Introduction to Digital Signal and System Analysis

Weiji Wang

Weiji Wang

Introduction to
Digital Signal and System Analysis

Contents

Preface

Since the 1990s, digital signals have been increasingly used not only in various industries and engineering equipments but also in everybody's daily necessities. Mobile phones, TV receivers, music CDs, multimedia computing, etc, are the indispensable items in modern life, in which digital formats are taken as a basic form for carrying and storing information. The major reason for the advancement in the use of digital signals is the big leap forward in the popularization of microelectronics and computing technology in the past three decades. Traditional analogue broadcast is being widely upgraded to digital. A general shift from analogue to digital systems has taken place and achieved unequivocal benefits in signal quality, transmission efficiency and storage integrity. In addition, data management advantage in digital systems has provided users with a very friendly interface. A typical example is the popular pull-down manual, easy to find, make choices and more choices are made available.

As marching into the digital era, many people in different sectors are quite keen to understand why this has happened and what might be the next in this area. They hope to obtain basic principles about digital signals and associated digital systems. Instead of targeting advanced or expert level, they as beginners often hope to grasp the subject as efficient and effective as possible without undertaking impossible task under usually limited time and effort available.

This book is written for those beginners who want to gain an overview of the topic, understand the basic methods and know how to deal with basic digital signals and digital systems. No matter the incentive is from curiosity, interest or urgently acquiring needed knowledge for one's profession, this book is well suited. The output standards are equivalent to university year two which lays a good foundation for further studies or moving on to specialised topics, such as digital filters, digital communications, discrete time-frequency representation, and time-scale analysis. The required mathematics for the reader is basically at pre-university level, actually only junior high schools maths is mainly involved. The content of materials in this book has been delivered to second year engineering and IT students at university for more than 10 years. A feature in this book is that the digital signal or system is mainly treated as originally existing in digital form rather than always regarded as an approximation version of a corresponding analogue system which gives a wrong impression that digital signal is poor in accuracy, although many digital signals come from taking samples out of analogue signals. The digital signal and system stand as their own and no need to use the analogue counter part to explain how they work.

To help understanding and gaining good familiarity to the topic, it will be very helpful to do some exercises attached to each chapter, which are selected from many and rather minimal in term of work load.

Weiji Wang
University of Sussex
Brighton, England
January 2012

1   Digital Signals and Sampling

1.1   Introduction

Digital signal processing (DSP) has become a common tool for many disciplines. The topic includes the methods of dealing with digital signals and digital systems. The techniques are useful for all the branches of natural and social sciences which involve data acquisition, analysis and management, such as engineering, physics, chemistry, meteorology, information systems, financial and social services. Before the digital era, signal processing devices were dominated by analogue type. The major reason for DSP advancement and shift from analogue is the extraordinary growth and popularization of digital microelectronics and computing technology.

The reason that digital becomes a trend to replace analogue systems, apart from it is a format that microprocessors can be easily used to carry out functions, high quality data storage, transmission and sophisticated data management are the other advantages. In addition, only 0s and 1s are used to represent a digital signal, noise can easily be suppressed or removed. The quality of reproduction is high and independent of the medium used or the number of reproduction. Digital images are two dimensional digital signals, which represent another wide application of digital signals. Digital machine vision, photographing and videoing are already widely used in various areas.

In the field of signal processing, a signal is defined as a quantity which carries information. An analogue signal is a signal represented by a continuous varying quantity. A digital signal is a signal represented by a sequence of discrete values of a quantity. The digital signal is the only form for which the modern microprocessor can take and exercise its powerful functions. Examples of digital signals which are in common use include digital sound and imaging, digital television, digital communications, audio and video devices.

To process a signal is to make numerical manipulation for signal samples. The objective of processing a signal can be to detect the trend, to extract a wanted signal from a mixture of various signal components including unwanted noise, to look at the patterns present in a signal for understanding underlying physical processes in the real world. To analyse a digital system is to find out the relationship between input and output, or to design a processor with pre-defined functions, such as filtering and amplifying under applied certain frequency range requirements. A digital signal or a digital system can be analysed in time domain, frequency domain or complex domain, etc.

1.2   Signal representation and processing

Representation of digital signals can be specific or generic. A digital signal is refereed to a series of numerical numbers, such as:

…, 2, 4, 6, 8, …

where 2, 4, 6 are samples and the whole set of samples is called a signal. In a generic form, a digital signal can be represented as time-equally spaced data

$$..., \quad x[-1] \quad x[0] \quad x[1] \quad x[2] \quad ...$$

where -1, 0, 1, 2 etc are the sample numbers, x[0], x[1], x[2], etc are samples. The square brackets represent the digital form. The signal can be represented as a compact form

$$x[n] \qquad\qquad -\infty < n < \infty$$

(1.1)

In the signal, x[-1], x[1], x[100], etc, are the samples, n is the sample number. The values of a digital signal are only being defined at the sample number variable n, which indicates the occurrence order of samples and may be given a specific unit of time, such as second, hour, year or even century, in specific applications.

We can have many digital signal examples:

- Midday temperature at Brighton city, measured on successive days,
- Daily share price,
- Monthly cost in telephone bills,
- Student number enrolled on a course,
- Numbers of vehicles passing a bridge, etc.

Examples of digital signal processing can be given in the following:

Example 1.1 To obtain a past 7 day's average temperature sequence. The averaged temperature sequence for past 7 days is

$$y[n] = \frac{1}{7}\left(x[n] + x[n-1] + x[n-2]... + x[n-6]\right).$$

For example, if n=0 represents today, the past 7 days average is

$$y[0] = \frac{1}{7}\left(x[0] + x[-1] + x[-2]... + x[-6]\right)$$

where $x[0]$, $x[-1]$, $x[-2]$, ... represent the temperatures of today, yesterday, the day before yesterday, …; $y[0]$ represents the average of past 7 days temperature from today and including today. On the other hand,

$$y[1] = \frac{1}{7}\left(x[1] + x[0] + x[-1] + ... + x[-5]\right)$$

represents the average of past 7 days temperature observed from tomorrow and including tomorrow, and so on. In a shorter form, the new sequence of averaged temperature can be written as

9

$$y[n] = \frac{1}{7}\sum_{k=1}^{7} x[n-k+1]$$

where x[n] is the temperature sequence signal and y[n] is the new averaged temperature sequence. The purpose of average can be used to indicate the trend. The averaging acts as a low-pass filter, in which fast fluctuations have been removed as a result. Therefore, the sequence y[n] will be smoother than x[n].

Example 1.2. To obtain the past M day simple moving averages of share prices, let x[n] denotes the close price, $y_M[n]$ the averaged close price over past M days.

$$y_M[n] = \frac{1}{M}\left(x[n] + x[n-1] + x[n-2]... + x[n-M+1]\right)$$

or

$$y_M[n] = \frac{1}{M}\sum_{k=1}^{M} x[n-k+1]$$

(1.2)

For example, M=20 day simple moving average is used to indicate 20 day trend of a share price. M=5, 120, 250 (trading days) are usually used for indicating 1 week, half year and one year trends, respectively. Figure 1.1 shows a share's prices with moving averages of different trading days.


Figure 1.1 S share prices with moving averages

1.3   Analogue-to-digital conversion

Although some signals are originally digital, such as population data, number of vehicles and share prices, many practical signals start off in analogue form. They are continuous signals, such as human's blood pressure, temperature and heart pulses. A continuous signal can be first converted to a proportional voltage waveform by a suitable transducer, i.e. the analogue signal is generated. Then, for adapting digital processor, the signal has to be converted into digital form by taking samples. Those samples are usually equally spaced in time for easy processing and interpretation. Figure 1.2 shows a analogue signal and its digital signal by sampling with equal time intervals. The upper is the analogue signal x(t) and the lower is the digital signal sampled at time t = nT, where n is the sample number and T is the sampling interval. Therefore,

$$x[n] = x(nT)$$
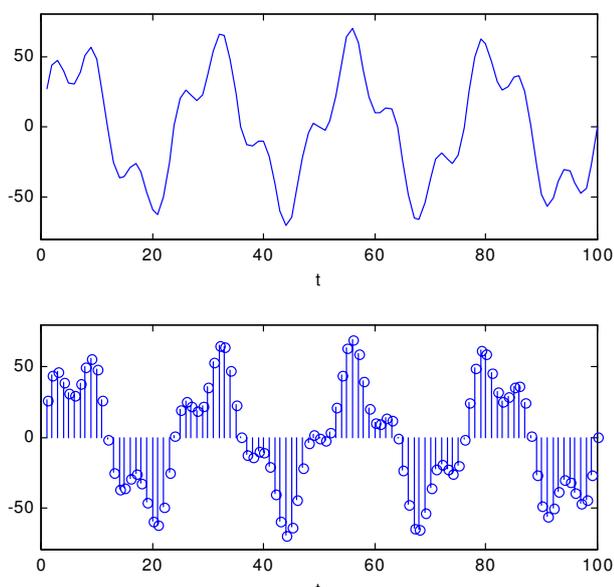
Figure 1.2 An analogue signal x(t) and digital signal x[n]. The upper is the analogue signal and the lower is the digital signal sampled at t = nT.

1.4   Sampling theorem

For ease of storage or digital processing, an analogue signal must be sampled into a digital signal. The continuous signal is being taken sample at equal time interval and represented by a set of members. First of all, a major question about it is how often should an analogue signal be sampled, or how frequent the sampling can be enough to represent the details of the original signal. It is obvious that too often will cause redundancy which will reduce the processing efficiency and cause an unnecessarily large size of data storage, but too sparse will cause a loss of signal details.

- **Shannon's sampling theorem**

Claude E. Shannon 1916-1949) established the sampling theorem that an analogue signal containing components up to maximum frequency f_c Hz may be completely represented by samples, provided that the sampling rate f_s is at least $2f_c$ (i.e. at least 2 samples are to present per period). That is

$$f_s \geq 2f_c$$

(1.3)

Let the sampling interval $T = \dfrac{1}{f_s}$, the sampling requirement is equivalently represented as

$$T \leq \frac{1}{2f_c}$$

(1.4)

Given sampling frequency f_s , the maximum analogue frequency allowed in the signal is

$$f_c = \frac{1}{2}f_s = \frac{1}{2T}$$

(1.5)

Under sampling will cause aliasing. That is, details of original signal will be lost and high frequency waveforms may be mistakenly represented as low frequency ones by the sampled digital signal. See Figure 1.3. It is worth noting that use of minimum sampling frequency is not absolutely safe, as those samples may just been placed at all zeros-crossing points of the waveform.

Example 1.3 An analogue signal is given as

$$x(t) = \sin 3000t + 2\cos 350t + \sin 200t \cos 20t$$

where *t* is the time in seconds, determine the required minimum sampling frequency for the signal and calculate the time interval between any two adjacent samples.

Solution:

The third term is equivalent to 2 components of frequencies 200+20 hz and 200-20 hz. The highest frequency in the signal therefore is $3000/2\pi = 477.5 hz$. Required minimum sampling frequency is $2 \times 477.5 hz = 955 hz$, or the sampling interval T is $1/955 = 0.001047 \text{seconds}$.
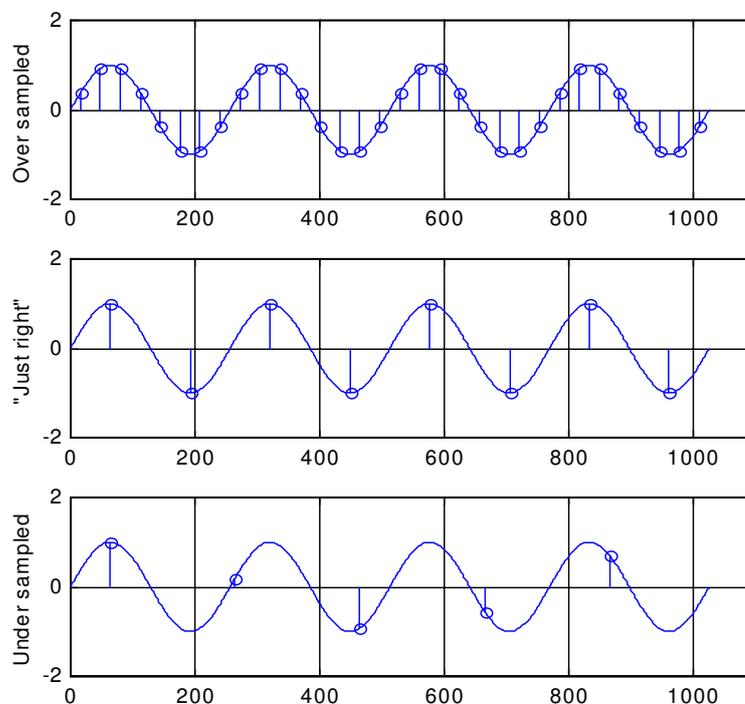
Figure 1.3 Over sampling and under sampling

1.5   Quantization in an analogue-to-digital converter

The quality of a digital signal is dependent on the quality of the conversion processes. An analogue signal takes on a continuous range of amplitudes. However, a practical electronic analogue-to-digital converter has limited levels of quantization. An n-bit analogue-to-digital converter has 2^n levels, i.e. only as many as 2^n different values can be presented in the sampling.

- **Quantization error**

During an unlimited level of analogue signal being converted into a limited level of digital signal, all possible values have to be rounded to those limited 2^n levels. This means a quantization error (or equivalently termed as quantization noise) has been introduced. In practice, n in the 2^n needs to be chosen to be big enough to satisfy the quantization accuracy. When n=3, 2^n =8 provides 8 quantization levels. Obviously, there exists big quantization errors in representing the original continuous signal by a small number of levels. But when taking *n=12*, it gives as many as 4096 quantization levels, which satisfies many industrial applications.

The following Figure 1.4 illustrates the quantization process in which the analogue to digital convertor has 8 levels. A continuous signal is sampled to a digital signal as …,1, 6, 6, 5, 5, 4, 4, 4, 4, 6,… which have difference, i.e. the error, at each sampling point between the analogue and digital values.
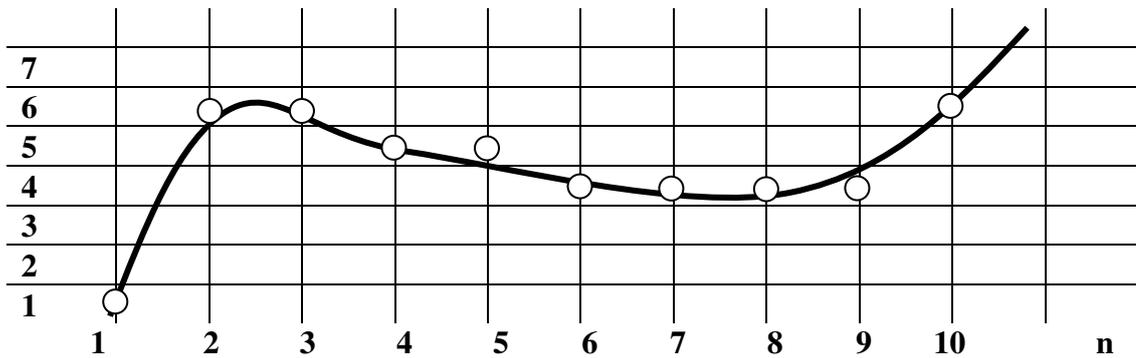

Figure 1.4 Continuous signal is sampled as 8 levels of digital signal.

Problems

Q1.1 Observe the signals in Fig. Q1.1 and answer the following questions:

   a) What is the frequency of the analogue sinusoidal signal (solid line) ? _____.
   b) How many samples have been taken from the analogue signal within one second (the sampling frequency)? _____.
   c) Does the digital signal (the dotted line) represent the original analogue signal correctly? _____ . What has happened? _____.
   d) What is the frequency of the digital sinusoidal signal? _____.
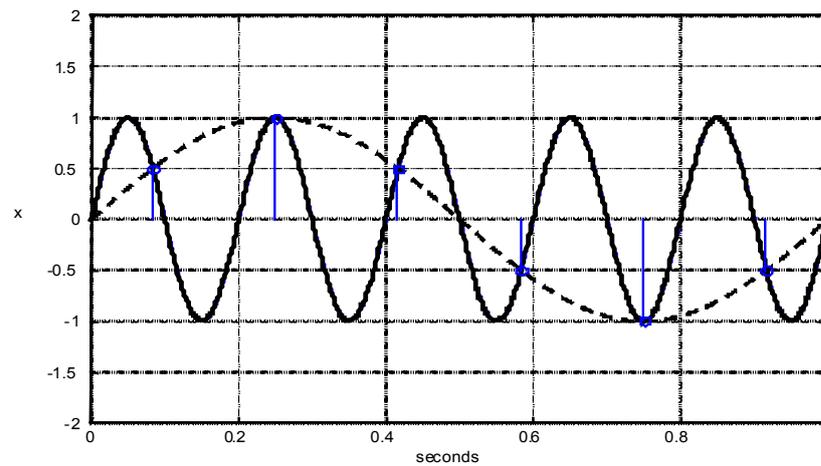   e) What should be a required minimum sampling frequency for the original analogue signal? _____.

Figure Q1.1

Q1.2 A stock price can be described by a digital series $x[n]$, where n indicates the day: $-\infty<n<\infty$. The past M day (including the present day) simple moving average (SMA) series is often used for trend analysis.

Write down the M day SMA for series $y_M[n]$ for indicating price trend. Also, when the 5 day SMA is crossing its 50 day SMA, the stock may be taken as changed in a medium term trend. Write down an expression to be used to alert this change in the trend.

2   Basic Types of Digital Signals

2.1   Three basic signals

Unit impulse

$$\delta[n] = \begin{cases} 1 & n = 0 \\ 0 & n \neq 0 \end{cases}$$

(2.1)

i.e. the unit impulse has only one non-zero value 1 at n=0, and all other samples are 0. It is the simplest signal but will be seen later very important.

Unit step

$$u[n] = \begin{cases} 1 & n \geq 0 \\ 0 & n < 0 \end{cases}$$

(2.2)

where the sample value rises at n=0 from 0 to 1 and keeps it to $n \rightarrow \infty$.

Ramp

$$r[n] = \begin{cases} n & n \geq 0 \\ 0 & n < 0 \end{cases}$$

(2.3)

Figure 2.1 illustrates the unit impulse, unit step and ramp signals. Alternatively, the three basic signals can be expressed by a tabular form as below:

n	…	-3	-2	-1	0	1	2	3	…
$\delta[n]$	…	0	0	0	1	0	0	0	…
u[n]	…	0	0	0	1	1	1	1	…
r[n]	…	0	0	0	0	1	2	3	…

The unit impulse, unit step and ramp signals are simple and basic signals. Unit impulse is the basic of basic signals. The unit step can be regarded as a sum of unit impulse on the positive side of n-axis. The ramp signal can be a set of unit impulses shifted to the right axis and scaled by sample number *n*. Or it can be regarded as the unit step scaled by the corresponding sample number n. Actually, later, we will know all signals can be regarded as a sum of shifted and scaled unit impulses.
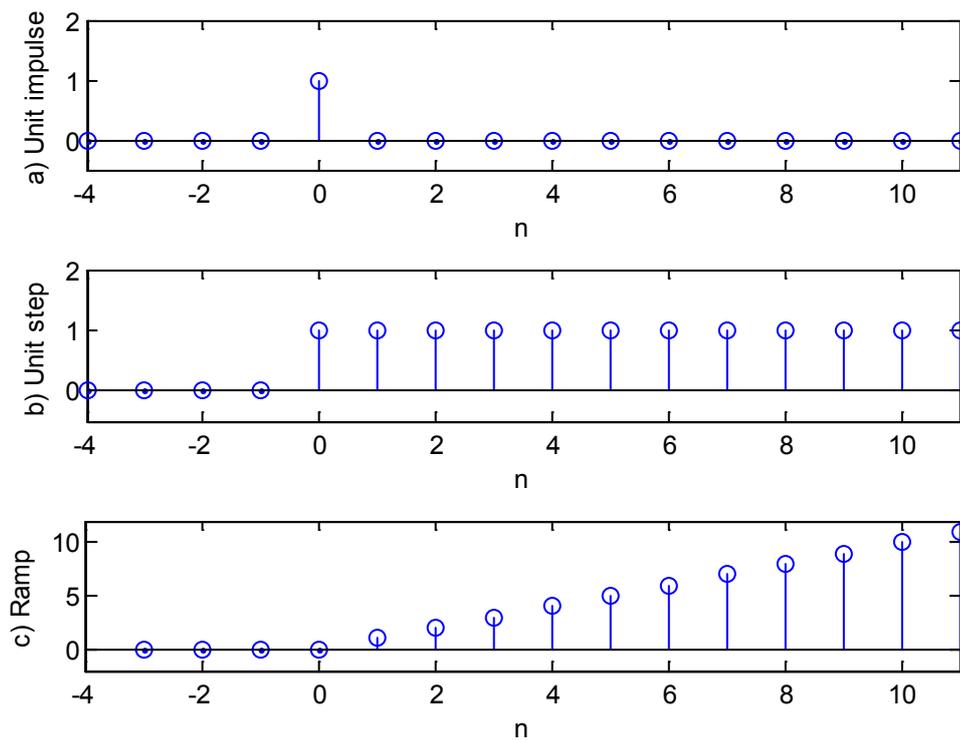
Figure 2.1 Unit impulse, unit step and ramp signals

2.2   Other basic signals

- Sinusoidal signals

Sinusoidal signals are referred to the sine and cosine functions. In digital format, they are

$$x[n] = A\cos(n\Omega)$$
$$x[n] = A\sin(n\Omega)$$

(2.4)

where, it is worth noting, Ω is the frequency with a unit of radians/sample, n is the sample number. The sinusoidal functions have a period of 2π.

- Exponential signal

$$x[n] = Ae^{\beta n}$$

or

$$x[n] = A\exp(\beta n)$$

(2.5)

where $A,$ β are constants and n the sample number.

- Complex signals z

A digital signal x[n] can be complex, a simple example is

$$x[n] = \cos(n\Omega) + j\sin(n\Omega)$$

where $j = \sqrt{-1}$, and the signal has real and imaginary parts. Note that according to Euler formula $\exp(jn\Omega) = \cos(n\Omega) + j\sin(n\Omega),$, there are relationships:

$$\cos(n\Omega) = \frac{1}{2}\{\exp(jn\,\Omega) + \exp(-jn\,\Omega)\}$$

$$\sin(n\Omega) = \frac{1}{2j}\{\exp(jn\,\Omega) - \exp(-jn\,\Omega)\}$$

$$(2.6)$$

2.3   Signal shifting, flipping and scaling

- Signal shifting

A signal can be shifted to left or right by any number of samples. In Figure 2.2 (a), a unit impulse has been shifted to the right by one sample; and in Figure 2.2 (b), a unit impulse $\delta[n]$ has been shifted to the left by one sample. They should be represented by $\delta[n-1]$ (shifted to the right) and $\delta[n+1]$ (shifted to the left), respectively. For the general case of a signal $x[n]$, shifting to the right and left by n_0 samples generates new signals $x[n-n_0]$ and $x[n+n_0]$. They are a delayed signal and an advanced signal, respectively.

