = \frac{\max\limits_{t} u^2(k)}{\lim\limits_{N\to\infty} \frac{1}{N}\sum_{k=1}^{N} u^2(k)}$$

The amplitude factor represents the ratio of the maximum amplitude to the average amplitude;

In the best case, $C_r = 1$.

A symmetrical binary signal conforms to $C_r = 1$ if $u(k) = \pm\bar{u}$.

♦ **Filtered Gaussian white noise**

A simple choice of input signals is Gaussian white noise passing through a filter.

Theoretically, signals in any spectrum range can be obtained by selecting appropriate filters.

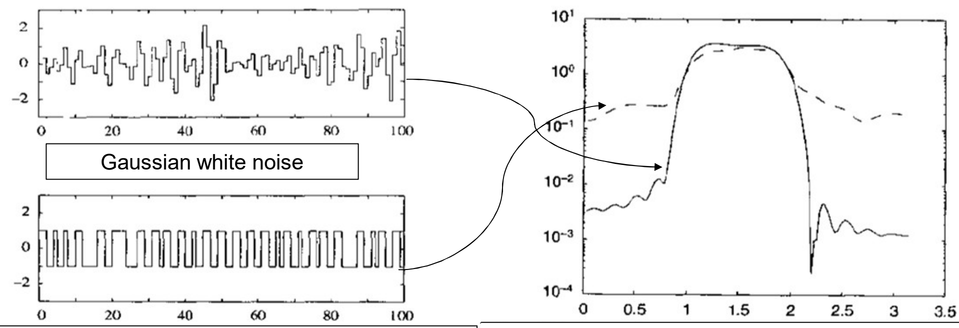
Gaussian white noise has no boundary in theory. So, the white noise must be limited to a certain range.

♦ **Random binary signal**

A random binary signal is a random signal sequence with only two values.

There are many ways to generate this signal.

The most common way is to generate a Gaussian white noise signal, pass it through a filter, and then take the positive and negative sign of the signal.



♦ **Pseudo-random binary signal**

PRBS (pseudo-random number sequence) is a group of periodic and deterministic data sequences with characteristics similar to a white noise.

It can be generated using the following equation

$$u(t) = \mathrm{rem}(A(q^{-1})u(t), 2) = \mathrm{rem}(a_1 u(t-1) + \cdots + a_n u(t-n), 2)$$

where $\mathrm{rem}(x, 2)$ represents the remainder of x divided by 2 and

$$A(q^{-1}) = a_1 q^{-1} + a_2 q^{-2} + \cdots + a_n q^{-n}$$

Because the past input vector $[u(t-1), \cdots, u(t-n)]$ has only 2^n different values, if the input has n consecutive 0, it will only result in the subsequent $u(t)$ being 0.

The longest period of the generated pseudorandom signal can only be $M=2^n-1$, and the specific period is determined by the value of $A(q^{-1})$.

♦ **Mixed sine signal**

The form of mixed sine signals is a combination of a series of sine waves

$$u(t) = \sum_{k=1}^{d} a_k \sin(\omega_k t + \varphi_k)$$

where ω_k is the k-th frequency, φ_k is the k-th phase and a_k is the coefficient.

This can be seen as a combination of multiple sine waves with different frequencies and phases.

57

♦ **Chirp signals**

A chirp signal is a sine wave signal whose frequency changes continuously with time in a range. The frequency changes from ω_1 to ω_2.

The form of a chirp signal is as follows:

$$u(t) = Asin(\omega_1 t + (\omega_2 - \omega_1)t^2/(2M))$$

From this formula, the "instant frequency" of the signal is

$$\omega = \omega_1 + \frac{t}{2M}(\omega_2 - \omega_1)$$

58

Self Assessment Question

Which of the following statements is true?

A. The input signal of dynamic system identification can accurately be designed according to the identification requirements.

B. Pseudo-random number sequence is a white noise.

C. Mixed sine signal and pseudorandom number sequence can produce the same identification results.

D. The system input signal plays a very important role in system identification.

■ Method of identifying a closed-loop system

The closed-loop system identification method considers three situations:

- It is assumed that the feedback system is unknown and the reference input signal cannot be used.
- It is assumed that both the reference input of the system and the feedback controller of the system are known.
- It is assumed that the feedback controller of the system is unknown, but the reference input of the system can be used.

Based on these three assumptions, there are three identification methods:

Direct identification method: use the input u(t) and output y(t) of the system to identify the mathematical model of the system using a method similar to the open-loop identification.

Indirect identification method: use the system reference input r(t) and output y(t) to identify the mathematical model of the closed-loop system, and then use the information of the closed-loop feedback controller to calculate the open-loop mathematical model of the system.

Joint input and output method: assume that the output y(t) and input u(t) of the system are driven by the reference input r(t). The mathematical models of the system and the feedback controller are identified simultaneously.

61

♦ **Direct identification method**

Direct identification is the most natural closed-loop system identification method, which is selected for the following reasons:

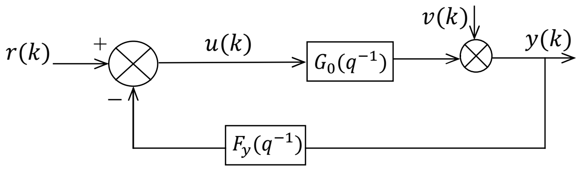
- It does not need to know any information about the feedback controller, that is, the information of the feedback controller is independent of identification.
- The open-loop identification method can be used directly without special algorithm.
- It can be used to identify unstable systems as long as the closed-loop system is stable.

The only direct disadvantage is that a more accurate disturbance model needs to be obtained.