(a)   (b)
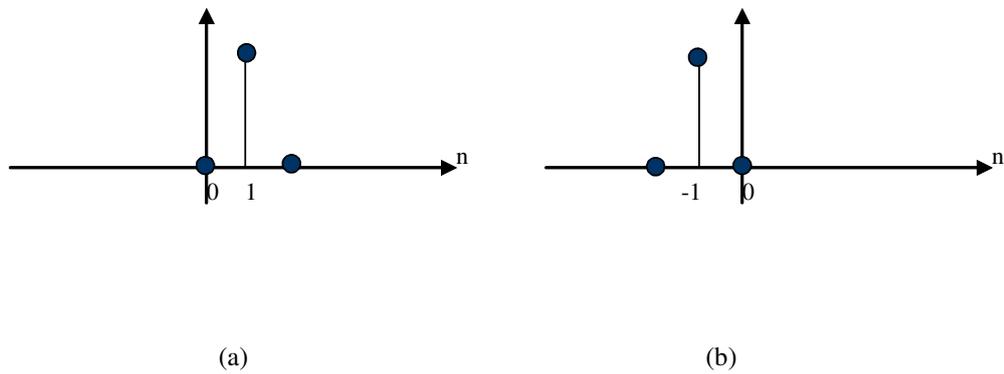
Figure 2.2 Unit impulse is shifted to left: d[n+1] (a) and Shifted to right: d[n-1] (b)

- Signal flipping

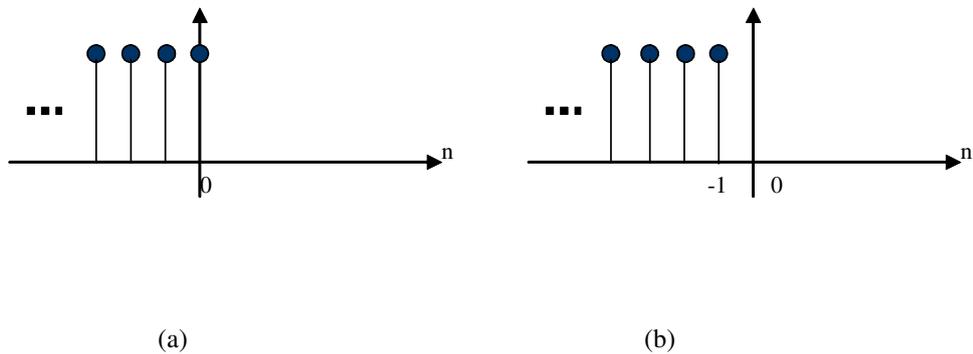A unit step $u[n]$ flipped about y-axis can generate a new signal $u[-n]$ shown in Figure 2.3(a). It can also be flipped and then shifted to left by one sample as $u[-(n+1)]$ or $u[-n-1]$ as shown in Figure 2.3(b). In general cases, a flipped signal of $x[n]$ is $x[-n]$.

(a)   (b)

Figure2.3 Unit step is flipped as u[-n] (a) and shifted to the left: u[-n-1] = u[-(n+1)] (b)

- Signal scaling

In Figure 2.4, a shifted unit impulse $\delta[n-1]$ is been scaled by -2 to $-2\,\delta[n-1]$. For general cases, a signal $x[n]$ scaled by number a is $ax[n]$.
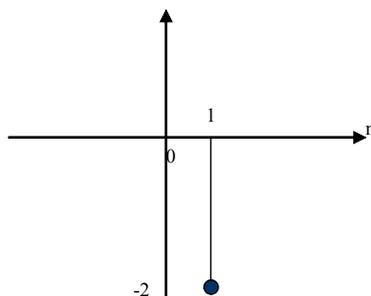

F**igure 2.4** A shifted unit impulse is being scaled by -2.

2.4   Periodic signals

A periodic signal satisfies the following relationship

$$x[n \pm kN] = x[n]$$

(2.7)

where k is an arbitrary integer and N the period. The above relationship indicates that a periodic signal can remain the same shape if it shifts to left or right by any integer number of periods. Typical periodic signals are sine and cosine waves.

e.g. For the signal $x[n] = \sin\left(\dfrac{n\pi}{1}\right)$, we can find the period by following steps:

We know that the sine function has a period of 2π. Therefore,

$$\sin\left(\frac{n\pi}{1}\right) = \sin\left(\frac{n\pi}{1} \pm 2\pi\right) = \sin\left(\frac{(n \pm 2\ )\pi}{1}\right)$$

This means that on the n-axis, a new signal after being shifted to left or right by 22 samples is still identical to the original signal. Therefore, *N=22* (samples) is the period.

2.5   Examples of signal operations

For 6 signals in Figure 2.5, the expressions using basic signals, including the unit impulse, unit step and ramp, can be found as

a) x[n]=-2 u[n]   *b) x[n]=-5 u[-n-4]*
c) x[n]= u[n+3] - u[n-5]   *d) x[n]= 5 d[n-6]*
e) x[n]= d[n-6]-u[-n]   *f) x[n]= 2 r[n+6] - 2 r[n+2]*

In b), the signal has been flipped, scaled by -5 and shifted to the left by 4 samples $-5u[-n-4] = -5u[-(n+4)]$ . . c) is an rectangular function or a window function. In f) the gradient has been changed by scaling factor 2.
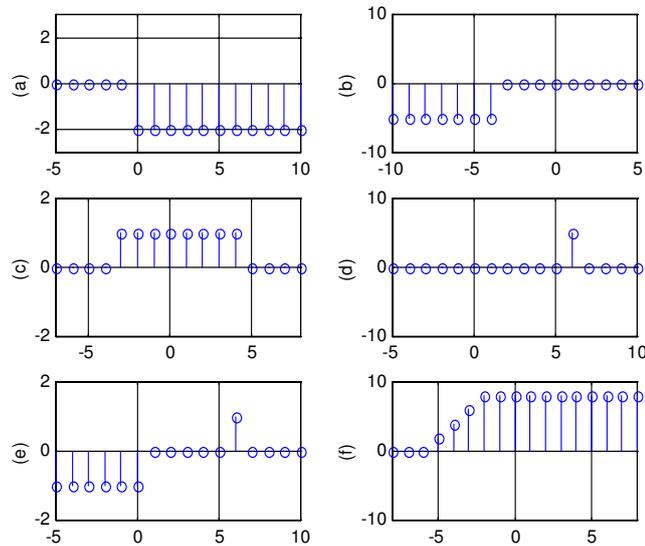

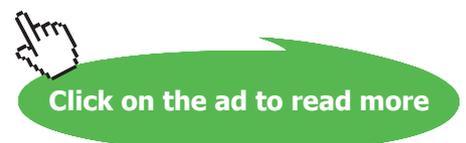**Figure 2.5** Examples of signal operations

The unit step consists of infinite number of unit impulses on the positive side of axis. The following are the representations between the unit impulse and unit step. The unit step is represented by unit impulses as

$$u[n] = \delta[n] + \delta[n-1] + \delta[n-2] + \ldots = \sum_{k=0}^{\infty} \delta[n-k] = \sum_{m=-\infty}^{n} \delta[m]$$

(2.8)

And, the unit impulse can be represented by the unit steps as

$$u[n] - u[n-1] = \sum_{m=-\infty}^{n} \delta[m] - \sum_{m=-\infty}^{n-1} \delta[m] = \delta[n] + \sum_{m=-\infty}^{n-1} \delta[m] - \sum_{m=-\infty}^{n-1} \delta[m] = \delta[n]$$

i.e. the impulse response is the difference of two unit steps.

$$\delta[n] = u[n] - u[n-1]$$

(2.9)

Problems

Q2.1   Sketch and label the following digital signals

    a) $-u[n+1]$

    b) $u[-n+1]$

    c) $u[n+2] + 2\delta[n-3]$

    d) $3u[n+4] - 3u[n-5]$

    e) $r[n+1] - 2r[n-2]$

Q2.2 Sketch and label the digital signal

    a) $x[n] = \dfrac{1}{n} u[n] \sin \dfrac{\pi\, n}{2}$

    b) $x[n] = u[n+2] - u[n-3] + r[n-3] - r[n-6] + 2\delta[n]$

where $\delta[n]$, $u[n]$ and $r[n]$ are the unit impulse, unit step and ramp functions, respectively.

Q2.3 Let $\delta[n]$, $u[n]$ and $r[n]$ be the unit impulse, unit step and ramp functions, respectively. Given

$$x_1[n] = u[n] - u[n-6]$$
$$x_2[n] = r[n] + \delta[n] + \delta[n-2] - \delta[n-4] - r[n-5]$$

Sketch and label the digital signals $x_1[n] + x_2[n]$ and $x_1[n] \cdot x_2[n]$.

Q2.4 Let $\delta[n]$ $u[n]$ and $r[n]$ be the unit impulse, unit step and ramp functions, respectively. Given

$$x_1[n] = u[n-2] - u[n-4]$$
$$x_2[n] = r[n-2] - \delta[n]$$

Sketch and label the digital signals $x_1[n] + x_2[n]$ and $x_1[n] \cdot x_2[n]$.

Q2.5 Find the period of the following digital signal:

(a) $x[n] = \sin\dfrac{\pi\, n}{11}$

(b) $x[n] = 3\sin\left(\dfrac{\pi\, n}{3} + \pi\right) + \cos\left(\dfrac{2\pi\, n}{15} - \pi\right)$

(c) $x[n] = 2 + 3\sin\dfrac{\pi\, n}{3} + 4\sin\dfrac{\pi\, n}{6} + \cos\dfrac{2\pi\, n}{15}$.

Q2.6 Two pperiodic digital signals are given as

$$x_1[n] = 1 + \cos\frac{\pi\, n}{16}, \qquad x_2[n] = \sin\frac{2\pi\, n}{23}$$

Find the period of $x_1[n]$ $x_2[n]$ and $x_1[n] - x_2[n]$.

3   Time-domain Analysis

3.1   Linear time-invariant (LTI) systems

A digital system is also refereed as a digital processor, which is capable of carrying out a DSP function or operation. The digital system takes variety of forms, such as a microprocessor, a programmed general-purpose computer, a part of digital device or a piece of computing software.

Among digital systems, linear time-invariant (LTI) systems are basic and common. For those reasons, it will be restricted to address about only the LTI systems in this whole book.

The linearity is an important and realistic assumption in dealing with a large number of digital systems, which satisfies the following relationships between input and output described by Figure 3.1. i.e. a single input $x_1[n]$ produces a single output $y_1[n]$, Applying sum of inputs $x_1[n] + x_2[n]$ produces $y_1[n] + y_2[n]$, and applying input $ax_1[n] + bx_2[n]$ generates $ay_1[n] + by_2[n]$.

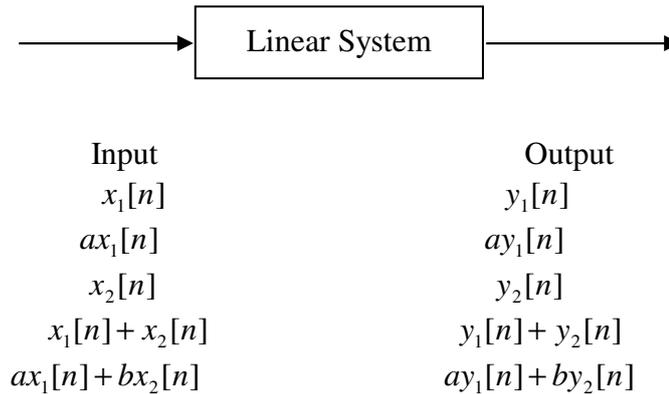| Input	Output
$x_1[n]$	$y_1[n]$
$ax_1[n]$	$ay_1[n]$
$x_2[n]$	$y_2[n]$
$x_1[n] + x_2[n]$	$y_1[n] + y_2[n]$
$ax_1[n] + bx_2[n]$	$ay_1[n] + by_2[n]$

Figure 3.1 Linearity of a system

The linearity can be described as the combination of a scaling rule and a superposition rule. The time-invariance requires the function of the system does not vary with the time. e.g. a cash register at a supermarket adds all costs of purchased items $x[n], x[n-1], \ldots$ at check-out during the period of interest, and the total cost $y[n]$ is given by

$$y[n] = x[n] + x[n-1] + x[n-2] + \ldots$$

(3.1)

where $y[n]$ is the total cost, and if $x[0]$ is an item registered at this moment, $x[-1]$ then is the item at the last moment, etc. The calculation method as a simple sum of all those item's costs is assumed to remain invariant at the supermarket, at least, for the period of interest.

3.2   Difference equations

Like a differential equation is used to describe the relationship between its input and output of a continuous system, a difference equation can be used to characterise the relationship between the input and output of a digital system. Many systems in real life can be described by a continuous form of differential equations. When a differential equation takes a discrete form, it generates a difference equation. For example, a first order differential equation is commonly a mathematical model for describing a heater's rising temperature, water level drop of a leaking tank, etc:

$$\frac{dy(t)}{dt} + ay(t) = bx(t)$$

(3.2)

where $x[n]$ is the input and $y[n]$ is the output. For digital case, the derivative can be described as

$$\frac{dy(t)}{dt} = \frac{y[n] - y[n-1]}{T} \; ,$$

(3.3)

i.e. the ratio of the difference between the current sample and one backward sample to the time interval of the two samples. Therefore, the differential equation can be approximately represented by a difference equation:

$$\frac{y[n] - y[n-1]}{T} + ay[n] = bx[n]$$

or

$$(1 + Ta)y[n] = y[n-1] + Tbx[n]$$

yielding a standard form difference equation:

$$y[n] = a_1 y[n-1] + b_1 x[n] \tag{3.4}$$

where $a_1 = \dfrac{1}{1+Ta}$ and $b_1 = \dfrac{Tb}{1+Ta}$ are constants.

For input's derivative, we have similar digital form as

$$\frac{dx(t)}{dt} = \frac{x[n] - x[n-1]}{T}.$$

Further, the second order derivative in a differential equation contains can be discretised as

$$\frac{d^2 y(t)}{dt^2} = \frac{\dfrac{y[n]-y[n-1]}{T} - \dfrac{y[n-1]-y[n-2]}{T}}{T} = \frac{1}{T^2}\left(y[n] - 2y[n-1] + y[n-2]\right). \tag{3.5}$$

When the output can be expressed only by the input and shifted input, the difference equation is called non-recursive equation, such as

$$y[n] = b_1 x[n] + b_2 x[n-1] + b_3 x[n-2] \tag{3.6}$$

On the other hand, if the output is expressed by the shifted output, the difference equation is a recursive equation, such as

$$y[n] = a_1 y[n-1] + a_2 y[n-2] + a_3 y[n-3] \tag{3.7}$$

where the output $y[n]$ is expressed by it shifted signals $y[n-1]$, $y[n-2]$, etc. In general, an LTI processor can be represented as

$$y[n] = a_1 y[n-1] + a_2 y[n-2] + ... + b_1 x[n-1] + b_2 x[n-2] + ...$$

or a short form

$$y[n] = \sum_{k=1}^{N} a_k y[n-k] + \sum_{k=0}^{M} b_k x[n-k] \tag{3.8}$$

or

$$\sum_{k=0}^{N} a_k' \, y[n-k] = \sum_{k=0}^{M} b_k \, x[n-k]$$

(3.9)

A difference equation is not necessarily from the digitization of differential equation. It can originally take digital form, such as the difference equation in Eq.(3.1).

3.3   Block diagram for LTI systems

Alternatively, equivalent to the difference equation, an LTI system can also be represented by a block diagram, which also characterises the input and output relationship for the system.

For example, to draw a block diagram for the digital system described by the difference equation:

$$y[n] + 0.7y[n-1] + 0.8y[n-2] = x[n] - 0.5x[n-1] - 0.6x[n-2]$$

The output can be rewrite as

$$y[n] = -0.7y[n-1] - 0.8y[n-2] + x[n] - 0.5x[n-1] - 0.6x[n-2]$$

The block diagram for the system is shown in Figure 3.2.

In the bock diagram, *T* is the sampling interval, which acts as a delay or right-shift by one sample in time. For general cases, instead of Eq.(3.9), Eq. (3.8) is used for drawing a block diagram. It can easily begin with the input, output flows and the summation operator, then add input and output branches.
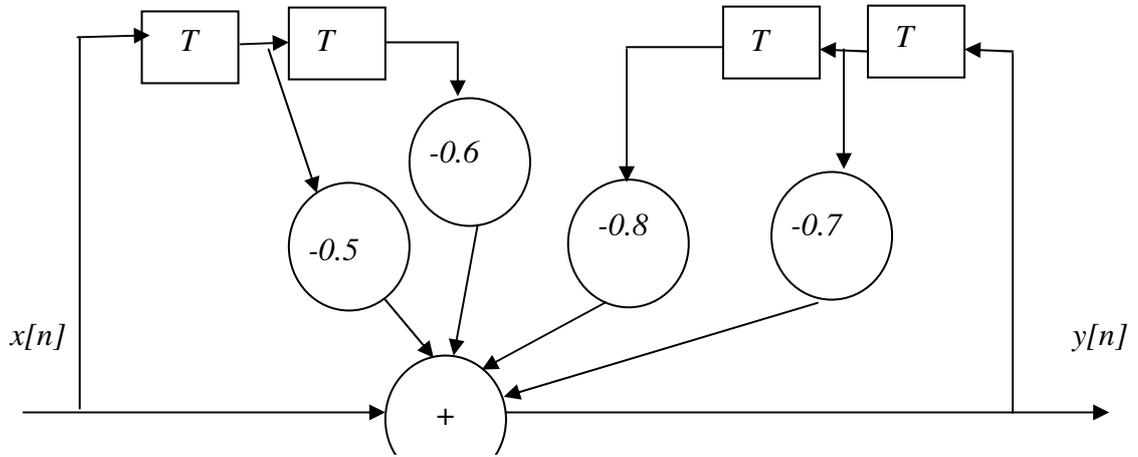

Figure 3.2 Block diagram of an LTI system

3.4   Impulse response

Both the difference equation and block diagram can be used to describe a digital system. Furthermore, the impulse response can also be used to represent the relationship between input and output of a digital system. As the terms suggest, impulse response is the response to the simplest input – unit impulse. Figure 3.2 illustrates a digital LTI system, in which the input is the unit impulse and the output is the impulse response.
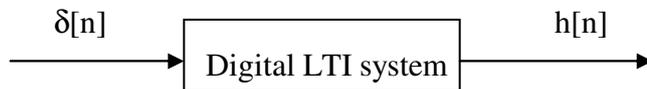

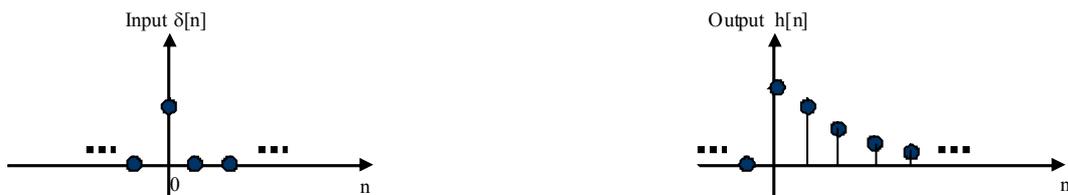**Figure 3.2** Unit impulse and impulse response

Figure 3.3 Unit impulse and impulse response of a causal system

Once the impulse response of a system is known, it can be expected that the response to other types of input can be derived.

An LTI system can be classified as causal or non-causal. A causal system is refereeing to those in which the response is no earlier than input, or h[n] =0 before n=0. This is the case for most of practical systems or the systems in the natural world. However, non-causal system can exist if the response is arranged, such as programmed, to be earlier than the excitation. See the illustration in Figure 3.4 below.

Figure 3.4 Unit impulse and impulse response of a non-causal system

The impulse response of a system can be evaluated from its difference equation. Following are the examples of finding the values of impulse responses from difference equations

Example 3.1 Evaluating the impulse response for the following systems

a) y[n]=3 x[n] + x[n-1] + 4 x[n-2]

We know that when the input is the simplest unit impulse d[n], the output response will be the impulse response. Therefore, replacing input x[n] by d[n] and response y[n] by h[n], the equation is still holding and has become special:

h[n]=3 d[n] + d[n-1] + 4 d[n-2]

It is easy to evaluate the impulse response by letting n=-1, 0,1,2,3,…

When n=-1, h[-1]=0
When n=0, h[0]= 3 d[0] +d[-1] + 4 d[-2]=3
When n=1, h[1]= 1,
When n=2, h[2]= 4,
When n=3, h[3]= 0,
…

Therefore, the impulse response h[n]=[3, 1, 4, 0, …]
↑

where ↑indicates the position of origin n=0.

b) Assume the system is causal. With the difference equation

y[n]=1.5 y[n-1] -0.85 y[n-2] + x[n]

We have

h[n]=1.5 h[n-1] -0.85 h[n-2] + δ[n]

Let n=0, 1, 2, 3,…

h[0]=0-0+1=1
h[1]=1.5 ´1-0 -0=1.5
h[2]=1.5´1.5-0.85´1+0=1.4
h[3]=1.5´1.4-0.85´1.5+0 =0.825

...

Therefore, the impulse response h[n]=[1, 1.5, 1.4, 0.825, …].
$$\uparrow$$

Generally for the difference equation:

$$y[n] = a_1 y[n-1] + a_2 y[n-2] + a_3 y[n-3] + ... + b_1 x[n-1] + b_2 x[n-2] + b_3 x[n-3] + ...$$

The impulse response can evaluated by the special equation with the simple unit impulse input:

$$h[n] = a_1 h[n-1] + a_2 h[n-2] + ... + b_1 \delta[n-1] + b_2 \delta[n-2] + ...$$
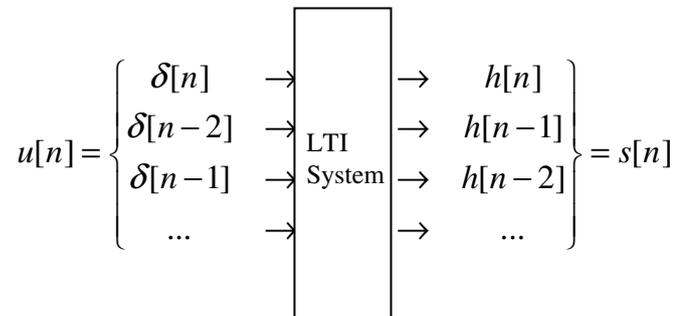
(3.10)

The step response is also commonly used to characterize the relationship between the input and output of a system. To find the step response using the impulse response, we know that the unit step can be expressed by unit impulses as

$$u[n] = \delta[n] + \delta[n-1] + \delta[n-2] + ...$$

(3.11)

The linear system satisfies the superposition rule. Therefore, the step response is a sum of a series of impulse responses excited by a series of shifted unit impulses. i.e., the step response is a sum of impulse responses

$$s[n] = h[n] + h[n-1] + h[n-2] + ... = \sum_{k=0}^{\infty} h[n-k] = \sum_{m=\infty}^{n} h[m]$$

(3.12)

To better understand Eq. (3.12), we can make use of the linearity of the LTI systems. In Figure 3.5, it has been shown that the input is decomposed in to impulses according to Eq.(3.11), and the output is the responses of all individual impulse responses described in Eq. (3.12).

$$u[n] = \begin{cases} \delta[n] & \to \\ \delta[n-2] & \to \\ \delta[n-1] & \to \\ ... & \to \end{cases} \boxed{\begin{array}{c} \text{LTI} \\ \text{System} \end{array}} \begin{array}{c} \to \\ \to \\ \to \\ \to \end{array} \left. \begin{array}{c} h[n] \\ h[n-1] \\ h[n-2] \\ ... \end{array} \right\} = s[n]$$

Figure 3.5 Multiple unit impulse inputs to an LTI system

Example 3.2: Find the step response s[n] for a system described by

y[n]=0.6y[n-1] +x[n].

Solution: From the difference equation, h[n]=0.6h[n-1]+δ[n], the samples of impulse response can be evaluated as

> h[0]=0+1=1
> h[1]=0.6´1+0=0.6
> h[2]=0.6´0.6+0=0.6´0.6
> h[3]=0.6´0.6´0.6+0=0.6´0.6´0.6
> ...

The samples of step response are

> s[0]=h[0]=1
> s[1]=h[0]+h[1]=1+0.6
> s[2]=h[0]+h[1]+h[2]=1+0.6+0.6´0.6

s[3]= h[0]+h[1]+h[2]+h[3]=1+0.6+0.6´0.6+0.6´0.6´0.6

$$s[\infty]=1+0.6+0.6´0.6+0.6´0.6´0.6+… = \frac{1}{1-0.6} = 2.5 ,$$

where the following series summation formula is applied.

$$1+a+a^2+a^3+…=\frac{1}{1-a} \qquad |a|<1$$

(3.13)

3.5   Convolution

In order to derive the convolution formula based on clear understanding, a signal is expressed by impulse functions as following:

For a signal $x[n]$ $\infty < n < \infty$, using the rules of the signal shifting and scaling described Section 2.3, it is decomposed into a series of unit impulses scaled by the sample values:

$$x[n]=…+x[-2]\delta[n+2]+x[-1]\delta[n+1]+x[0]\delta[n]+x[1]\delta[n-1]+x[2]\delta[n-2]+…$$

or

$$x[n]=\sum_{k=\infty}^{\infty}x[k]\delta[n-k]$$

(3.14)

The above expression can be illustrated as following. For example, a signal $x[n]$ has 2 non-zero samples, which can be represented as a sum of shifted and scaled unit impulses $x[1]\delta[n-1]+x[2]\delta[n-2]$, as illustrated in Figure 3.6.
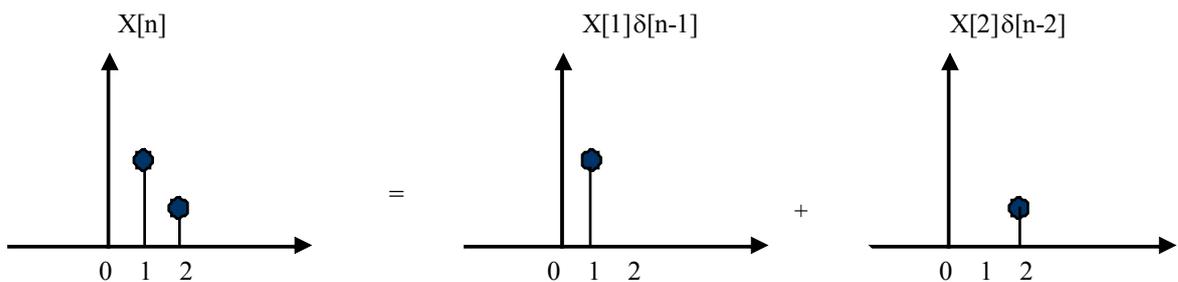

Figure 3.6 A signal can be decomposed into simple sequences.
(Except those non-zero samples, all other samples have zero values.)

In general cases, assuming the system is causal, if the input x[n] is

δ[n]: … 0 0 1 0 0 …
          ↑

and ,the impulse response is

h[n]: ... 0 0 0 h[0] h[1] h[2] ...

$\uparrow$

Then, the output y[n]is h[n].

On the other hand, if the input is

x[0]δ[n]: ... 0 x[0] 0 0   ...

The, the output y[n] is x[0]h[n]

Finally, if the input is

x[n]: ... x[-1] x[0] x[1] x[2] ...

$\uparrow$

Using the linearity again, the output y[n] will be a sum of all responses to individual shifted and scaled impulses as in Eq.(3.14), shown in Figure 3.7.

$$x[n] = \begin{cases} \cdots \\ x[-1]\delta[n+1] \\ x[0]\delta[n] \\ x[1]\delta[n-1] \\ x[2]\delta[n-2] \\ \cdots \end{cases} \boxed{\begin{array}{c} \text{LTI} \\ \text{System} \end{array}} \begin{cases} \cdots \\ \rightarrow x[-1]h[n+1] \\ \rightarrow x[0]h[n] \\ \rightarrow x[1]h[n-1] \\ \rightarrow x[2]h[n-2] \\ \cdots \end{cases} = y[n]$$
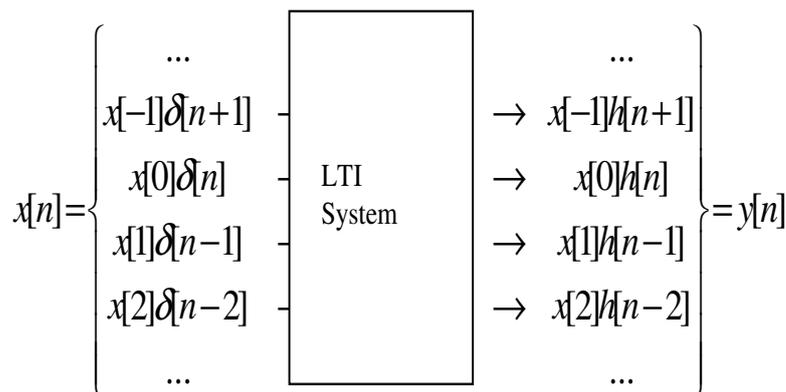
.

. **Figure 3.7** Decomposed inputs generate decomposed outputs.

i.e. the output of an LTI system will be

$$y[n] = \ldots + x[-1]h[n+1] + x[0]h[n] + x[1]h[n-1] + x[2]h[n-2] + \ldots \tag{3.15}$$

or $\qquad y[n] = \sum_{k=\infty}^{\infty} x[k]\ h[n-k] \tag{3.16}$

Eq. (3.15) or (3.16) is called the convolution sum or convolution, which describes how the input and impulse response are engaged to generate the output in an LTI system. For short, the convolution sum is also represented by

$$y[n] = x[n] * h[n] \tag{3.17}$$

where '*' represents the convolution operation. Explicitly, the samples of the response by the convolution are

$$
\begin{aligned}
y[0] &= \ldots + x[-1]\ h[1] + x[0]\ h[0] + x[1]\ h[-1] + x[2]\ h[-2] + \ldots \\
y[1] &= \ldots + x[-1]\ h[2] + x[0]\ h[1] + x[1]\ h[0] + x[2]\ h[-1] + \ldots \\
y[2] &= \ldots + x[-1]\ h[3] + x[0]\ h[2] + x[1]\ h[1] + x[2]\ h[0] + \ldots \\
&\ldots \ldots
\end{aligned} \tag{3.18}
$$

Eq. (3.18) describes the way of calculating a convolution. For manually calculating the convolution, put x[n] in normal order and put h[n] in a flipped order. x[n] and h[n] are aligned with their origin. The output sample y[0] can be calculated by a sum of multiplications between corresponding samples. Shifting h'[n] right by one sample, y[1] can also be calculated by a sum of multiplications between new corresponding samples. The following is an example.

Example 3.3 Obtain the output of a system using manual convolution:

n	...	-1	0	1	2	3	4	...
x[n]	...	0	4	5	6	7	8	...
h[n]	...	0	0.5	0.25	0.125	0.0625	0.03125	…

k	-4	-3	-2	-1	0	1	2	3	4	...
x[n]	0	0	0	0	4	5	6	7	8	...

h'[n]	0.03125	0.0625	0.125	0.25	0.5	0	0	0	0	... y[0]
		0.03125	0.0625	0.125	0.25	0.5	0	0	0	...y[1]
			0.03125	0.0625	0.125	0.25	0.5	0	0	...y[2]
				0.031	0.062	0.125	0.25	0.5	0	... y[3]

∘ ∘ ∘

The convolution satisfies the exchange rule. From Eq. (3.16), let *r=n-k*, then *k=n-r*. When *k* runs from -∞ to ∞, *r* runs from ∞ to -∞:

$$y[n] = \sum_{r=\infty}^{\infty} x[n-r] \; h[r]$$

or

$$y[n] = \sum_{k=\infty}^{\infty} h[k] \; x[n-k]$$

i.e.

$$y[n] = h[n] * x[n]$$

Therefore,

$$y[n] = x[n] * h[n] = x[n] * h[n] \tag{3.19}$$

i.e. there is no difference if *x[n]* and *h[n]* swap their places. Eq.(3.19) is applicable to any 2 signals $x_1[n]$ and $x_2[n]$:

$$x_1[n] * x_2[n] = x_2[n] * x_1[n]$$

<div align="right">(3.20)</div>

Example 3.4: A filter's difference equation is y[n]-0.5 y[n-1]=0.5 x[n], where x[n]: ...0 0 4 5 6 7 8 9 10 ...

<div align="right">↑</div>

   a) Find the impulse response of the filter and to calculate samples of response y[0], y[1], ... , y[4].

   b) Check y[n] can be found by the difference equation.

Solution:

   a) From

y[n]=0.5 y[n-1] + 0.5 x[n]

we know

h[n]=0.5 h[n-1]+ 0.5 d[n]

Then, the impulse response can be evaluated

h[0]= 0.5 ´ 1=0.5

h[1]= 0.5 ´ 0.5 =0.25

h[2]=0.5 ´ 0.25=0.125

Using the manual convolution method,

y[0]=4 ×0.5=2

y[1]=4 ×0.25+5 ×0.5=3.5

y[2]=4 ×0.125+5 ×0.25+6 ×0.5=4.75

y[3]=4 ×0.0625+5 ×0.125+6 ×0.25+7 ×0.5=5.875

y[4]=4 ×0.03125+5×0.0625+6 ×0.125+7 ×0.25+7 ×0.5=6.9375

c) Direct calculation

Causality is assumed in this case …, i.e. x[-2]=0, x[-1]=0. Therefore, it can be determined

…, y[-2], y[-1]=0,

using

y[n]=0.5 y[n-1] + 0.5 x[n]

y[-1]=0

y[0]=0.5×0+0.5×4=2

y[1]= 0.5×2+0.5×5=3.5

y[2]=0.5×3.5+0.5×6=4.75

y[3]=0.5×4.75+0.5×7=5.875

y[4]=0.5×5.875+0.5×8=6.9375

3.6   Graphically demonstrated convolution

The following Figure 3.8 illustrates how the convolution between the input and impulse response is carried out.

a) and b) are the unit impulse and impulse response, respectively.

c) is the input 1 : x[-1]d[n+1]. The response is in d) : h[n+1].

e) and f) are the input 2 : x[0]]d[n] and response 2 :x[0]h[n].

g) and h) are the input 3 : x[1]]d[n-1] and response 3 : x[1]h[n-1].

i) is the total input x[n]=… x[-1]d[n+1]+ x[0]]d[n]+ x[1]]d[n-1]+…

j) is the total response y[n]= … x[-1]h[n+1]+ x[0]]h[n]+ x[1]]h[n-1]+…

i.e.
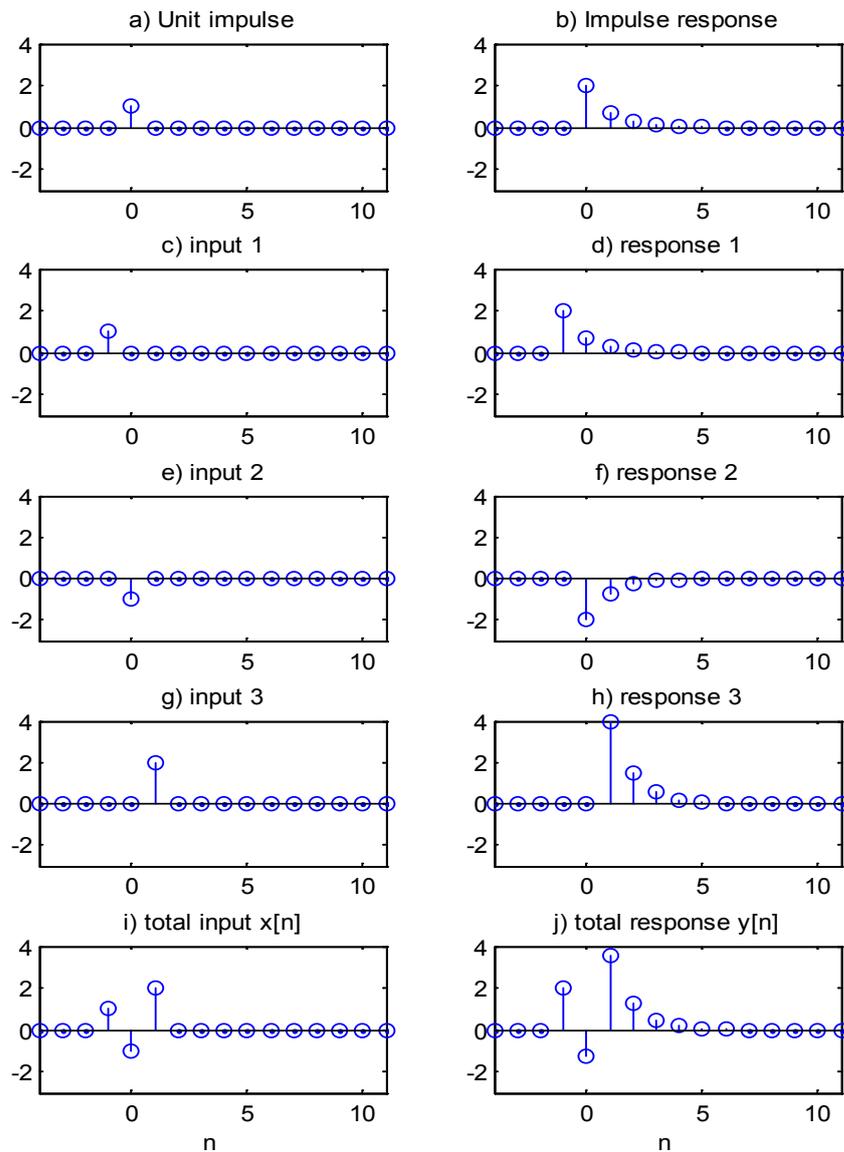$$y[n] = \sum_{k=\infty}^{\infty} x[k] \ h[n-k]^{\cdot}$$

Figure 3.8 Graphic expression for digital convolution

Problems

Q3.1 Many systems can be described by a continuous form of first order differential equation:

$$\frac{dy(t)}{dt} + ay(t) = b\frac{dx(t)}{dt} + cx(t)$$

Derive the corresponding discrete form of difference equation if the sampling interval is *T*.

Q3.2 A second order differential equation is:

$$\frac{dy^2(t)}{dt^2} + 2\frac{dy(t)}{dt} + 3y(t) = \frac{dx(t)}{dt} + 4x(t)$$

Derive the corresponding discrete form of difference equation if the sampling interval is T.

Q3.3 Draw a block diagram for the digital system described by the difference equation:

a) $y[n] = 0.7y[n-1] + x[n] - 0.5x[n-1]$

b) $y[n] = 0.35y[n-1] + 0.65y[n-2] + 2x[n] - 0.55x[n-2]$.

Q3.4. Evaluate the impulse responses of the following filters upto n=6:

a) $y[n] = 3y[n-1] + x[n]$

b) $y[n] = 3y[n-1] + 2x[n-2]$

c) $y[n] = 3y[n-1] + x[n] - 2x[n-1] + 3x[n-2]$

Q3.5 Assuming that a digital system described by the following difference equation is causal, evaluate the impulse response for the system from $n = -\infty$ up to $n = 4$.

$$y[n] = 0.5 y[n-1] + 0.25 y[n-2] + x[n]$$

Q3.6 For LTI digital systems, the output is the convolution between the input and the impulse response. Calculate the output $y[n]$ by manual convolution up to *n=5* for the given impulse responses and inputs:

a)
$$h[n] = \begin{bmatrix} 2 & 1 & 3 & 4 & ... & \end{bmatrix} \qquad x[n] = \begin{bmatrix} 1 & 2 & 3 & ... \end{bmatrix}$$
$$\uparrow \qquad\qquad\qquad\qquad\qquad \uparrow$$

b)
$$h[n] = \begin{bmatrix} 4 & 3 & 2 & 1 & 0 & ... \end{bmatrix} \qquad x[n] = \begin{bmatrix} 2 & -1 & 0 & ... \end{bmatrix}$$
$$\uparrow \qquad\qquad\qquad\qquad\qquad \uparrow$$

c)
$$h[n] = (-0.5)^n u[n]$$
$$x[n] = [10, \ 10]$$
$$\uparrow$$

41

4   Frequency Domain Analysis

4.1   Fourier series for periodic digital signals

Consider a periodic digital signal x[n], n=0,1,2,...,N-1, where N is the number of sample values in each period. From Euler's complex exponential equation:

$$\exp\left(-j\frac{2\pi\,kn}{N}\right) = \cos\left(\frac{2\pi\,kn}{N}\right) - j\sin\left(\frac{2\pi\,kn}{N}\right) \quad k=0,1,2,...,N\text{-}1 \tag{4.1}$$

for each frequency k, Eq. (4.1) contains 2 sinusoidal functions in real and imaginary parts, respectively, with π/2 difference in phase. The frequencies of the functions are *k=0,1,2,...,N-1*. The fundamental sinusoidal function is when the frequency k=1. The other higher sinusoidal functions are called harmonics. The coefficients of Fourier Series for a digital signal can be calculated by

$$a_k = \frac{1}{N}\sum_{n=0}^{N-1} x[n]\exp\left(-j\frac{2\pi\,kn}{N}\right) \quad k=1,2,\ldots,N\text{-}1 \tag{4.2}$$

where a_k is *k*th spectral coefficient, indicating the strength of the *k*th harmonic function. The original digital signal x[n] can be represented by its constituent harmonics as the form of discrete Fourier series:

$$x[n] = \sum_{k=0}^{N-1} a_k \exp\left(j\frac{2\pi\,kn}{N}\right). \tag{4.3}$$

The following Figure 4.1 illustrates how many complex multiplications are required in calculating those coefficients a_k and from a_k to obtaining $x[n]$. It is basically N^2 in each occasion.

$$x[0] \qquad a_0$$
$$x[1] \qquad a_1$$
$$x[2] \qquad a_2$$
$$\ldots \qquad \ldots$$
$$x[N-1] \qquad a_{N-1}$$

$$a_0 \qquad x[0]$$
$$a_1 \qquad x[1]$$
$$a_2 \qquad x[2]$$
$$\ldots \qquad \ldots$$
$$a_{N-1} \qquad x[N-1]$$
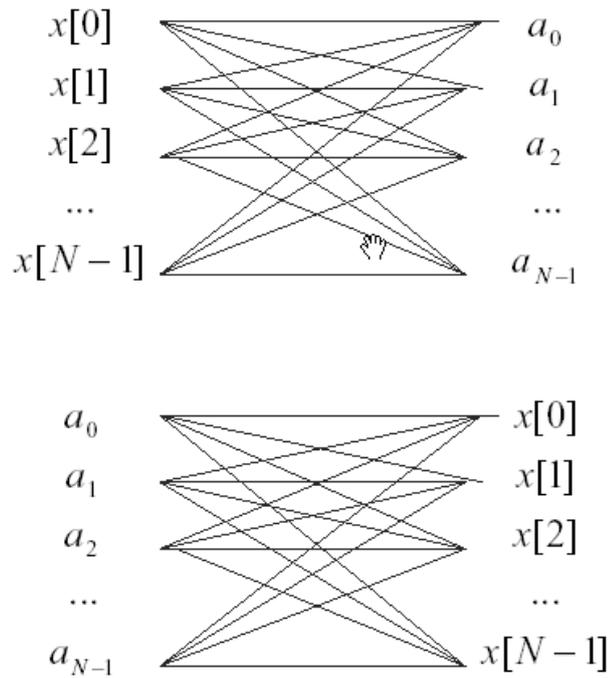
Figure 4.1 Number of complex multiplications

Example 4.1 A periodic signal $x[n] = \begin{bmatrix} 0 & 2 & 4 & 6 \end{bmatrix}$ has the period N=4, its coefficients of Fourier series can be calculated by the formula

$$a_k = \frac{1}{4} \sum_{n=0}^{3} 2n \exp\left(-j\frac{2\pi kn}{4}\right) \qquad for\ 0 \le k \le 3$$

Therefore,

$$a_0 = \frac{1}{4}\left[0 + 2\times1\exp(0) + 2\times2\exp(0) + 2\times3\exp(0)\right] = \frac{1}{4}(0+2+4+6) = 3 ,$$

$$a_1 = \frac{1}{4}\left[0 + 2\times1\exp\left(-j\frac{2\pi}{4}\right) + 2\times2\exp\left(-j\frac{2\pi2}{4}\right) + 2\times3\exp\left(-j\frac{2\pi3}{4}\right)\right]$$
$$= \frac{1}{4}\left(0 + 2(-j) + 4(-1) + 6(j)\right)$$
$$= -1 + j$$

$$a_2 = \frac{1}{4}\left[0 + 2\times1\exp\left(-j\frac{2\pi2}{4}\right) + 2\times2\exp\left(-j\frac{2\pi2\times2}{4}\right) + 2\times3\exp\left(-j\frac{2\pi2\times3}{4}\right)\right]$$
$$= \frac{1}{4}\left(0 + 2(-1) + 4(1) + 6(-1)\right)$$
$$= -1$$

$$a_3 = \frac{1}{4}\left[0 + 2\times1\exp\left(-j\frac{2\pi3}{4}\right) + 2\times2\exp\left(-j\frac{2\pi3\times2}{4}\right) + 2\times3\exp\left(-j\frac{2\pi3\times3}{4}\right)\right]$$
$$= \frac{1}{4}\left(0 + 2(j) + 4(-1) + 6(-j)\right)$$
$$= -1 - j$$

The spectral coefficients are

$$a_k = [3,\ -1+j,\ -1,\ -1-j]$$, and the modulus is

$$\left|a_k\right| = [3,\ 1.4142,\ 1,\ 1.4142]$$.