62

♦ **Indirect identification method**

For a closed-loop system, there is the following block diagram



The transfer function of its closed-loop system is

$$\begin{aligned} y(k) &= G_{cl}(q^{-1})r(k) + v_{cl}(k) \\ &= \frac{G_0(q^{-1})}{1 + F_y(q^{-1})G_0(q^{-1})} r(k) + \frac{1}{1 + F_y(q^{-1})G_0(q^{-1})} v(k) \end{aligned}$$

63

The data of y(t) and r(t) can be used to identify G_{cl}, and then the model G_0 of the open-loop system can be identified from the following equation

$$G_{cl}(q^{-1}) = \frac{G_0(q^{-1})}{1 + F_y(q^{-1})G_0(q^{-1})}$$

The advantage of indirect identification is that any identification method can be used to identify G_{cl}.

But, the disadvantage is that any error related to the controller $F_y(q^{-1})$, including some nonlinear characteristics in $F_y(q^{-1})$, will be brought into the identification of G_0 .

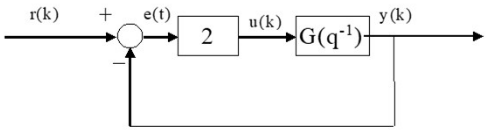
64

Self Assessment Question

The closed-loop input-output relationship is obtained using the closed-loop identification method:

$$2y(k) + 4y(k+1) + 3r(k-1) = 0$$

If the closed-loop system is as shown in the figure below, determine its open-loop transfer function $G(q^{-1})$.



■ Data pre-processing

When the data are collected through the experimental method, it may not be directly applied to the identification work because of the following reasons:

- There may be some high-frequency interference in the collected data, and the frequency of these interference is significantly higher than the dynamic frequency of the system.
- Obvious wrong data and data loss, or data interruption.
- Data may be mixed with drift or DC component, or low-frequency interference, with periodic characteristics.

For offline identification, before identification, it is necessary to draw the trend chart of the data, carefully check the possible defects of these data, and carry out some preprocessing.

♦ **DC component**

For a standard linear model, describe the relationship between its input $u(k)$ and output $y(k)$ with the following equation.

$$A(q^{-1})y(k) = B(q^{-1})u(k) + v(k)$$

This description not only includes the dynamic characteristics between $u(k)$ and $y(k)$, that is, how $y(k)$ changes with the change of $u(k)$.

The steady-state characteristics between them are also described.

If there is a fixed input $u(k)$, there is a fixed $y(k)$ in the steady state, that is:

$$u(k) \equiv \bar{u}$$
$$y(k) \equiv \bar{y}$$
$$A(1)\bar{y} = B(1)\bar{u}$$

In a practical system, the original data $u^m(t)$ and $y^m(t)$ obtained from sensors and actuators may not fully reflect the static characteristics of the system, because there are always some offsets in it.

To solve this problem, there are the following solutions (four methods):

- Obtain y(t) and u(t) from the experiment, and then remove average
- Identify the DC component as a parameter
- Use differential data
- Use a high-pass filter

♦ **Obtain y(t) and u(t) from the experiment and remove the average**

The most direct way is to do such an experiment and apply a fixed input $\bar{u}$ to the system, Then obtain the output $\bar{y}$ of the system under steady-state conditions;

This steady state is close to the working point of the system;

The data can be preprocessed in this way

$$y(k) = y^m(k) - \bar{y}$$
$$u(k) = u^m(k) - \bar{u}$$

69

Calculation of average terms

Another way is to remove the average term in the data.

In this way, obtain $\bar{u}$ and $\bar{y}$

$$\bar{y} = \frac{1}{N}\sum_{t=1}^{N} y^m\,(k)$$

$$\bar{u} = \frac{1}{N}\sum_{t=1}^{N} u^m\,(k)$$

70

♦ **Identify the DC component as a parameter**

Add a fixed parameter to the original mathematical model α to process DC components.

When identifying the system model, consider α as a parameter of θ to identify, *i.e.*,

$$A(q^{-1})y^m(k) = B(q^{-1})u^m(k) + \alpha + v(k)$$

This method has no essential difference from the previous one.

71

♦ **Use differential data**

For input and output data, a differential filter $L(q^{-1})$ is used for processing

$$L(q^{-1}) = 1 - q^{-1}$$
$$y_F^m(k) = L(q^{-1})y^m(k) = y^m(k) - y^m(k-1)$$
$$u_F^m(k) = L(q^{-1})u^m(k) = u^m(k) - u^m(k-1)$$

By using the difference method, the DC component can be removed.

♦ **Use high-pass filter**

Any high-pass filter with a DC gain of 0 (or close to 0) can be used to remove the DC component in the signal.

72

♦ **Comparison of methods**

For any off-line identification, the first method to remove the DC component is the preferred method.

If the steady state experiment cannot be carried out, the calculation of the average terms will be used.

Compared with the first method, the second method that uses the identification method to remove the DC component is a more complex and time-consuming method.

Differential data has a strong amplification effect on high-frequency noises. So, it may not be used in some cases.

73

Self Assessment Question

Which one below is true?

A. In system identification, all high-frequency signals in the collected signals must be filtered out;

B. The DC component in a collected signal can not be removed;

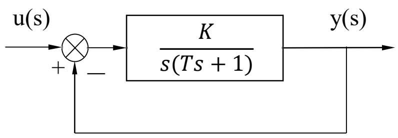
C. The differential filter has little effect on a noise;

D. The high-pass filter has no effect on a low-frequency noise.

74

Exercise 1.3

Consider the following system with K=16 and T=0.25:



If an M-sequence is used as the input u(t), determine how many shift registers are needed to identify the parameters K and T of the system.