Figure 4.2 illustrates one period of the signal and its spectral coefficients.

Example 4.2 A sine wave $x[n] = \sin\left(\frac{2\pi\ 4n}{64}\right)$.. With k=4 and N=64, the sine wave and its Fourier spectrum are shown in the following Figure 4.3. It can be seen that the single sine wave gives a single spectral line at frequencies of 4 hz and 64-4=60 hz, or $\frac{\pi}{8}$ and $2\pi - \frac{\pi}{8} = \frac{15\pi}{8}$ radians/sample.

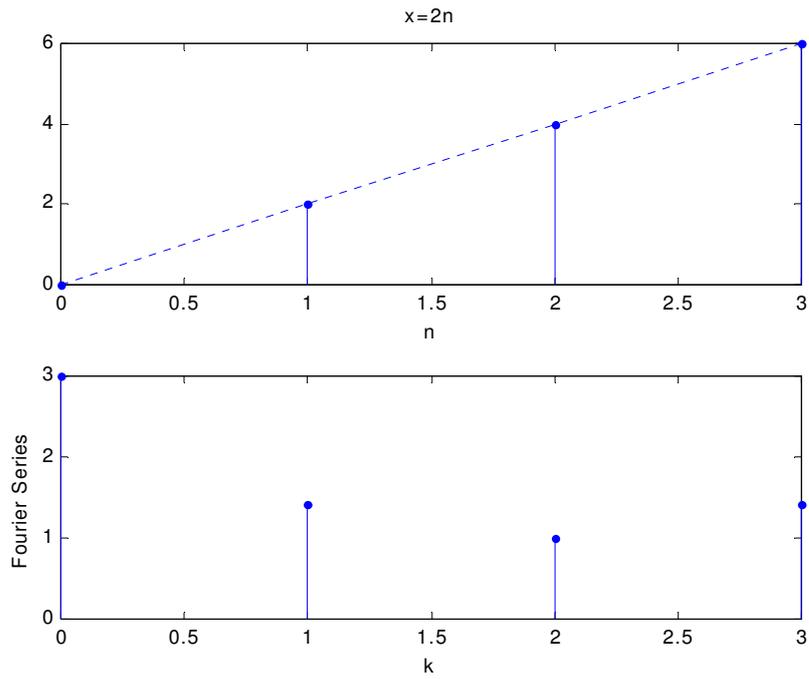**Figure 4.2** Spectral coefficients of Fourier series
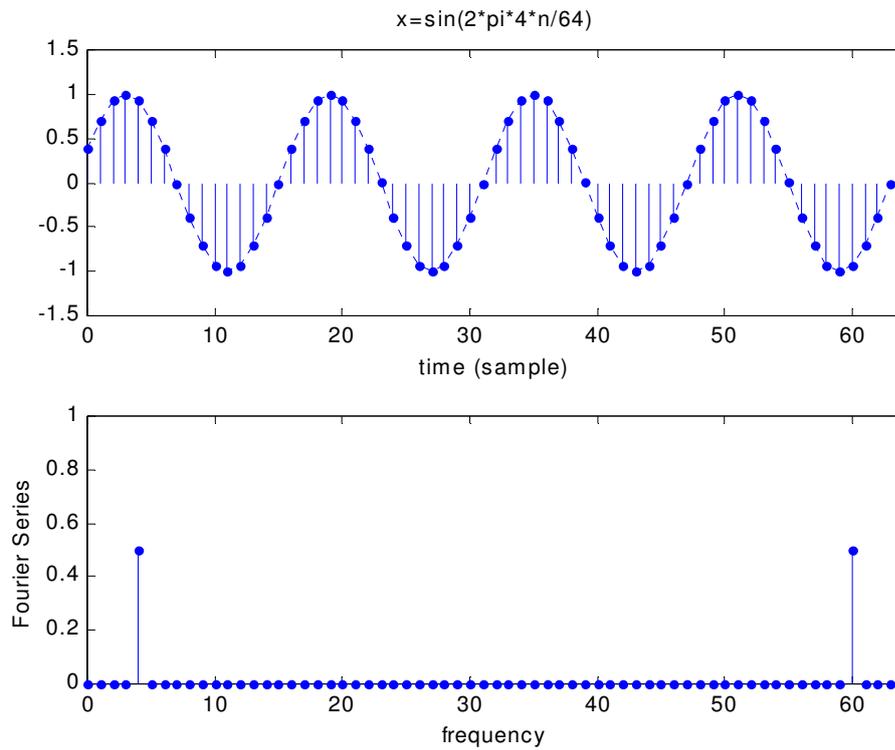

Figure 4.3 Sine wave and its spectral coefficients of Fourier series

4.2   Fourier transform for non-periodic signals

The Fourier series introduced in previous section is only applicable to periodic signals. In general cases, the signal is not periodic, and the Fourier series formula Eq. (4.2) is not applicable. Therefore, a Fourier Transform for non-periodic signals must be introduced.

Let frequency $\Omega = \dfrac{2\pi\,k}{N}$ (rad/sample) changes continuously between 0 and 2π. This can be achieved by assuming N is big enough and the fundamental frequency $\Omega_0 = \dfrac{2\pi}{N}$ is very small. Define a Fourier transform as

$$X(\Omega) = \sum_{n=\infty}^{\infty} x[n]\exp(-j\Omega n) \tag{4.4}$$

where $X(\Omega)$ is the Fourier spectrum, which is a continuous function, therefore, usual round brackets are used. It is also a periodic function with a period of 2π as sinusoidal functions are multiplied in the transform.

Assuming N is big and using Eq.(4.3), the inverse Fourier transform can be defined:

$$x[n] = \sum_{k=0}^{N-1} a_k \exp\left(j\frac{2\pi\, k}{N} \right)$$

$$= \sum_{k=0}^{N-1} \left\{ \frac{1}{N} X(k\Omega_0) \right\} \exp(-k\,\Omega_0 n)$$

$$= \sum_{k=0}^{N-1} \left\{ \frac{\Omega_0}{2\pi} X(k\Omega_0) \right\} \exp(-k\,\Omega_0 n)$$

$$= \frac{1}{2\pi} \sum_{k=0}^{N-1} X(k\Omega_0) \exp(-k\,\Omega_0 n)\Omega_0$$

With $N \to \infty$, $\Omega_0 \to d\Omega$ and $k\Omega_0 \to \Omega$, the above inverse Fourier transform can be expressed as

$$x[n] = \frac{1}{2\pi} \int_0^{2\pi} X(\Omega) \exp(j\Omega n)\, d\Omega$$

$$(4.5)$$

Example 4.3 Find the Fourier transform for the unit impulse signal $x[n] = \delta\,[n]$.

From the definition Eq.(4.4),

$$X(\Omega) = \sum_{n=\infty}^{\infty} \delta[n] \exp(-j\Omega n) = 1$$

If the unit impulse is shifted,

$$x[n] = \delta\,[n-1]$$

$$X(\Omega) = \sum_{n=\infty}^{\infty} \delta[n-1] \exp(-j\Omega n) = \exp(-j\Omega)$$

$$\left| X(\Omega) \right| = 1, \qquad \angle X(\Omega) = -\Omega .$$

Example 4.4 Find the Fourier transform of the rectangle signal

$$x[n] = \frac{1}{5}\left\{ \delta[n-2] + \delta[n-1] + \delta[n] + \delta[n+1] + \delta[n+2] \right\}$$

$$X(\Omega) = \sum_{n=\infty}^{\infty} x[n] \exp(-j\Omega n)$$

$$= \sum_{n=\infty}^{\infty} 0.2\left\{ \delta[n-2] + \delta[n-1] + \delta[n] + \delta[n+1] + \delta[n+2] \right\} \exp(-j\Omega n)$$

i.e.

$$X(\Omega) = 0.2\{\exp(-j2\Omega) + \exp(-j\Omega) + 1 + \exp(j\Omega) + \exp(j2\Omega)\}$$
$$= 0.2\{1 + 2\cos(2\Omega) + 2\cos(\Omega)\}$$

It is a real function. $\left|X(\Omega)\right| = 0.2\left|1 + 2\cos(2\Omega) + 2\cos(\Omega)\right|$, $\angle X(\Omega) = 0$.

Example 4.5 A signal is described by

$$x[n] = 0.5\delta[n] + 0.5^2\delta[n-1] + 0.5^3\delta[n] + \ldots$$

Apply the Fourier transform

$$X(\Omega) = \sum_{n=\infty}^{\infty} x[n]\exp(-j\Omega n)$$

$$= 0.5 + 0.25\exp(-j\Omega) + 0.125\exp(-j2\Omega) + \ldots$$
$$= 0.5\sum_{n=0}^{\infty}\{0.5\exp(-j\Omega)\}^n = \frac{0.5}{1 - 0.5\exp(-j\Omega)}$$

It is a complex function. The magnitude is

$$\left|X(\Omega)\right| = \frac{0.5}{\left|1 - 0.5\exp(-j\Omega)\right|}$$

$$= \frac{0.5}{\left\{(1 - 0.5\cos\Omega)^2 + (0.5\sin\Omega)^2\right\}^{1/2}}$$

$$= \frac{0.5}{\left\{1 - \cos\Omega + 0.25(\cos\Omega)^2 + 0.25(\sin\Omega)^2\right\}^{1/2}}$$

$$= \frac{0.5}{\left\{1.25 - \cos\Omega\right\}^{1/2}}$$

The phase:

$$\Phi_X(\Omega) = \angle(0.5) - \angle\left(1 - 0.5\exp(-j\Omega)\right)$$
$$= -\tan^{-1}\frac{0.5\sin\Omega}{1 - 0.5\cos\Omega}$$

4.3   Properties of Fourier transform

The linearity is held by the Fourier transform, i.e.:

If $x_1[n] \leftrightarrow X_1(\Omega)$ and $x_2[n] \leftrightarrow X_2(\Omega)$

Then $Ax_1[n] + Bx_2[n] \leftrightarrow AX_1(\Omega) + BX_2(\Omega)$

where $\leftrightarrow$ denotes the pair between the Fourier transform and inverse Fourier transform.

The time-shifting property is also held, i.e.:

If $x\ [n] \leftrightarrow X(\Omega)$

Then $x[n - n_0] \leftrightarrow X(\Omega)\exp(-j\Omega n_0)$

The Fourier transform of a convolution in time-domain is a product of Fourier transforms in the frequency domain, i.e.:

If $x_1[n] \leftrightarrow X_1(\Omega)$ and $x_2[n] \leftrightarrow X_2(\Omega)$,

Then $x_1[n] * x_2[n] \leftrightarrow X_1(\Omega)X_2(\Omega)$.

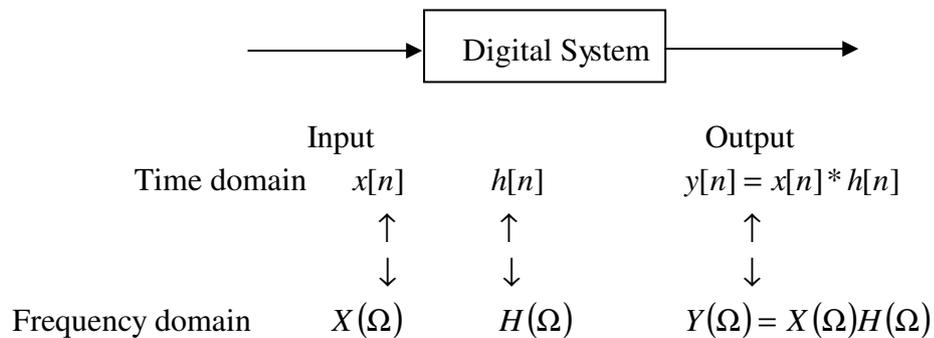The Fourier Transform properties and transform pairs are listed in Table 4.1.

4.4　Frequency response

Consider digital LTI systems, take Fourier transform for the input, impulse response and output:

$$X(\Omega) = \sum_{n=\infty}^{\infty} x[n]\exp\left(-j\Omega n\right)$$
,

$$H(\Omega) = \sum_{n=\infty}^{\infty} h[n]\exp\left(-j\Omega n\right)$$
,

$$Y(\Omega) = \sum_{n=\infty}^{\infty} y[n]\exp\left(-j\Omega n\right)$$
,

	Input		Output
Time domain	$x[n]$	$h[n]$	$y[n] = x[n] * h[n]$
	↑	↑	↑
	↓	↓	↓
Frequency domain	$X(\Omega)$	$H(\Omega)$	$Y(\Omega) = X(\Omega)H(\Omega)$

Figure 4.4 Input-output relationships in time and frequency domains

where $H(\Omega)$ is known as the frequency response, which describes the gain of a system at different frequencies and can be obtained by taking Fourier transform to the impulse response. The input-output relationship in the time and frequency domains, and the time-frequency relationships of the three quantities are shown in Figure 4.4.

Using the convolution property of the Fourier transform, for the input-output relationship $y[n] = \sum_{k=\infty}^{\infty} h[k]\ x[n-k]$, apply the Fourier transform on both sides, it can be obtained

$$Y(\Omega) = H(\Omega)X(\Omega)$$

(4.6)

The frequency response can be obtained from the transforms of input and output:

$$H(\Omega) = \frac{Y(\Omega)}{X(\Omega)}$$

(4.7)

Table 4.1: Fourier Transform Properties and Pairs

Signal $x[n]$	Fourier Transform $X(\Omega)$	Property
$ax[n]+by[n]$	$aX(\Omega)+bY(\Omega)$	Linearity
$x[n-n_0]$	$X(\Omega)\exp(-j\Omega n_0)$	Time-shift
$x[n]*y[n]$	$X(\Omega)Y(\Omega)$	Convolution
$x[n]y[n]$	$X(\Omega)*Y(\Omega)$	Modulation
$\delta[n]$	1	
$\delta[n-n_0]$	$\exp(-j\Omega n_0)$	
$u[n]$	$(1-\exp(-j\Omega))^{-1} + \sum_{k=-\infty}^{\infty} \pi\delta(\Omega-2\pi k)$	
$a^n u[n] \quad |a|<1$	$(1-a\exp(-j\Omega))^{-1}$	
$x[n]=1 \quad |n|\le m$ $x[n]=0 \quad |n|>m$	$\dfrac{\sin\{(m+1/2)\Omega\}}{\sin(\Omega/2)}$	

Example 4.6: Find the frequency response from the impulse responses:

     a)   h[n]=0.2 {δ[n-2]+ δ[n-1]+ δ[n]+ δ[n+1]+ δ[n+2]}

Taking the Fourier transform, the frequency response is

$$H(\Omega) = \sum_{n=\infty}^{\infty} h[n]\exp(-j\Omega n)$$
$$= 0.2\{\exp(-j2\Omega)+\exp(-j\Omega)+1+\exp(j\Omega)+\exp(j2\Omega)\}$$
$$= 0.2\{1+2\cos\Omega+2\cos 2\Omega\}$$

$|H(\Omega)| = 0.2|1+2\cos\Omega+2\cos 2\Omega|$ and the phase $\Phi_H(\Omega)=0$.
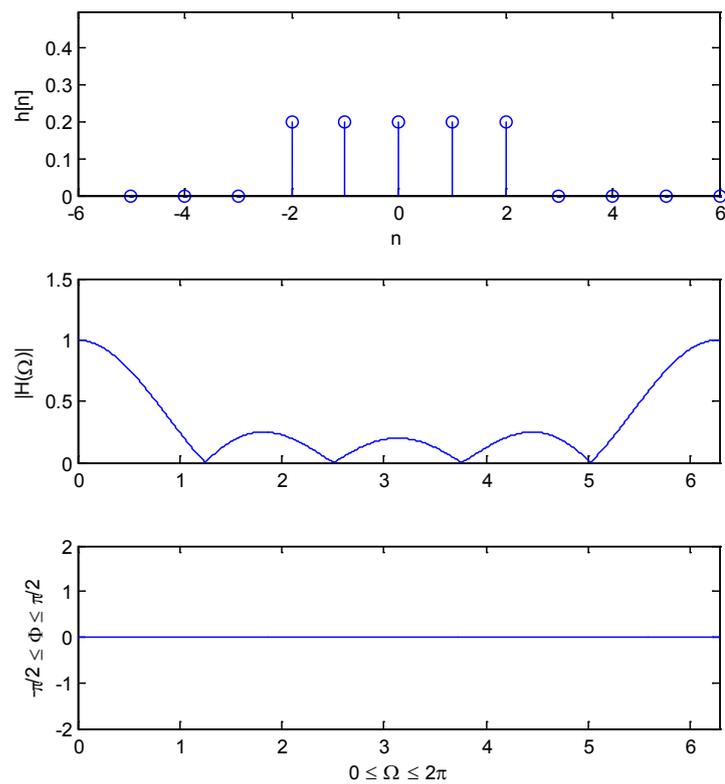
Figure 4.5 Impulse response and frequency response in Example 4.6 (a)

b) $h[n] = 0.5\delta[n] + 0.25\delta[n-1] + 0.125\delta[n-2] + ...$

$$H(\Omega) = \sum_{n=-\infty}^{\infty} h[n]\exp(-j\Omega n) = 0.5 + 0.25\exp(-j\Omega) + 0.125\exp(-j2\Omega) + ...$$

$$= 0.5\sum_{n=0}^{\infty}\{0.5\exp(-j\Omega)\}^{n}$$

i.e. $H(\Omega) = \dfrac{0.5}{1 - 0.5\exp(-j\Omega)}$.
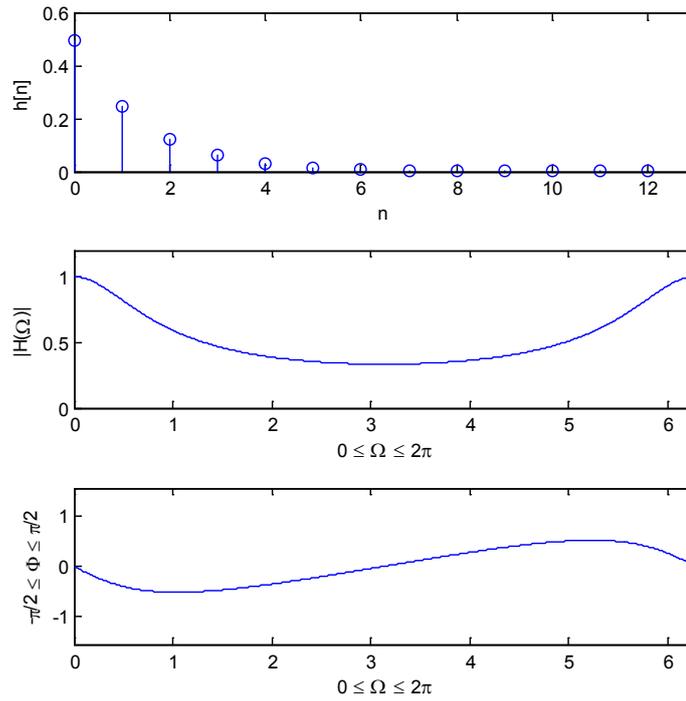
The modulus and phase are shown in Figure 4.6.

Figure 4.6 Impulse response and frequency response in Example 4.6 (b)

In general cases, a system can be described by the difference equation

$$\sum_{k=0}^{N} a_k\, y[n-k] = \sum_{k=0}^{M} b_k\, x[n-k]$$

Take Fourier transform for both sides,

$$\sum_{k=0}^{N} a_k \exp(-jk\Omega) Y(\Omega) = \sum_{k=0}^{M} b_k \exp(-jk\Omega) X(\Omega)$$

or

$$Y(\Omega)\sum_{k=0}^{N} a_k \exp(-jk\Omega) = X(\Omega)\sum_{k=0}^{M} b_k \exp(-jk\Omega)$$

Therefore, the frequency response can be obtained by

$$H(\Omega) = \frac{Y(\Omega)}{X(\Omega)} = \frac{\displaystyle\sum_{k=0}^{M} b_k \exp(-j k\,\Omega)}{\displaystyle\sum_{k=0}^{N} a_k \exp(-j k\,\Omega)} \tag{4.8}$$

Example 4.7: A digital high-pass filter is described by

y[n]= - 0.8 y[n-1] + x[n]- x[n-1]

Find the frequency response and sketch its magnitude and phase over the range 0<Ω<π.

Solution: Re-arrange the difference equation as

y[n] + 0.8 y[n-1] = x[n]- x[n-1]

Take Fourier transform for both sides

$$Y(\Omega) + 0.8\exp(-j\Omega)Y(\Omega) = X(\Omega) - \exp(-j\Omega)X(\Omega)$$

or

$$(1 + 0.8\exp(-j\Omega)\, Y(\Omega) = (1 - \exp(-j\Omega)\, X(\Omega)$$

The frequency response is:

$$H(\Omega) = \frac{1 - \exp(-j\Omega)}{1 + 0.8\exp(-j\Omega)} \quad or \quad H(\Omega) = \frac{1 - \cos\Omega + j\sin\Omega}{1 + 0.8\cos\Omega - 0.8\,j\sin\Omega}$$

Its magnitude is

$$\left|H(\Omega)\right| = \frac{\left|1 - \exp(-j\Omega)\right|}{\left|1 + 0.8\exp(-j\Omega)\right|} = \frac{\left\{(1 - \cos\Omega)^2 + (\sin\Omega)^2\right\}^{1/2}}{\left\{(1 + 0.8\cos\Omega)^2 + (0.8\sin\Omega)^2\right\}^{1/2}} = \left\{\frac{2 - 2\cos\Omega}{1.64 + 1.6\cos\Omega}\right\}^{1/2}$$

And the phase

$$\Phi_H(\Omega) = arc\ \tan\left(\frac{\sin\Omega}{1 - \cos\Omega}\right) - arc\ \tan\left(\frac{-0.8\sin\Omega}{1 + 0.8\cos\Omega}\right)$$. See Figure 4.7.
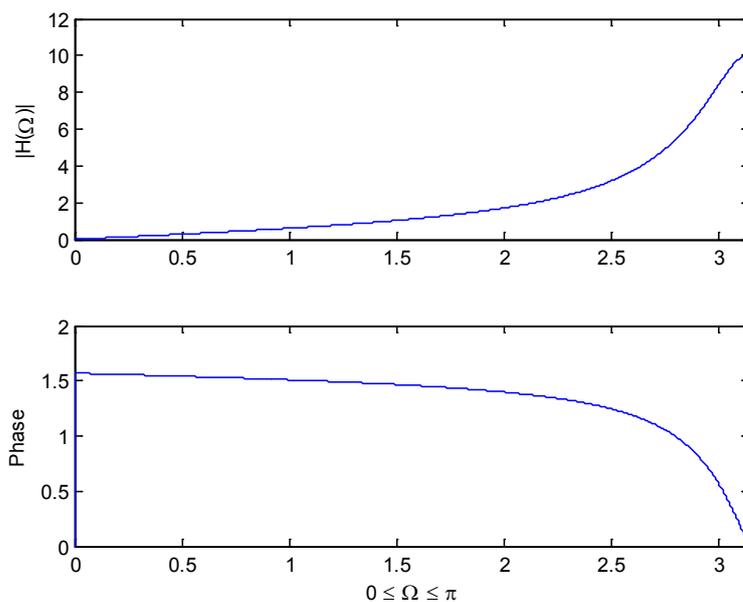
Figure 4.7 Modulus and phase of frequency response

4.5   Frequency correspondence when sampling rate is given

Let x[n] be a signal discretized with a sampling rate of f_s Hz, we are about to find the position of peak for a complex harmonic component with frequency f_0 Hz $\exp\left(j\dfrac{2\pi f_0}{f_s}n\right)$ in the frequency Ω (rad/sample) domain. Its Fourier transform is given by

$$X(\Omega) = \sum_{n=\infty}^{\infty} \exp\left(j\frac{2\pi f_0}{f_s}n\right)\exp\left(-j\Omega n\right)$$
$$= \sum_{n=\infty}^{\infty} \exp\left(j\left(\frac{2\pi f_0}{f_s}-\Omega\right)n\right)$$

i.e.

$$X(\Omega) = 2\pi\delta\left(\Omega - \frac{2\pi f_0}{f_s}\right)$$

Therefore, the peak appears at $\Omega = 2\pi\dfrac{f_0}{f_s}$. According to the Nyquist sampling theorem, the maximum frequency in the signal is $f_0 = f_s/2$, corresponding to $\Omega = 2\pi\dfrac{f_s/2}{f_s} = \pi$ (rad/sample).

Note the following properties exist:

$$X(\Omega + 2\pi) = \sum_{n=\infty}^{\infty} x[n]\exp(-j\Omega n - j2\pi n) = X(\Omega)$$

, and

$$\left|X(\pi + \Omega)\right| = \left|X(\pi - \Omega)\right|$$

Therefore, with given sampling rate f_s (Hz), the correspondence between frequency Ω (rad/sample) and f (Hz) is

$$\Omega = 2\pi \frac{f}{f_s} \text{ (rad/sample)}$$

or

$$f = \frac{\Omega}{2\pi} f_s \text{ (Hz)}$$

Particularly, for one period in the frequency domain:

$$
\begin{array}{ccc}
\Omega & & f \\
0 & \rightarrow & 0 \\
\pi & \rightarrow & f_s/2 \\
2\pi & \rightarrow & f_s
\end{array}
$$

When sampling rate is given, the frequency correspondence between frequency Ω and f domains is depicted in Figure 4.8.
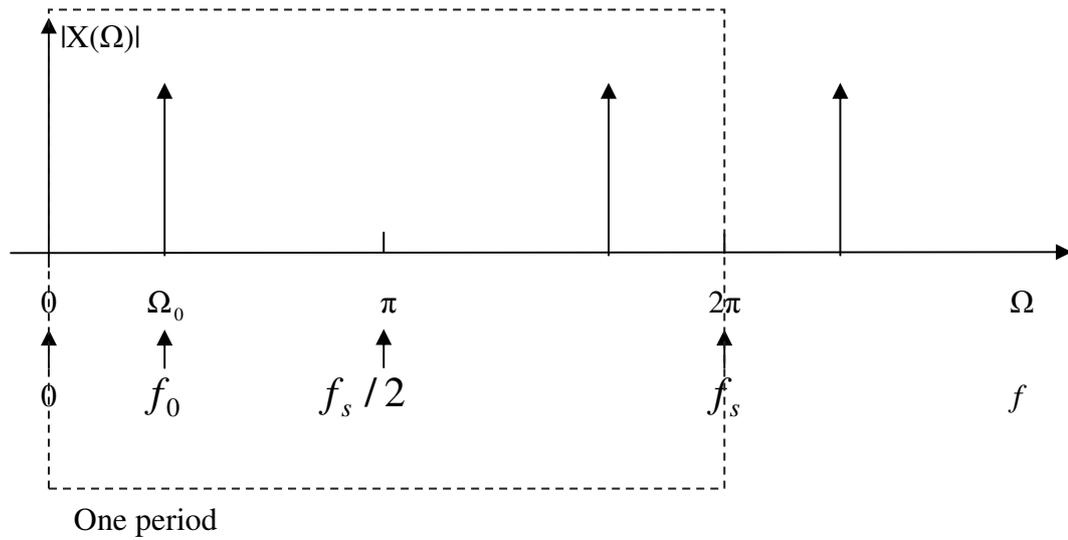
Figure 4.8 Frequency correspondence between frequencies in different units.

Problems

Q4.1 Find the frequency response from the impulse response

$h[n]= \delta[n]+ 0.5\delta[n-1]+ 0.25\delta[n-2]+...$
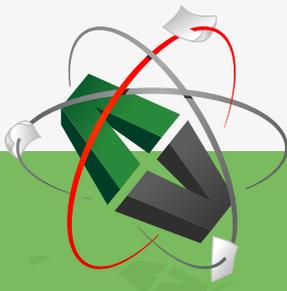
Q4.2 A digital system has the difference equation

$$y[n] = 0.5y[n-1] + y[n-2] + x[n]$$

Find the frequency response.

Q4.3 A digital system has the frequency response:

$$H(\Omega) = \frac{\exp(j\Omega)}{\exp(j\Omega) + 0.8}$$

Sketch the magnitude of its frequency response $\left|H(\Omega)\right|$ over $0 < \Omega < \pi$.

5 Z Domain Analysis

5.1   z-transform and inverse z-transform

In the previous chapters, time domain and frequency domain analysis have been introduced. In each of those domains different insights of digital signals are revealed. It is useful to introduce another domain: the z domain. A digital time signal can be transferred into z domain by the z-transform. The z-transform is defined as

$$X(z) = \sum_{n=0}^{\infty} x[n]z^{-n}$$

(5.1)

where z is a complex variable. The transform defined by Eq. (5.1) is a unilateral transform as defined on one side of the axis $0 \le n < \infty$. In the transform, each sample $x[n]$ is multiplied by the complex variable z^{-n}, i.e.

$$
\begin{array}{ccccc}
x[0] & x[1] & x[2] & x[3] & ... \\
z^0 & z^{-1} & z^{-2} & z^{-3} & ... \\
\uparrow & & & &
\end{array}
$$

There is advantage in this unilateral transform definition as it can avoid mathematical inconvenience. One can shift the signal of interest to obtain a required origin in its analysis, thus usually causing no trouble in applications.

The inverse z-transform can be found by

$$x[n] = \frac{1}{2\pi \ j} \oint X(z)z^{n-1}dz$$

(5.2)

It involves contour integration, and further discussion is beyond the scope of this basic content. However, an alternative approach is available using partial fractions together with z-transform formulas of basic functions. Table 5.1 lists the basic properties of the z transform and Table 5.2 lists some basic z- transform pairs.

Example 5.1 Find the z transform for a signal and reconstruct a signal from its z-transform.

a) x[n]=[$\begin{array}{cccc}1 & 0.8 & 0.8^2 & 0.8^3 \end{array}$ ...] is shown in Figure 5.1(a), find the z-transform.
$\qquad\uparrow$

Using the definition Eq.(5.1),

$$X(z) = \sum_{n=0}^{\infty} x[n]z^{-n} = 1 + 0.8z^{-1} + 0.8^2 z^{-2} + 0.8^3 z^{-3} + \ldots$$

$$= 1 + 0.8z^{-1} + \left(0.8z^{-1}\right)^2 + \left(0.8z^{-1}\right)^3 + \ldots$$

$$= \frac{1}{1 - 0.8z^{-1}} = \frac{z}{z - 0.8}$$

    b) Reconstruct the signal corresponding to the z-transform: $X(z) = \dfrac{1}{z + 1.2}$

$$X(z) = \frac{1}{z + 1.2} = \frac{z^{-1}}{1 + 1.2z^{-1}} = z^{-1}\left(\frac{1}{1 - (1.2z^{-1})}\right)$$

Using the series summation formula in Eq.(3.13) (where $\left|1.2z^{-1}\right| < 1$ is required.),

$$X(z) = z^{-1}\left\{1 + (-1.2z^{-1}) + (-1.2z^{-1})^2 + (-1.2z^{-1})^3 + \ldots\right\}$$

$$= z^{-1} - 1.2z^{-2} \_ 1.44z^{-3} - 1.728z^{-4} + \ldots$$

Therefore, the reconstructed original signal can be obtained as

      x[n]=[0 1 -1.2 1.44 -1.728 ...].
        ↑

The signal is shown in Figure 5.1(b).

Figure 5.1 Signals in the examples
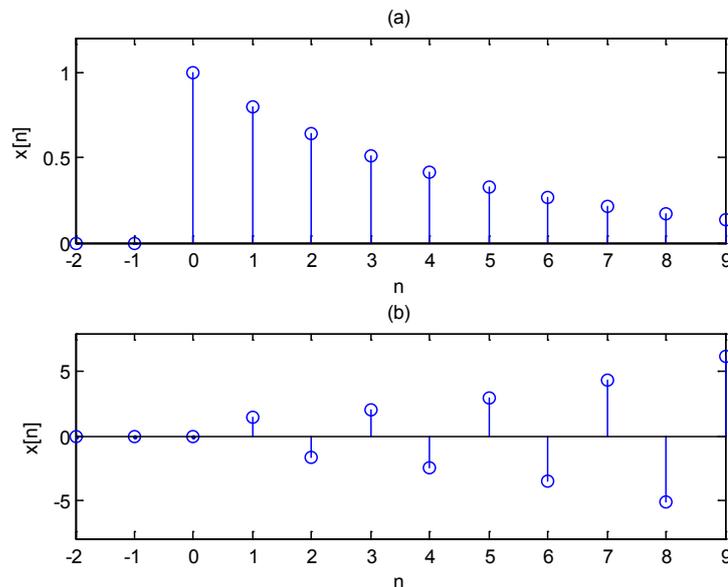
5.2   Relationship between z-transform and Fourier transform

Let $z = \exp(j\Omega)$, i.e. the complex variable z is only allowed on the unit circle, the z-transform becomes a unilateral Fourier transform

$$X(\Omega) = \sum_{n=0}^{\infty} x[n]\exp(-j\Omega n)$$

Obviously, apart from on the unit circle, the complex operator z can be specified into other curves or region, if necessary. Later, it will be shown the unit circle is important boundary on the z-domain.

5.3   Z as time shift operator

Multiplying by z implies a time advance and dividing by z, or multiplying by z^{-1}, is to cause a time delay. For the unit impulse,

$$X(z) = \sum_{n=0}^{\infty} \delta[n]z^{-n} = z^{-n}\Big|_{n=0} = 1$$
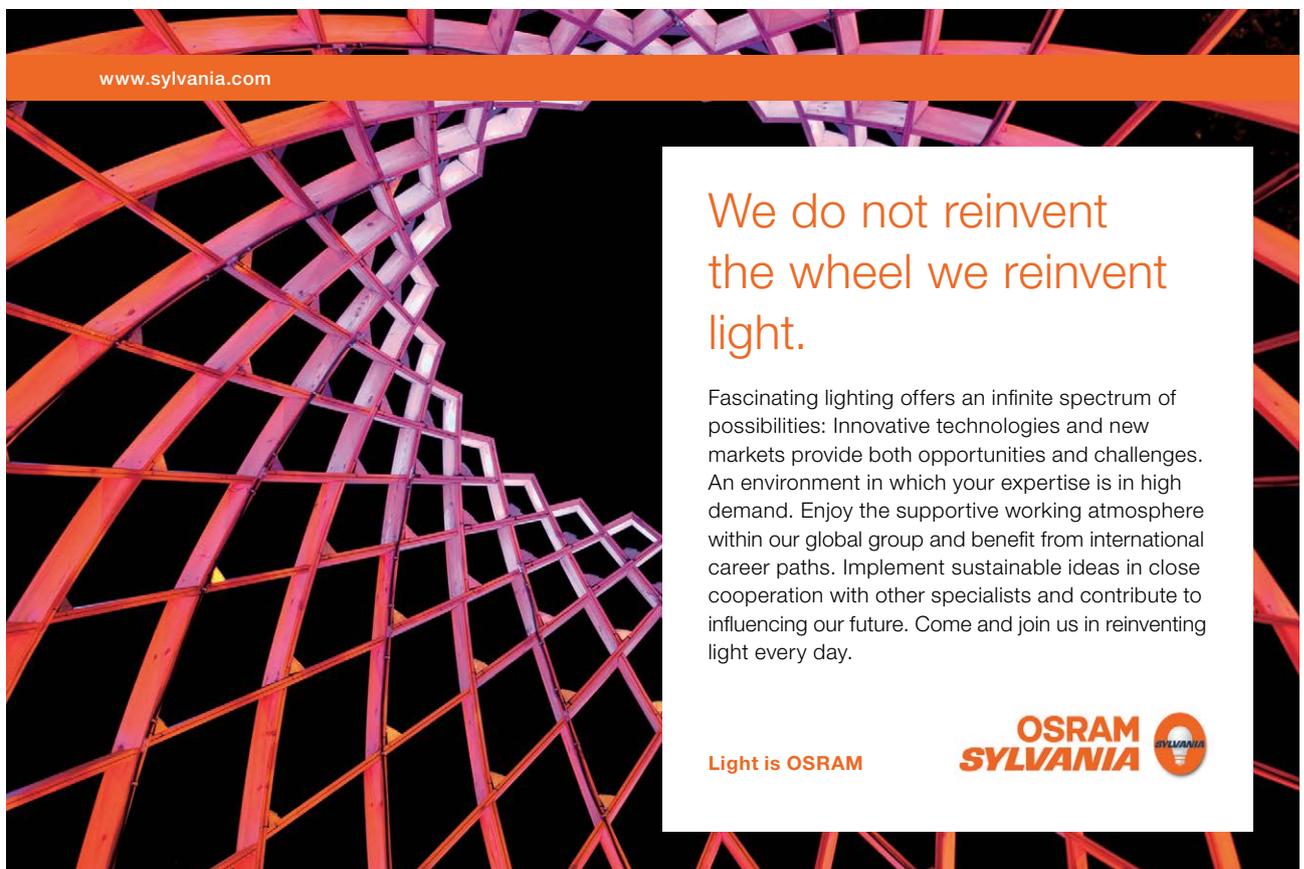
For the delayed unit impulse,

$$X(z) = \sum_{n=0}^{\infty} \delta [n-1]z^{-n} = z^{-n}\Big|_{n=1} = z^{-1}$$

For more general cases of shifting by n_0 samples,

$$X(z) = \sum_{n=0}^{\infty} \delta [n-n_0]z^{-n} = z^{-n}\Big|_{n=n_0} = z^{-n_0}$$

For shifted signal x[n], i.e. delayed by n_0 samples, the z-transform

$$\sum_{n=0}^{\infty} x [n-n_0] u[n-n_0]z^{-n} = X(z)z^{-n_0}$$

5.4   Transfer function

The transfer function describes the input-output relationship, or the transmissibility between input and output, in the z-domain. Applying the z-transform to the output of a system, the relationship between the z-transforms of input and output can be found:

$$Y(z) = \sum_{n=0}^{\infty} y[n]z^{-n} = \sum_{n=0}^{\infty} \sum_{r=\infty}^{\infty} x[r]h[n-r]z^{-n}$$

$$= \sum_{r=\infty}^{\infty} x(r)z^{-r} \sum_{n=0}^{\infty} h[n-r]z^{-(n-r)}$$

$$= \sum_{r=\infty}^{\infty} x(r)z^{-r} \sum_{m=-r}^{\infty} h[m]z^{-m} = X(z)H(z)$$

Therefore,

$$H(z) = \frac{Y(z)}{X(z)}$$

(5.3)

i.e., the transfer function can be obtained from the z-transforms of input and output.

Alternatively, the transfer function $H(z)$ can be obtained by applying z-transform directly to the impulse response $h[n]$. The relationships of input ($x[n]$ and $X(z)$), output ($y[n]$ and $Y(z)$) and system function ($h[n]$ and $H(z)$) in the time and z domains are depicted in Figure 5.2.

$$\boxed{\text{Digital System}}$$

	Input		Output
Time domain	$x[n]$	$h[n]$	$y[n] = x[n] * h[n]$
	↑	↑	↑
	↓	↓	↓
z-domain	$X(z)$	$H(z)$	$Y(z) = X(z)H(z)$
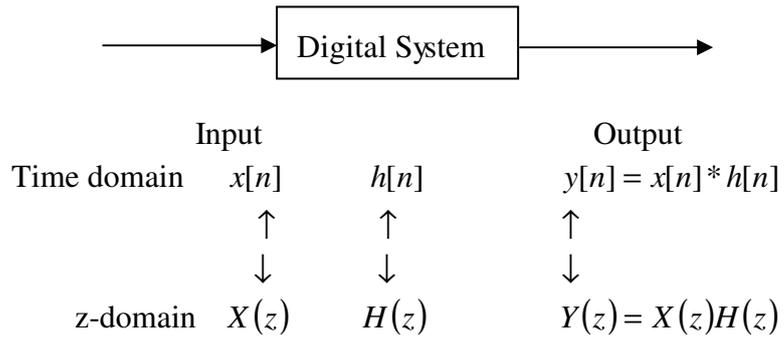
Figure 5.2 Input-output relationship in time and z domains

Instead of using Eq.(5.2), the inverse z-transform can be made through partial fractions. The following are examples.

Example 5.2 A signal has a z-transform $X(z) = \dfrac{1}{z(z-1)(2z-1)}$, find the corresponding original signal $x[n]$.

The z-transform can be represented by partial fractions as

$$X(z) = \frac{1}{z(z-1)(2z-1)} = \frac{A}{z} + \frac{B}{z-1} + \frac{C}{2z-1}$$

where A, B, C are constants, which can be determined by the following steps.

$$X(z) = \frac{A}{z} + \frac{B}{z-1} + \frac{C}{2z-1}$$

$$= \frac{A(z-1)(2z-1) + Bz(2z-1) + Cz(z-1)}{z(z-1)(2z-1)}$$

$$= \frac{A(2z^2 - 3z + 1) + B(2z^2 - z) + C(z^2 - z)}{z(z-1)(2z-1)}$$

$$= \frac{(2A + 2B + C)z^2 + (-3A - B - C)z + A}{z(z-1)(2z-1)}$$

i.e.

$$\frac{1}{z(z-1)(2z-1)} = \frac{(2A + 2B + C)z^2 + (-3A - B - C)z + A}{z(z-1)(2z-1)}$$

The numerator of the left hand side is forced to equal to the right hand side, yielding the following simultaneous equations:

$$\begin{cases} 2A + 2B + C = 0 \\ 3A + B + C = 0 \\ \qquad A = 1 \end{cases}$$

Solving the equations, the constants can be found as *A=1, B=1, C=-4*. Therefore,

$$X(z) = \frac{1}{z} + \frac{1}{z-1} - \frac{4}{2z-1} = z^{-1}\left\{ 1 + \frac{z}{z-1} + \frac{2z}{z-0.5} \right\}.$$

According to basic z-transform pair formulas in Table 5.2,

$$\begin{array}{ccc} 1 & \dfrac{z}{z-1} & \dfrac{2z}{z-0.5} \\ \updownarrow & \updownarrow & \updownarrow \\ \delta[n] & u[n] & 2(0.5)^n u[n] \end{array}$$

Combining the time-shift factor z^{-1}, the signal in time domain can be obtained as

$$x[n] = \delta[n-1] + u[n-1] - 2(0.5)^{n-1} u[n-1].$$

At the values of n this is

I   …
         x[0]=0
         x[1]=0
         x[2]=0
         x[3]=0.5
         x[4]=0.75
         x[5]=0.875
         ...

There is an alternative way of finding the coefficients for the partial fractions. For the above example,

$$X(z) = \frac{1}{z(z-1)(2z-1)} = \frac{A}{z} + \frac{B}{z-1} + \frac{C}{2z-1}$$

A, B and C can be found by the following operations

$$A = z \frac{1}{z(z-1)(2z-1)}\bigg|_{z=0} = 1$$

$$B = (z-1) \frac{1}{z(z-1)(2z-1)}\bigg|_{z=1} = 1$$

$$C = (2z-1) \frac{1}{z(z-1)(2z-1)}\bigg|_{z=1/2} = -4$$

Therefore,

$$X(z) = \frac{1}{z(z-1)(2z-1)} = \frac{1}{z} + \frac{1}{z-1} - \frac{4}{2z-1}.$$

Thus, the signal can be found in the same way:

$$x[n] = \delta[n-1] + u[n-1] - 2(0.5)^{n-1}u[n-1]$$

Table 5.1 Z-transform definition and properties

Definition or property	Signal	z-transform
z-transform definition	$x[n]$	$X(z) = \sum_{n=0}^{\infty} x[n]z^{-n}$
Inverse z-transform	$x[n] = \dfrac{1}{2\pi\ j}\oint X(z)z^{n-1}dz$	$X(z)$
Linearity	$ax_1[n] + bx_2[n]$	$aX_1(z) + bX_2(z)$
Time-shifting property	$x\ [n-n_0]u[n-n_0]$	$X(z)z^{-n_0}$
Convolution	$x[n]*y[n]$	$X(z)Y(z)$

5.5   Z-plane, poles and zeros

For the z-transform of a digital signal or a transfer function of an LTI system, generally it can be expressed as factorised form for both the numerator and denominator:

$$X(z) = \frac{N(z)}{D(z)} = \frac{K(z-z_1)(z-z_2)(z-z_3)...}{(z-p_1)(z-p_2)(z-p_3)...}$$

(5.4)

where $z_1, z_2, z_3 ...$ are called the zeros as which make X(z) =0; and $p_1, p_2, p_3, ...$ are the poles as which make X(z) ° ¥. Zeros and poles are either a real number or complex conjugate pairs.
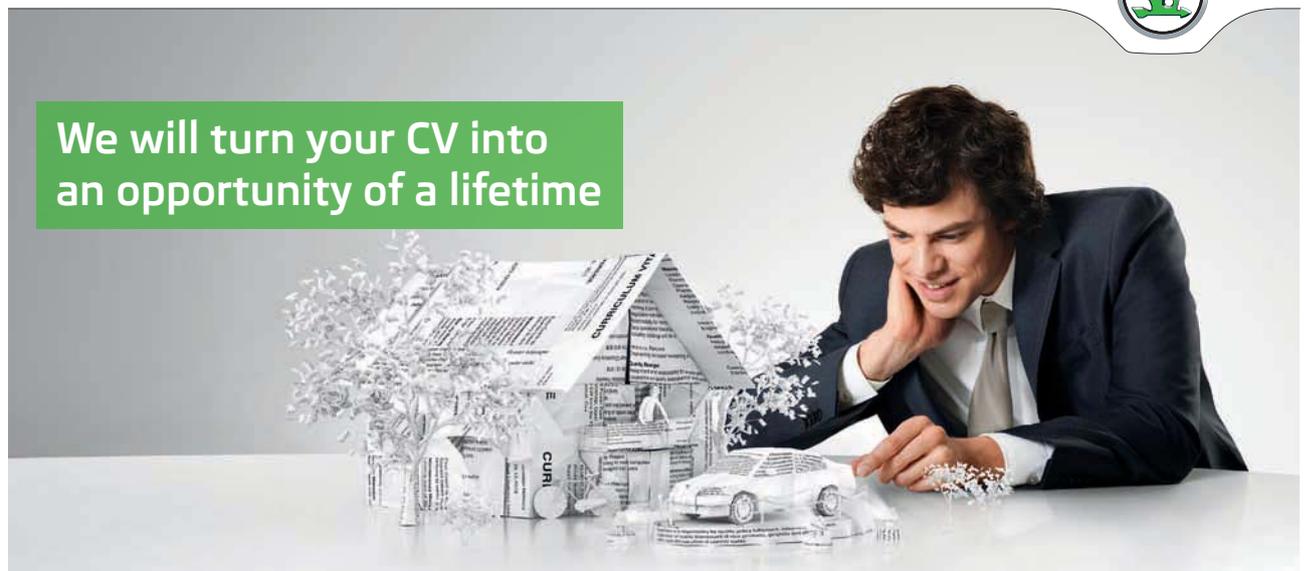
The z-plane is a complex plane in which the zeros and poles of a z-transform are plotted, which is used to visualise the properties of a signal or a system. The positions of the poles and zeros on a z-plane determine the frequency properties and degree of stability. On the other hand, in designing a digital system, the poles and zeros can be chosen to put in appropriate locations for achieving certain required performance.

z-plane
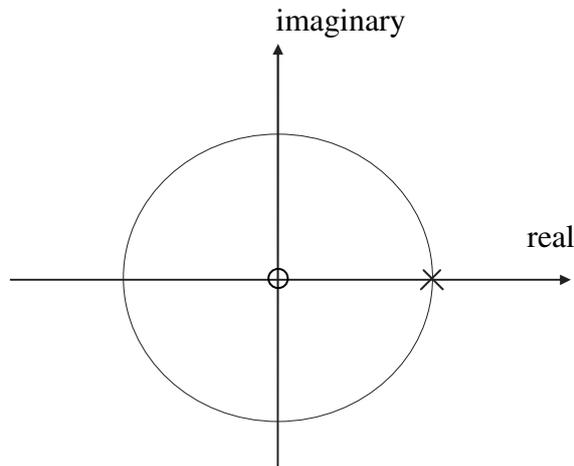
Figure 5.3 z-plane

Example 5.3: From the z -transform pair table, we know the unit step pair as

$$u[n] \leftrightarrow \frac{z}{z-1}$$

The z-transform of the unit step has one zero at origin as $X(z)\big|_{z=0} = 0$ and one pole z =1 as $X(z)\big|_{z=1} \to \infty$, shown in Figure 5.3 in which the pole is represented by a cross and the zero is represented by a circle.

Example 5.4: Find zeros and poles for a z-transform

$$X(z) = \frac{z^2(z-1.2)(z+1)}{(z-0.5+j0.7)(z-0.5-j0.7)(z-0.8)} .$$

Re-write it as

$$X(z) = \frac{(z-0)(z-0)(z-1.2)(z-(-1))}{\{z-(0.5-j0.7)\}\{z-(0.5+j0.7)\}(z-0.8)}$$

It can be obtained: 4 zeros: 0,0, 1.2, -1; and 3 poles: (0.5-j0.7), (0.5+j0.7), 0.8.

They are plotted in Figure 5.4.

Table 5.2 z-Transform pairs

Signal $x[n]$	Z-Transform $X(z)$
$\delta[n]$	1
$u[n]$	$\dfrac{z}{z-1}$
$r[n]$	$\dfrac{z}{(z-1)^2}$
$a^n u[n]$	$\dfrac{z}{z-a}$
$(1-a^n)u[n]$	$\dfrac{z(1-a)}{(z-a)(z-1)}$
$\cos(n\Omega_0)u[n]$	$\dfrac{z(z-\cos\Omega_0)}{z^2-2z\cos\Omega_0+1}$
$\sin(n\Omega_0)u[n]$	$\dfrac{z\sin\Omega_0}{z^2-2z\cos\Omega_0+1}$
$a^n\sin(n\Omega_0)u[n]$	$\dfrac{az\sin\Omega_0}{z^2-2az\cos\Omega_0+a^2}$
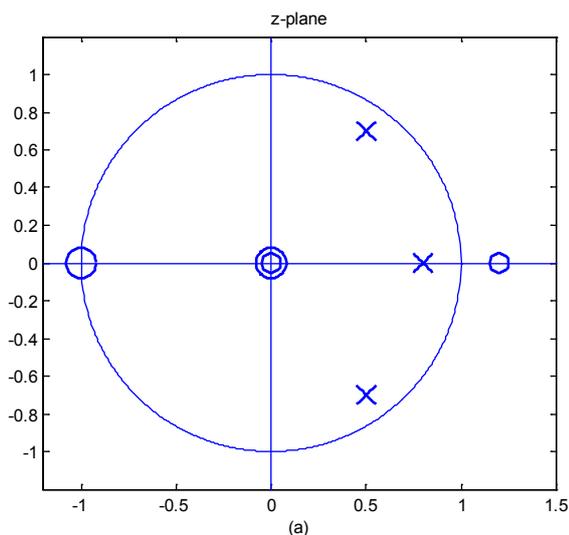
Figure 5.4 Zeros and poles of the signal

Example 5.5. Find the zeros and poles for a z-transform $X(z) = (z^5 - 1)(z^2 + 1)$

Note $z^5 - 1 = 0$ has 5 complex roots.

Since the rule for calculating roots for complex numbers is

$$\left(\exp(j\theta)\right)^{1/n} = \exp(j\frac{\theta + 2k\pi}{n}) \qquad k = 0,1,...,n-1$$

Therefore, the 5 roots are

$$(1)^{1/5} = \left(\exp(j0)\right)^{1/5} = \exp(j\frac{2k\pi}{5}) \qquad k = 0,1,...,4$$

In the z-transform, there are 7 zeros: $1, \exp(j\frac{2\pi}{5})$ $\exp(j\frac{4\pi}{5})$ $\exp(j\frac{6\pi}{5})$ $\exp(j\frac{8\pi}{5})$, j,-j

There is no pole in this case as denominator is 1. See Figure 5.5.
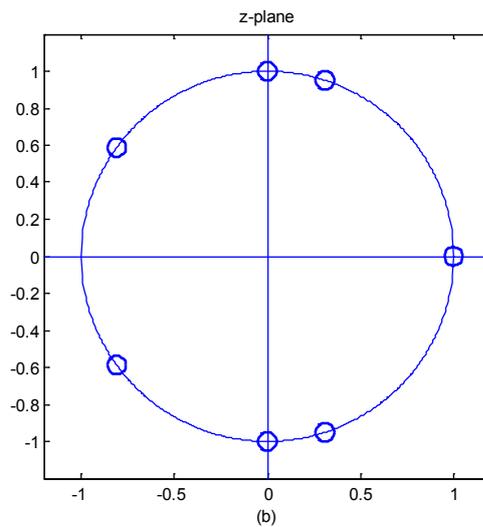
Download free eBooks at bookboon.com

Figure 5.5 Zeros of the z transform

5.6   Stability of a system

The term stability indicates the self-recovery capability of a system after disturbance is applied. If the impulse response decays with the time, the system has the capability of returning it original calm status. Therefore, the stability can be judged by whether the impulse response of the system decays to zero as $n \to \infty$. The following will be shown that the location of poles gives information about the stability of a system.

Consider a simple transfer function with one pole $z = \alpha$:

$$H(z) = \frac{Y(z)}{X(z)} = \frac{1}{z - \alpha} \qquad\qquad (5.5)$$

where α is a real constant. From the above,

$$zY(z) - \alpha Y(z) = X(z)$$

$$Y(z) - \alpha z^{-1}Y(z) = z^{-1}X(z)$$

Apply the inverse z-transform to both sides, yielding the difference equation

$$y[n] = \alpha y[n-1] + x[n-1] \qquad\qquad (5.6)$$

and by specifying an input as a unit impulse it can be known,

$$h[n] = \alpha h[n-1] + \delta[n-1] \qquad\qquad (5.7)$$

The impulse response function can be evaluated as

71

$$h[n] = 0, 1, \alpha, \alpha^2, \alpha^3, \alpha^4 \ldots \tag{5.8}$$

If the impulse response decays with time then the system will return to its initial state before disturbance. Therefore the stability requires $|\alpha| < 1$ in Eq.(5.8) as when $n \to \infty$ makes $h[\infty] \to 0$. i.e. in the z-plane, the real pole $z = \alpha$ must lie on the real axis and inside the unit circle.

Consider a transfer function with conjugate imaginary poles:

$$H(z) = \frac{Y(z)}{X(z)} = \frac{1}{(z - j\alpha)(z + j\alpha)} \tag{5.9}$$

Rearranging,

$$z^2 Y(z) + \alpha^2 Y(z) = X(z)$$

Or

$$Y(z) + \alpha^2 z^{-2} Y(z) = z^{-2} X(z)$$

Taking the inverse z-transform of both sides yields the difference equation

$$y[n] + \alpha^2 y[n-2] = x[n-2]$$

or

$$y[n] = -\alpha^2 y[n-2] + x[n-2] \tag{5.10}$$

From the above we use $h[n] = -\alpha^2 h[n-2] + \delta[n-2]$ to evaluate its impulse response, it can be obtained as

$$h[n] = 0, 0, 1, 0, -\alpha^2, 0, \alpha^4, 0, -\alpha^6, 0, \alpha^8, \ldots \tag{5.11}$$

It is clear that the stability condition also requires $|\alpha| < 1$, in order to achieve that when $n \to \infty$, $h[\infty] \to 0$.

In the case of expression in polar co-ordinates, consider a conjugate pole pair,

$$H(z) = \frac{Y(z)}{X(z)} = \frac{1}{(z - r\exp(j\theta))(z - r\exp(-j\theta))} = \frac{1}{z^2 - 2rz\cos\theta + r^2} \tag{5.12}$$

where r is the radius, θ is the angle on the complex plane. The difference equation can be obtained in the same way as

$$y[n] = 2r\cos\theta\ y[n-1] - r^2 y[n-2] + x[n-2].$$

or

$$h[n] = 2r\cos\theta\ h[n-1] - r^2 h[n-2] + \delta[n-2]$$

By evaluating the impulse response, assuming the system is causal, i.e. $h[n] = 0$ when $n < 0$

$$h[0] = 0,\ \ h[1] = 0,\ \ h[2] = 1,$$
$$h[3] = 2r\cos\theta\ ,$$

$$h[4] = (2r\cos\theta)^2 - r^2$$
$$h[5] = 2r\cos\theta\big((2r\cos\theta)^2 - r^2\big) - r^2 2r\cos\theta = 2r\cos\theta\big(((2r\cos\theta)^2 - r^2) - r^2\big)$$
$$\ldots$$

It can be found that the stable condition is he modulus |r|<1, i.e. the poles are required inside the unit circle.

Example 5.6: The transfer function of a digital system is given by

$$H(z) = \frac{z^2(z-1)(z^2+1)}{(z+0.8)(z^2-1.38593z+0.9604)(z^2+1.64545z+0.9025)},$$, find the zeros and poles.

Re-writing in polar form as:

$$H(z) = \frac{z^2(z-1)(z-\exp\left(j\frac{\pi}{2}\right)j)(z-\exp\left(-j\frac{\pi}{2}\right))}{(z+0.8)\left(z-0.98\exp\left(j\frac{\pi}{4}\right)\right)\left(z-0.98\exp\left(-j\frac{\pi}{4}\right)\right)\left(z-0.95\exp\left(j\frac{5}{6}\pi\right)\right)\left(z-0.95\exp\left(-j\frac{5}{6}\pi\right)\right)}$$

The zeros and poles are easily to be found and described as:

5 zeros: $0,\ 0\ ,1,\ \exp\left(\pm j\frac{\pi}{2}\right)$, and 5 poles: $-0.8,\ 0.98\exp\left(\pm j\frac{\pi}{4}\right),\ 0.95\exp\left(\pm j\frac{5}{6}\pi\right)$. .

Any zeros or poles at the origin only produce a time advance or delay. A minimum-delay system requires:

The number of poles = the number of zeros

If zeros are more than poles, a system becomes non-causal. Some poles are needed to be added to change into causal.

Example 5.7: If a transfer function $H(z) = \dfrac{1}{z^2 - 2rz\cos\theta + r^2} = \dfrac{z^{-2}}{1 - 2rz^{-1}\cos\theta + r^2z^{-2}}$, ,

It can be found

$$y[n] = 2r\cos\theta\ y[n-1] - r^2 y[n-2] + x[n-2].$$

or

$$h[n] = 2r\cos\theta\ h[n-1] - r^2 h[n-2] + \delta[n-2]$$

Assuming the system is causal, $h[n] = 0$ when $n < 0$

$$h[0] = 2r\cos\theta\ h[-1] - r^2 h[-2] + \delta[-2] = 0$$
$$h[1] = 2r\cos\theta\ h[0] - r^2 h[-1] + \delta[-1] = 0.$$

i.e., from the impulse response it can be seen that in the system, there is time delay of 2 samples. However, alternatively, if the z-transform is given by

$$H(z) = \frac{z^2}{z^2 - 2\varepsilon\ \cos\theta + r^2} = \frac{1}{1 - 2\varepsilon^{-1}\cos\theta + r^2z^{-2}},$$

$$y[n] = 2r\cos\theta \; y[n-1] - r^2 y[n-2] + x[n]$$

By evaluating the impulse response, it can be known that when the numbers of poles and zeros are equal, there is no longer a time delay.

5.7   Evaluation of the Fourier transform in the z-plane

Let z = exp(jW), i.e. as the phase angle W varies, but its modulus $|z| = 1$ keeps on the unit circle and takes values on the circle. From this viewpoint, the Fourier transform is a special z-transform.

Example 5.8 A z-transform function   is given by $H(z) = \dfrac{z - 0.8}{z + 0.8}$,

Its frequency response is $H(\Omega) = \dfrac{\exp(j\Omega) - 0.8}{\exp(j\Omega) + 0.8}$

The magnitude $|H(\Omega)| = \dfrac{\left\{(\cos\Omega - 0.8)^2 + (\sin\Omega)^2\right\}^{1/2}}{\left\{(\cos\Omega + 0.8)^2 + (\sin\Omega)^2\right\}^{1/2}}$

Therefore,

$$|H(0)| = |H(2\pi)| = \frac{1 - 0.8}{1 + 0.8} = 0.1.1111$$

and

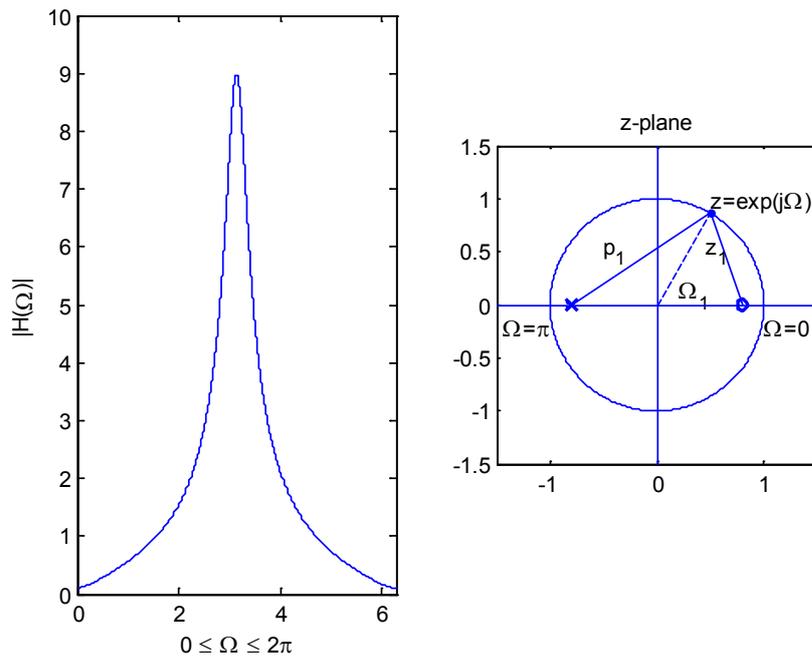$$|H(\pi)| = \left|\frac{-1 - 0.8}{-1 + 0.8}\right| = 9.0$$

Graphically, referring to the following Figure 5.6 (b), let

$$z_1 = |\exp(j\Omega) - 0.8|$$
$$p_1 = |\exp(j\Omega) + 0.8|$$

it can be seen that the quotient $|H(\Omega_1)| = \dfrac{z_1}{p_1}$ reaches the minimum at $\Omega_1 = 0$, and reaches the maximum at $\Omega_1 = \pi$. This has explained the frequency selective property of a high pass.

Figure 5.6 The modulus of frequency response and the illustration of its value as a function of frequency Ω.

Example 5.9: For the transfer function $H(z) = \dfrac{z^2(z-1.2)(z+1)}{(z-0.5+j0.7)(z-0.5-j0.7)(z-0.8)}$,, let $z = \exp(j\Omega)$, the frequency response is

Download free eBooks at bookboon.com

$$H(\Omega) = \frac{(\exp(j\Omega) - 0)(\exp(j\Omega) - 0)(\exp(j\Omega) - 1.2)(\exp(j\Omega) - (-1))}{\{\exp(j\Omega) - (0.5 - j0.7)\}\{\exp(j\Omega) - (0.5 + j0.7)\}(\exp(j\Omega) - 0.8)}$$

The modulus is

$$|H(j\Omega)| = \frac{\left|(\exp(j\Omega) - 1.2)(\exp(j\Omega) - (-1)\right|}{\left|\{\exp(j\Omega) - (0.5 - j0.7)\}\{\exp(j\Omega) - (0.5 + j0.7)\}(\exp(j\Omega) - 0.8)\right|}.$$

The modulus is shown in Figure 5.8. Poles or zeros at the origin do not affect the values of modulus.
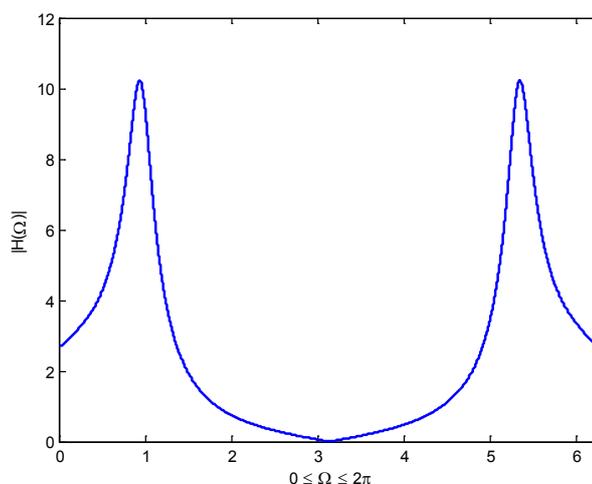

Figure 5.8 Modulus of frequency response

5.8   Characteristics of 1st and 2nd order systems

Digital systems can be classified into first, second and higher order systems according to the number of poles. The same number of zeros is normally chosen to minimise the system delay.

First-order $H_1(z) = \dfrac{z - z_1}{z - p_1}$ (5.13)

Second-order $H_2(z) = \dfrac{(z - z_2)(z - z_3)}{(z - p_2)(z - p_3)}$ (5.14)

Higher-order $H(z) = \dfrac{(z - z_1)(z - z_2)(z - z_3)(z - z_4)...}{(z - p_1)(z - p_2)(z - p_3)(z - p_4)...}$ (5.15)

A higher-order system can be built up by cascading first and second-order subsystems. The frequency selective properties can be controlled by choosing appropriate pole and zero locations.

77

For poles, those close to the unit-circle can produce sharp peaks in the frequency response. Therefore, high gain can be achieved by placing the poles close to the unit circle. Equal number of zeros and poles are normally placed to ensure no system delay or causing non-causality, i.e. the impulse response begins at n=0. Consider the first order system with a zero the origin

The first-order system $\qquad H_1(z) = \dfrac{z}{z - \alpha}$

Its frequency response $\qquad H_1(\Omega) = \dfrac{\exp(j\Omega)}{\exp(j\Omega) - \alpha}$

To achieve low-pass, adopting $0 < \alpha < 1$, i.e. the pole is on the positive axis. The denominator $1 - \alpha$ becomes the smallest, and the maximum gain value is

$$G_{max} = \left| \frac{\exp(0)}{\exp(0) - \alpha} \right| = \frac{1}{1 - \alpha}$$

and the minimum gain   is

$$G_{min} = \left| \frac{\exp(j\pi)}{\exp(j\pi) - \alpha} \right| = \frac{1}{1 + \alpha}$$

To achieve a high-pass, adopting $-1 < \alpha < 0$, i.e. the pole is on negative axis, the denominator $1 + \alpha$ becomes the biggest, and the maximum gain at peak value is

$$G_{max} = \left| \frac{\exp(j\pi)}{\exp(j\pi) - \alpha} \right| = \frac{1}{1 + \alpha} ,$$

and the minimum gain is

$$G_{min} = \left| \frac{\exp(0)}{\exp(0) - \alpha} \right| = \frac{1}{1 - \alpha}$$

When $|\alpha|$ is close to 1, the peak gain gets high, the bandwidth gets more narrow, and the impulse response decays more slowly. The following figures illustrate the above low and high pass filters.
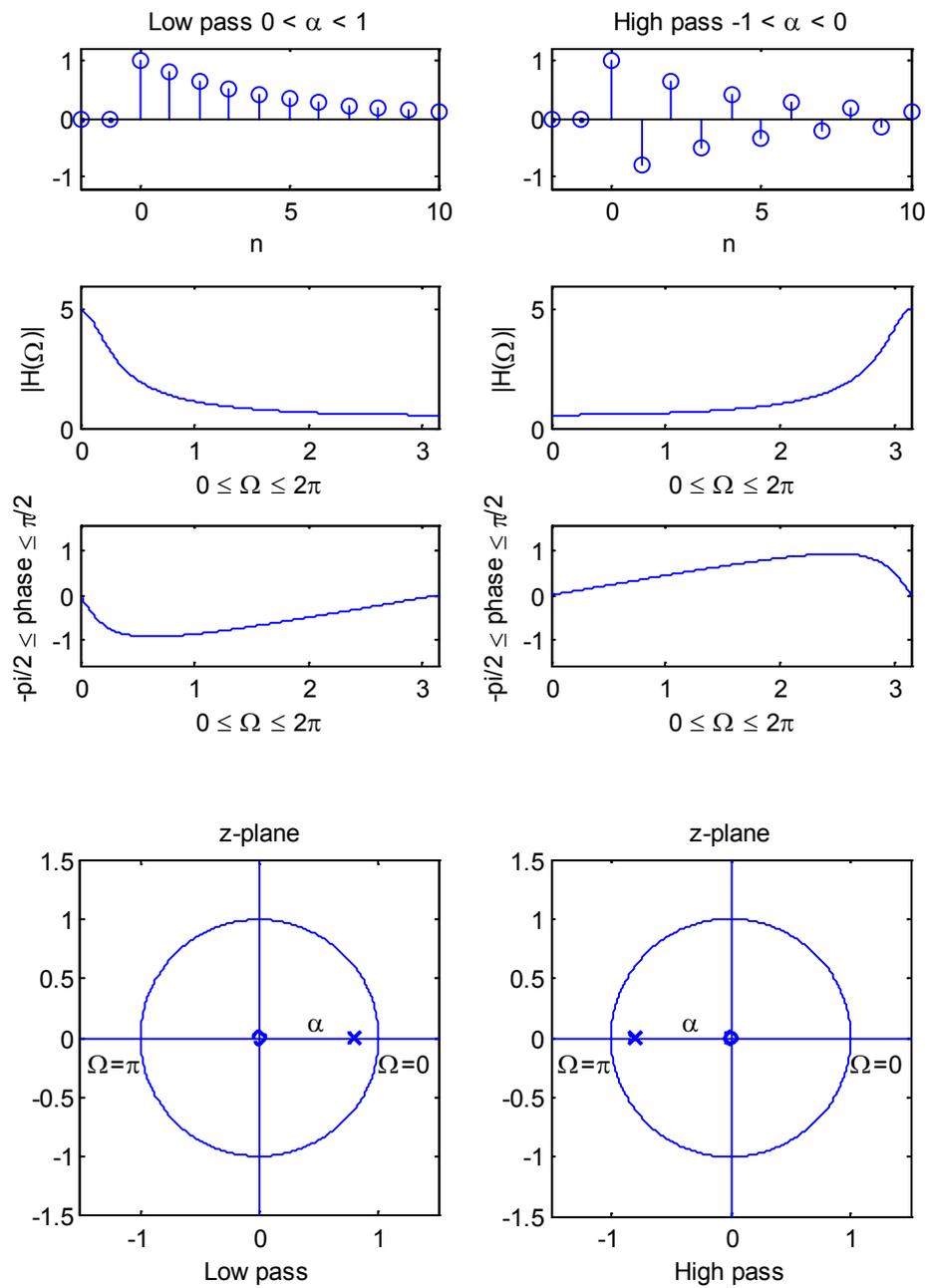
Figure 5.9 The frequency response, poles and zeros of low and high pass filters

For a second-order system, the transfer function in polar form is

$$H_2(z) = \frac{z^2}{(z - r\exp(j\theta))(z - r\exp(-j\theta))} = \frac{z^2}{z^2 - 2r\cos\theta + r^2}$$
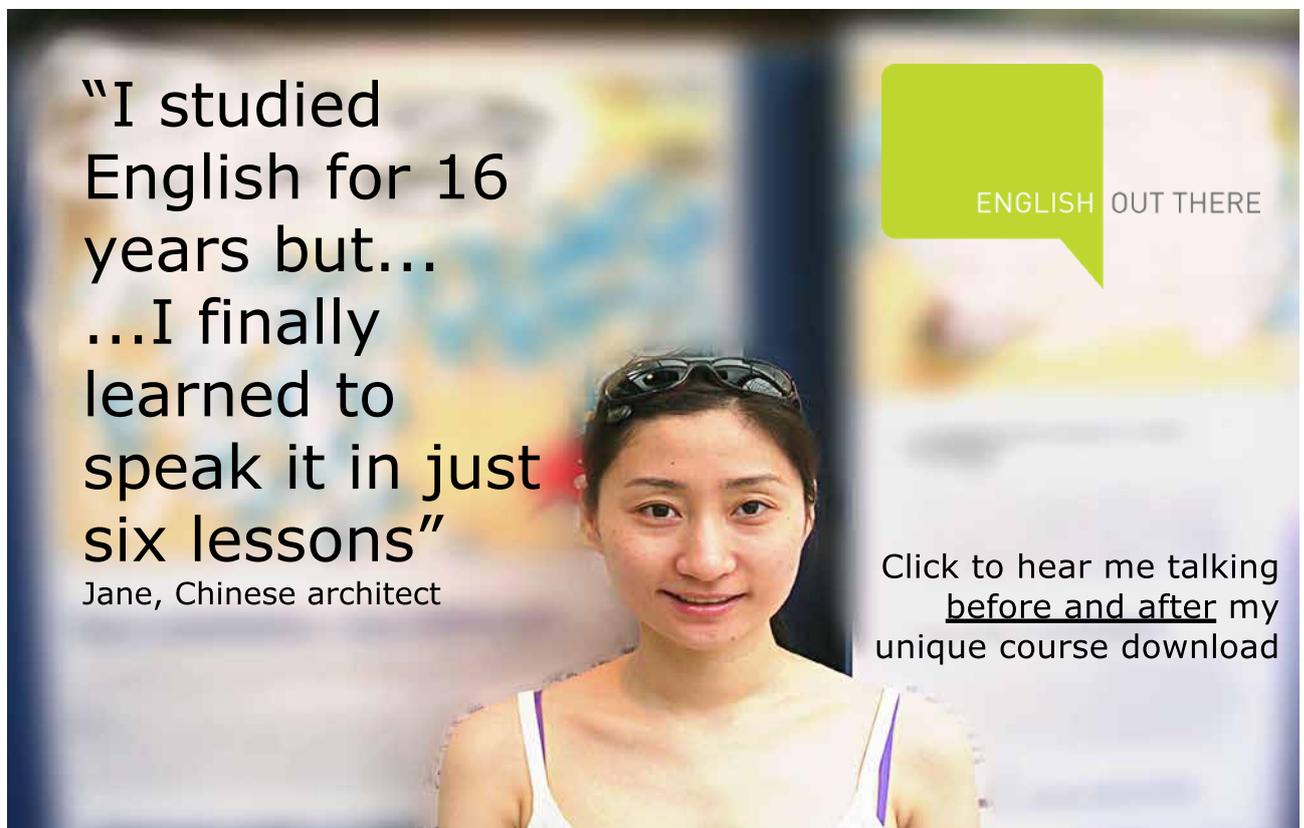
Its frequency response is

$$H_2(\Omega) = \frac{\exp(j2\Omega)}{\exp(j2\Omega) - 2r\cos\theta\exp(j\Omega) + r^2} = \frac{1}{1 - 2r\cos\theta\exp(-j\Omega) + r^2\exp(-j2\Omega)}$$

The gain is

$$|H_2(\Omega)| = \frac{1}{\left\{(1 - 2r\cos\theta\cos\Omega + r^2\cos 2\Omega)^2 + (2r\cos\theta\sin\Omega - r^2\sin 2\Omega)^2\right\}^{1/2}}$$
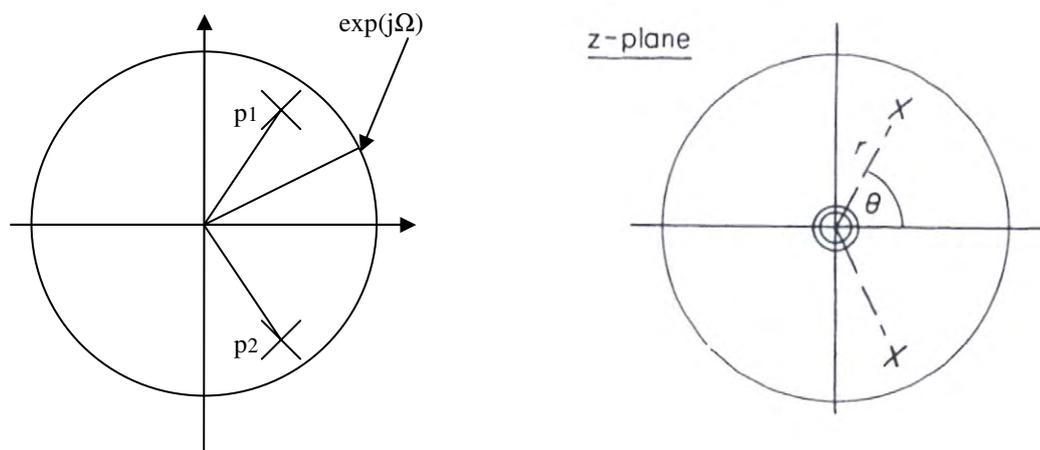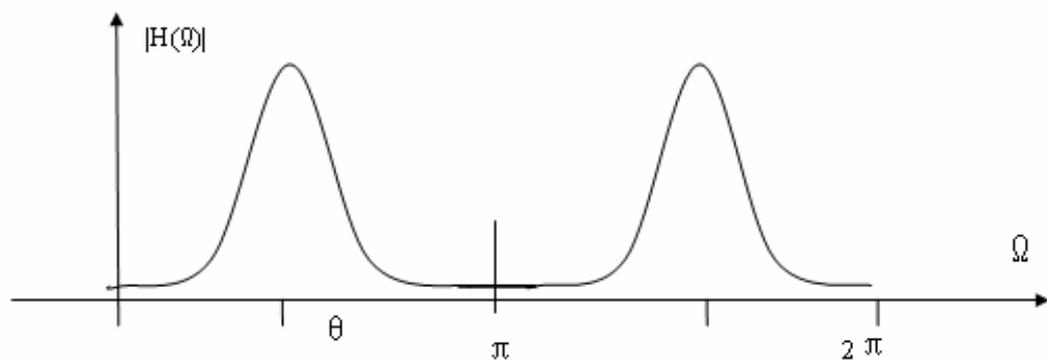
The peak gain occurs at $\Omega = \pm\theta$:

$$G_{max} = |H_2(\theta)| = \frac{1}{\left\{(1 - 2r\cos^2\theta + r^2\cos 2\theta)^2 + (2r\cos\theta\sin\theta - r^2\sin 2\vartheta)^2\right\}^{1/2}}$$

Figure 5.10 Poles of the 2nd order system

Figure 5.11 Modulus of 2nd order system

An example of a higher order transfer function is

$$H(z) = \frac{(z+1)(z^2 - z + 1)}{z(z^2 + 0.9z + 0.81)}.$$

Let $z = \exp(j\Omega)$, the frequency rejections are at $\Omega = \pm\dfrac{\pi}{3},\ \pi$ and frequency passes are at $\Omega = \pm\dfrac{2\pi}{3}$. The poles and zeros, and the frequency response are shown in Figure 5.12.
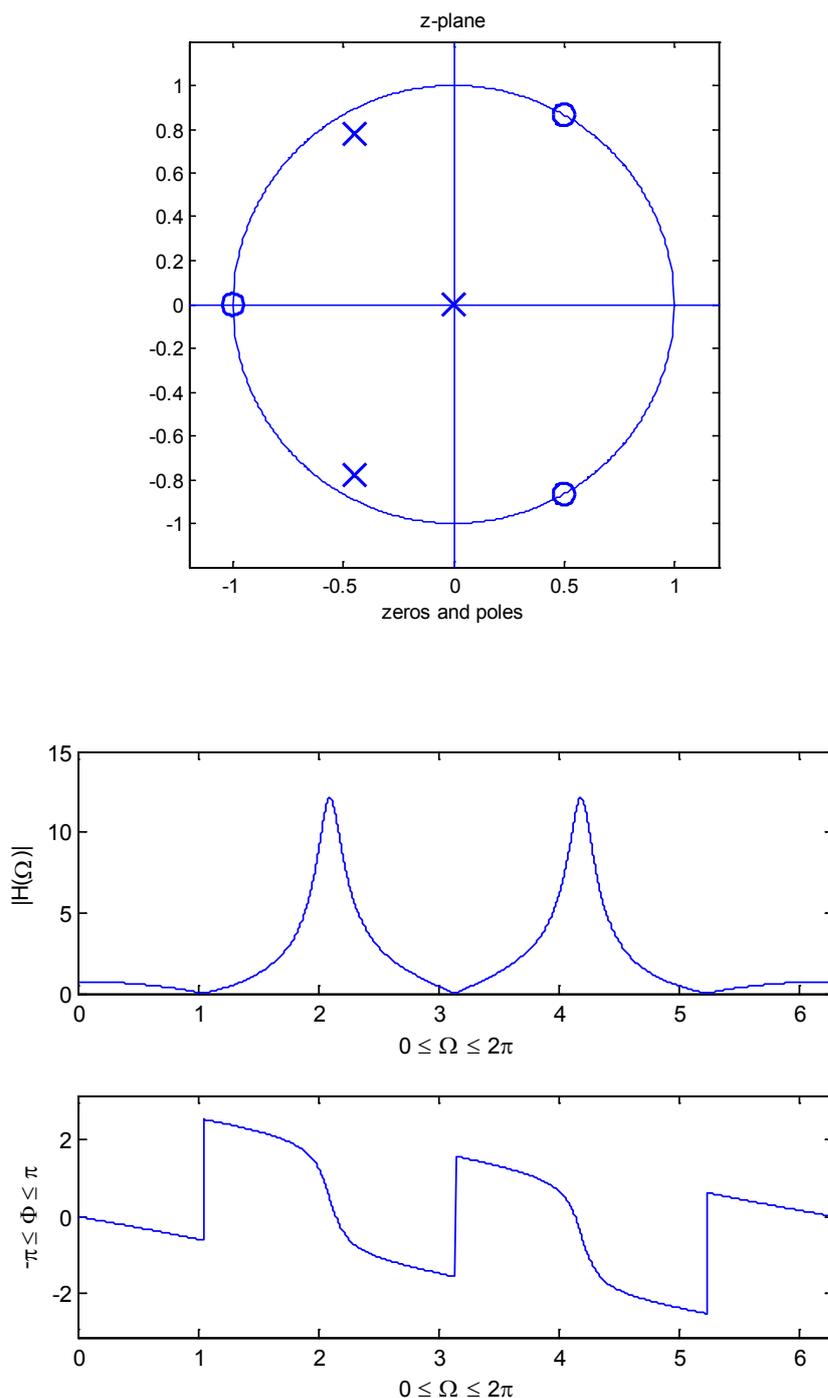
Figure 5.12 Zeros, poles, and frequency response

A more general expression for 2nd order systems, in which there are two zeros and two poles, is given by its transfer function

$$H(z) = \frac{(z - r_1 \exp(j\theta_1))(z - r_1 \exp(-j\theta_1))}{(z - r_2 \exp(j\theta_2))(z - r_2 \exp(-j\theta_2))} = \frac{z^2 - 2r_1 z \cos\theta_1 + r_1^2}{z^2 - 2r_2 z \cos\theta_2 + r_2^2}$$

The frequency response is given by

$$H(\Omega) = \frac{\exp(j2\Omega) - 2r_1 \cos\theta_1 \exp(j\Omega) + r_1^2}{\exp(j2\Omega) - 2r_2 \cos\theta_2 \exp(j\Omega) + r_2^2}$$

where if $r_1 = 1$ there will be a frequency rejection at $\Omega = \theta_1$.

The gain is

$$|H(\Omega)| = \frac{\left\{\left(1 - 2r_1\cos\theta_1\cos\Omega + r_1^2\cos 2\Omega\right)^2 + \left(2r_1\cos\theta_1\sin\Omega - r_1^2\sin 2\Omega\right)^2\right\}^{1/2}}{\left\{\left(1 - 2r_2\cos\theta_2\cos\Omega + r_2^2\cos 2\Omega\right)^2 + \left(2r_2\cos\theta_2\sin\Omega - r_2^2\sin 2\Omega\right)^2\right\}^{1/2}}$$

A summary if its frequency properties in relation to the zeros and poles is listed in Table 5.3.

Table 5.3 Summary of properties of 2nd order systems

	System function	About Zeros	About Poles
Transfer function	$H(z)$	Zeros: When z =Zero, H(z) = 0	Poles: When z =Pole, H(z) = ∞
Frequency response	$|H(\Omega)|$	When $\Omega = \theta_1$ $|H|$ = min If $r_1 = 1$ and $\Omega = \theta_1$, $|H|=0$ (reject frequency)	When $\Omega = \theta_2$, $|H|$ = max (peak gain)

Problems

Q5.1 Find the Z-transform for a digital signal given by

$$x[n] = [1, \ 0.35, \ (0.35)^2 \ ,(0.35)^3, \ ...]$$
$$\uparrow$$

and find the Z-transform if the signal has been changed to

$$x[n] = [1, \ 0.35, \ (0.35)^2 \ ,(0.35)^3, \ ...]$$
$$\uparrow$$

Q5.2 A digital signal is described as

$$x[n] = \begin{cases} 0 & for & n < 0 \\ (0.9)^n & for & n \geq 0 \end{cases}$$

Find the sum $\sum_{n=0}^{\infty} x[n]$ and the Z-transform of the signal.

Q5.3 Expend the following z-transform functions as power series in z^{-1}, and find their first five sample values (starting at n=1):

a) $X(z) = \dfrac{1}{z - 0.6}$

b) $X(z) = \dfrac{z}{z + 1.2}$

c) $X(z) = \dfrac{z+1}{z-1}$

d) $X(z) = \dfrac{0.5z}{z^2 - z + 0.5}$

e) $X(z) = \dfrac{z - 0.5}{z(z - 0.8)(z - 1)}$

Q5.4 Find the transfer function H(z) and frequency response H(Ω) of an LTI system whose impulse response is defined by:

$$h[n] = 0.9h[n-1] - 0.81h[n-2] + \delta[n] - \delta[n-1] + \delta[n-2].$$

Q5.5 Find the zeros and poles of the following transfer functions and identify their stability and causality:

a) $H(z) = \dfrac{z^2 - z - 2}{z^2 - 1.3z + 0.4}$

b) $H(z) = \dfrac{z^2 + 1.5z + 0.9}{z^2 - 1.5z + 1.1}$

c) $H(z) = \dfrac{z^2 - z + 1}{z^2 + 1}$

d) $H(z) = \dfrac{z^3 - z^2 + z - 1}{z^2 - 0.25}$

e) $H(z) = \dfrac{z^9 - 1}{z^8(z - 1)}$

f) $H(z) = \dfrac{z^5 - 2}{z^{10} - 0.8}$

Q5.6 Find the transfer function $H(z)$ and frequency response $H(\Omega)$ of a system whose impulse response is defined by:

$$h[n] = h[n-1] - 0.9h[n-2] + \delta[n] + \delta[n-2]$$

Q5.7 A digital system is described as

$$y[n] - \alpha y[n-1] + \alpha^2 y[n-2] = 2x[n]$$

By considering the pole locations of the associated transfer function, determine the range of the real number, α, for which the system is stable.

Q5.8 A digital system is described by:

$$y[n] = y[n-1] - \alpha y[n-2] + x[n] + \beta x[n-2].$$

By considering the pole locations of the associated transfer function, determine the ranges of the real numbers α and β for which the system can be stable and find the poles and zeros.

Q5.9 A digital filter has the transfer function:

$$H(z) = \frac{z}{z + \alpha}$$

Describe the frequency characteristic with reference to the range of values of α for which the system is stable.

Q5.10 A digital system has the transfer function:

$$H(z) = \frac{z^2}{z^2 - 0.8}$$

Sketch the magnitude of its frequency response $\left|H(\Omega)\right|$ over $0 < \Omega < \pi$.

Q5.11 sketch the magnitude and phase for the frequency responses of filters:

$$H(z) = \frac{z}{z^2 - 1.36z + 0.922}$$

a)

b) $$H(z) = \frac{z^4}{(z^2 - z + 0.8)(z^2 - 1.36z + 0.922)}$$

Q5.12 Find the pass and reject frequencies for the filter given by the transfer function:

$$H(z) = \frac{(z-1)(z^2 - z + 1)}{z(z^2 + 0.9z + 0.81)}.$$

Q5.13 Find the digital signal x[n] and y[n] given by the Z-transforms

$$X(z) = \frac{z}{z - 0.25}, \text{ and } Y(z) = \frac{z + 0.75}{z - 0.25}.$$

Q5.14 Using the smallest possible number of z-plane poles and zeros, design a filter with the following frequency performance:

         Complete rejection at $\Omega = 0$
         Complete rejection at $\Omega = \pi/3$
         A pass-band at $\Omega = 2\pi/3$

The poles are placed at radius $r = 0.9$ in the z-plane and there should be no unnecessary delay in the output signal. Determine the transfer function $H(z)$ for the filter.

Q5.15 Find and sketch the poles and zeros of the following transfer functions. Visualize and sketch the magnitude of the system's frequency response over $0 < \Omega < \pi$.

a) $$H(z) = \frac{z^3 - z^2 + 0.8z - 0.8}{z^2(z + 0.8)}$$

b) $$H(z) = \frac{z^3 - 2z^2 + 2z - 1}{z(z^2 + 0.9z - 0.81)}$$

6   Discrete Fourier Transform

6.1   Definition of discrete Fourier transform

For a digital signal $x[n]$, the discrete Fourier transform (DFT) is defined as

$$X[k] = \sum_{n=0}^{N-1} x[n]\exp\left(-j\frac{2\pi kn}{N}\right)$$

(6.1)

where the DFT $X[k]$ is a discrete periodic function of period N. Therefore one period of distinct values are only taken at $k = 0,1,2,...,N-1$.

Note that the DFT Eq.(6.1) only has defined the transform over $0 \le n \le N-1$, otherwise not known or not cared. This is different from Fourier series in which the signal is strictly periodic or the discrete version of Fourier transform in which the signal is non-periodic but defined over $-\infty < n < \infty$. The comparison is made in Table 6.1. The DFT matches most of the practical cases in which only limited record is available from a certain measurement.

The inverse discrete Fourier transform (IDFT) is

$$x[n] = \frac{1}{N}\sum_{k=0}^{N-1} X[k]\exp\left(j\frac{2\pi kn}{N}\right) \qquad\qquad n = 0,1,2,...,N-1$$

(6.2)

where *x[n]* is a periodic function of period N. Distinct values can be taken from one period at $n = 0,1,2,...,N-1$. For this reason, we can assume the original signal x[n] in the above DFT formula has been extended to a signal of periodic of N. i.e., the DFT considers a non-periodic signal *x[n]* to be periodic only for the purpose of mathematic convenience. Otherwise, the summation in the DFT formula is not to run for N samples $0 \le n \le N-1$ but the whole axis $-\infty < n < \infty$ and to obtain a continuous spectrum $X(\Omega)$. Table 6.1 has listed the definitions of Fourier series, discrete version of Fourier Transform and DFT for comparison.

Table 6.1 Comparison of 3 different transforms

Signal type	Transform	Forward	Inverse
Periodic	Fourier Series	$a_k = \frac{1}{N}\sum_{n=0}^{N-1} x[n]\exp\left(-j\frac{2\pi\,kn}{N}\right)$	$x[n] = \sum_{k=0}^{N-1} a_k\exp\left(j\frac{2\pi\,kn}{N}\right)$
Non-periodic	Discrete version of Fourier Transform	$X(\Omega) = \sum_{n=-\infty}^{\infty} x[n]\exp(-j\Omega n)$	$x[n] = \frac{1}{2\pi}\int_0^{2\pi} X(\Omega)\exp(-j\Omega n)d\Omega$
Length N	Discrete Fourier Transform	$X[k] = \sum_{n=0}^{N-1} x[n]\exp\left(-j\frac{2\pi kn}{N}\right)$	$x[n] = \frac{1}{N}\sum_{n=0}^{N-1} X[k]\exp\left(j\frac{2\pi kn}{N}\right)$

In the following Figure 6.2, the difference between the DFT and discrete version of FT are compared. The upper left is a non-periodic signal with N samples in $0 \le n \le N-1$ in which zeros are given to all outside the N records. The upper right is its discrete version of Fourier transform which is a continuous function. The lower left is the signal in which the N samples are regarded as one period and the record has been extended to the whole axis $-\infty < n < \infty$. Therefore, like Fourier series, its periodic discrete spectrum is shown in the lower right figure.
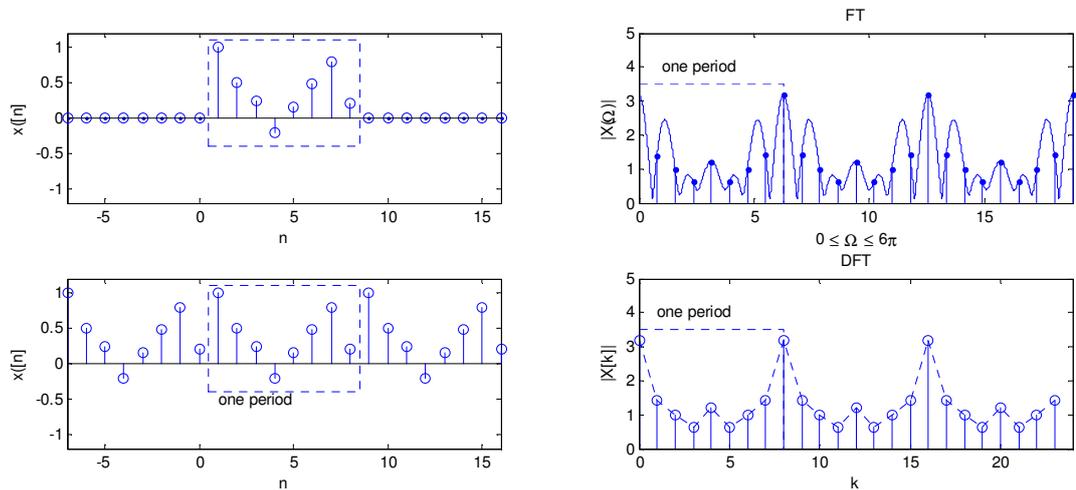

Figure 6.2 The discrete version of FT for a non-periodic signal and DFT for a periodic signal.

In essence, applying the DFT is to decompose a periodic signal to a series of cosine and sine functions represented by

$$\exp\left(-j\frac{2\pi kn}{N}\right) = \cos\frac{2\pi kn}{N} - j\sin\frac{2\pi kn}{N}$$

where k is the frequency of the sinusoidal function which runs through all possibilities from 0 (direct current) to N-1. The following figure shows the first few sinusoidal components.
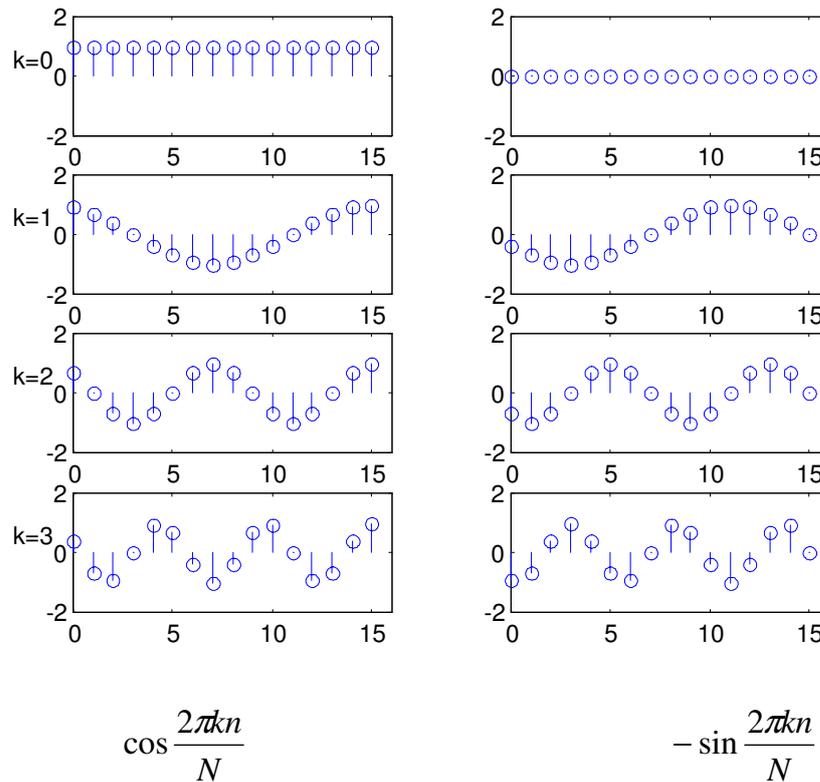

$$\cos\frac{2\pi kn}{N} \qquad\qquad\qquad -\sin\frac{2\pi kn}{N}$$

Figure 6.3 Decomposition of a periodic digital into cosine and sine waves.

6.2   Properties of DFT

1. Periodicity

In the time domain,

$$x[n \pm rN] = x[n]$$

(6.3)

and in the frequency domain

$$X[k \pm rN] = X[k]$$

(6.4)

where r is an arbitrary integer and N is the period. This property says that the shape of the signal stays the same when it is shifted to left or right by integer number of N samples.

2. Linearity

If $\qquad$ $x_1[n] \leftrightarrow X_1[k]$ and $x_2[n] \leftrightarrow X_2[k]$

then

$$Ax_1[n] + Bx_2[n] \leftrightarrow AX_1[k] + BX_2[k] \qquad (6.5)$$

where $\leftrightarrow$ represents the pair of DFT and IDFT, and A and B are constants. This property includes an equal magnification rule, and a superposition rule between the input and output.

3. Time-shifting

If $\qquad$ $x[n] \leftrightarrow X[k]$

then

$$x[n - n_0] \leftrightarrow X[k]\exp\left(-j\frac{2\pi k n_0}{N}\right) \qquad (6.6)$$

The time shifting will cause a change of spectrum in phase, not in the magnitude, because $\left|\exp\left(-j\frac{2\pi k n_0}{N}\right)\right| = 1$..

4. Convolution

If $\qquad$ $x_1[n] \leftrightarrow X_1[k]$ and $x_2[n] \leftrightarrow X_2[k]$

then

$$\sum_{m=0}^{N-1} x_1[n]x_2[m-n] \leftrightarrow X_1[k]X_2[k] \qquad (6.7)$$

The relationship of convolution between two signals in time domain can be simplified to a multiplication in the frequency domain. In the formula, the convolution is defined on one period.

5. Modulation

If $\qquad$ $x_1[n] \leftrightarrow X_1[k]$ and $x_2[n] \leftrightarrow X_2[k]$

then

$$x_1[n]x_2[n] \leftrightarrow \sum_{m=0}^{N-1} X_1[m]X_2[k-m] \qquad (6.8)$$

Likewise to the property 4), the relationship of convolution between two spectra in the frequency domain can be simplified to a multiplication in the time domain.

6. Even and odd signals

From the DFT definition

$$X[k] = \sum_{n=0}^{N-1} x[n] \left(\cos\frac{2\pi kn}{N} - j\sin\frac{2\pi kn}{N} \right)$$

where $\cos\dfrac{2\pi kn}{N}$ is an even function, and $\sin\dfrac{2\pi kn}{N}$ is an odd function.

Let

$$\text{Real}(X[k]) = \sum_{n=0}^{N-1} x[n]\cos\frac{2\pi kn}{N}$$

$$\text{Imag}(X[k]) = \sum_{n=0}^{N-1} x[n]\sin\frac{2\pi kn}{N}$$

When x[n] is real signal,

a) if x[n] is an even function,

Im(X[k]) =0   (6.9)

b) if x[n] is an even function,

Re(X[k]) =0   (6.10)

This property can be used to simplify and save the calculation.

7. Conjugation

If x[n] is real, $X[0] = \sum_{n=0}^{N-1} x[n]$ and $X[N/2] = \sum_{n=0}^{N-1} (-1)^n x[n]$ are real coefficients, and the other N-2 are complex coefficients.

$$X[-k] = X^*[k] \;\text{ or }\; X[N-k] = X^*[k] \tag{6.11}$$

$$\left| X[-k] \right| = \left| X[k] \right|, \quad \left| X[N-k] \right| = \left| X[k] \right| \tag{6.12}$$

Only X[0], X[N/2] and X(k), k=1,2,N/2-1 are needed to represent the whole X[k] (k=0,1,2,…,N-1). i.e. there are a total of 2 real and N/2-1 complex coefficients. It can also be proved

$$\left| X[\frac{N}{2} - k] \right| = \left| X[\frac{N}{2} + k] \right|$$

(6.13)

The proof is as below:

$$X[\frac{N}{2} - k] = \sum_{n=0}^{N-1} x[n] \exp\left(-j \frac{2\pi(N/2 - k)n}{N} \right)$$

$$= \sum_{n=0}^{N-1} x[n] \exp\left(-j \frac{2\pi(N/2 - k - N)n}{N} \right)$$

$$= \sum_{n=0}^{N-1} x[n] \exp\left(j \frac{2\pi(N/2 + k)n}{N} \right) = X^*[\frac{N}{2} + k]$$

Therefore, $\left| X[\frac{N}{2} - k] \right| = \left| X[\frac{N}{2} + k] \right|$. This property tells that the modules of the DFT is symmetrical about the vertical line $n = \frac{N}{2}$.

8. Complex signal $x[n]$

If the signal $x[n]$ is complex, there is no spectral symmetry, and all N coefficients are distinct in general.

6.3   The fast Fourier transform (FFT)

James W. Cooley and JohnW. Tukey in 1965 made a revolutionary invention in calculating the DFT (published in J.W.Cooley and J.W. Tukey in Math. Comput., vol. 19, April 1965, pp297-301). In the algorithm known as FFT, redundancy in direct calculating complex DFT due to periodicity in sinusoidal functions has been removed, therefore the computing time has been remarkably reduced. The principle can be explained in the following.

For the DFT $X[k] = \sum_{n=0}^{N-1} x[n] \exp\left(-j \frac{2\pi \ kn}{N} \right)$, let the complex function

$$W_N = \exp\left(-j \frac{2\pi}{N} \right)$$

then

$$W_N^{kn} = \exp\left(-j \frac{2\pi \ kn}{N} \right).$$

If separating $x[n]$ to an eve and an odd sequences

$$X[k] = \sum_{n=0}^{N-1} x[n] \, W_N^{kn} = \sum_{r=0}^{N/2-1} x[2r] \, W_N^{2rk} + \sum_{r=0}^{N/2-1} x[2r+1] \, W_N^{(2r+1)k}$$

$$= \sum_{r=0}^{N/2-1} x[2r] \, W_{N/2}^{rk} + W_N^k \sum_{r=0}^{N/2-1} x[2r+1] \, W_{N/2}^{rk}$$

(6.14)

Eq.(6.14) means a DFT of length N can be equivalent to 2 DFTs of length N/2. As an immediate result, the number of distinct complex numbers can be reduced from N W_N^{kn} to N/2 $W_{N/2}^{rk}$ in the above DFT summation; thus complex multiplications can be greatly reduced in computation. The heart of implementing FFT is to make above division further until each DFT has only 2 samples. A requirement is the length of data N is an integer power of 2.

Therefore, N is chosen to be an integer power of 2, N/2 is even. 2 N/2-point sequences can be decomposed into 2 shorter N/4-point sequences. This decomposition continues until all sequences are 2-point sub-sequences, each of which requires only a simple 2-point DFT. This procedure produces a radix-2 FFT algorithm.

For example: Let N=8, the DFT is

$$X[k] = \sum_{n=0}^{7} x[n] \, W_8^{kn}$$

The number of direct calculation of its DFT will be $8^2 = 64$, approximately. However, it can be divided into 2 length N=4 sequences:

$$X[k] = \sum_{r=0}^{3} x[2r] \ W_4^{rk} + W_8^k \sum_{r=0}^{3} x[2r+1] \ W_4^{rk}$$

Further, they can be divided in to 4 length N=2 sequences:

$$X[k] = \sum_{s=0}^{1} x[2(2s)] \ W_2^{sk} + W_4^k \sum_{r=0}^{1} x[2(2s+1)] \ W_2^{sk}$$
$$+ W_8^k \left(\sum_{s=0}^{1} x[2(2s)+1] \ W_2^{sk} + W_4^k \sum_{s=0}^{1} x[2(2s+1)+1] \ W_2^{sk} \right)$$

Explicitly, from the above,

$$X[k] = x[0] \ W_2^{0k} + x[4] \ W_2^{1k} + W_4^k \left(x[2] \ W_2^{0k} + x[6] W_2^{1k} \right)$$
$$+ W_8^k \left(x[1] \ W_2^{0k} + x[5] \ W_2^{1k} + W_4^k \left(x[3] \ W_2^{0k} + x[7] \ W_2^{1k} \right) \right)$$

where we know $W_2^0 = 1, \quad W_2^1 = -1$, therefore,

$$X[k] = x[0] \ + x[4] \ W_2^{1k} + W_4^k \left(x[2] \ + x[6] W_2^{1k} \right) + W_8^k \left(x[1] \ + x[5] \ W_2^{1k} + W_4^k \left(x[3] \ + x[7] \ W_2^{1k} \right) \right)$$

where only W_4^k and W_8^k are actually complex numbers, there are as many as only 3 + 7 = 12.

In original DFT, there are approximately N^2 multiplications in $\ W_N^{kn} = \exp\left(-j \dfrac{2\pi nk}{N} \right)$ (there are some unities when k or n=0). However, in the FFT algorithm, redundant computation in multiplying W_N^{kn} are reduced by re-arranging samples to shorter sequences to enable multiplication by much fewer distinct $\ W_N^{kn}, W_{N/2}^{kn}, W_{N/4}^{kn}, ... W_2^0 \ and \ W_2^1$ in a butterfly shaped flow chart. Figure 6.4 illustrates N^2 multiplications in a length N=8 DFT.
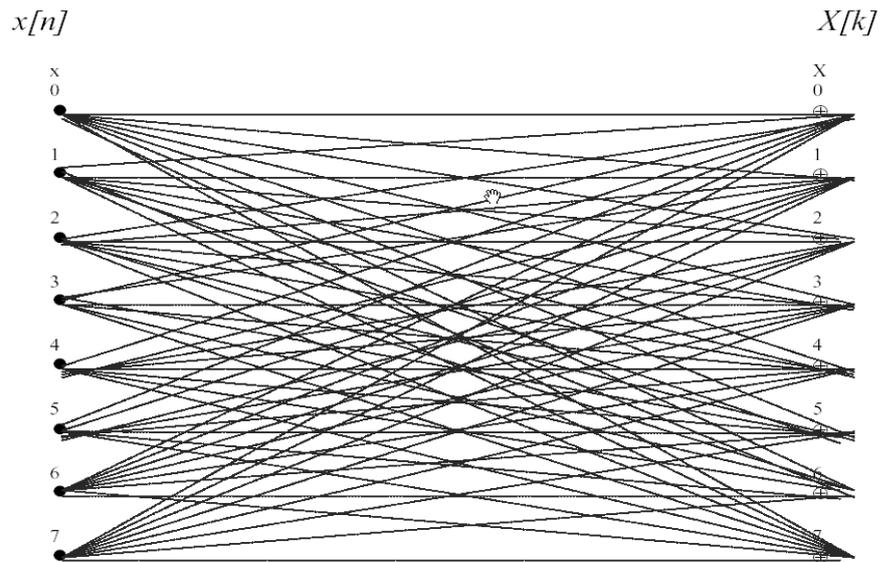
$x[n]$ $X[k]$

Figure 6.4 Number of complex multiplication in DFT

It can be divided into 2 length N=4 DFTs , i.e.

$$X[k] = \sum_{r=0}^{3} x[2r]W_4^{rk} + W_8^{k} \sum_{r=0}^{3} x[2r+1]W_4^{rk}$$

Download free eBooks at bookboon.com

where

$$W_8^k = \exp\left(-j\frac{2\pi k}{8}\right)$$

and

$$W_8^k = \exp\left(-j\frac{2\pi k}{8}\right)$$

In the following Figure 6.5, two boxes represents 2 length N=4 DFTs. The solid lines represent moves and the doted lines represent complex multiplications.
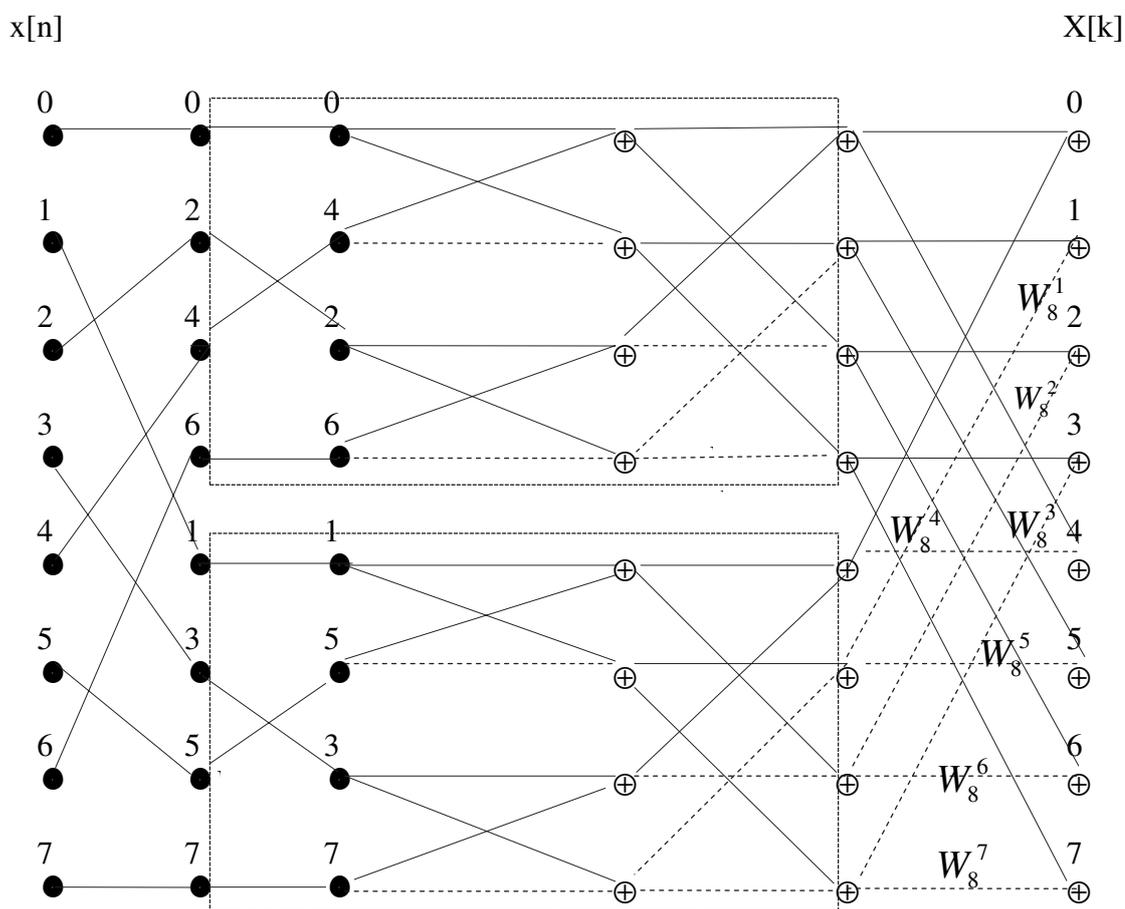

Figure 6.5 1 N=8 point DFT are changed to 2 N=4 point DFTs

In the two boxes, 4 N=2 DFTs can be obtained as:

$$X[k] = \sum_{s=0}^{1} x[2(2s)]W_2^{sk} + W_4^k \sum_{s=0}^{1} x[2(2s+1)]W_2^{sk}$$

$$+ W_8^k \left\{ \sum_{s=0}^{1} x[2(2s)+1]W_2^{sk} + W_4^k \sum_{s=0}^{1} x[2(2s+1)+1]W_2^{sk} \right\}$$

$$W_8^k = \exp\left(-j\frac{2\pi k}{8}\right), \; W_4^k = \exp\left(-j\frac{2\pi k}{4}\right)$$

$$W_2^{sk} = \exp\left(-j\frac{2\pi sk}{2}\right), \quad W_2^0 = 1, W_2^1 = -1$$

Using the following graphical illustration in Figure 6.6, we can see the complex multiplications are only 4+6+7=17 in this case, much fewer than that in the direct calculation of the DFT.
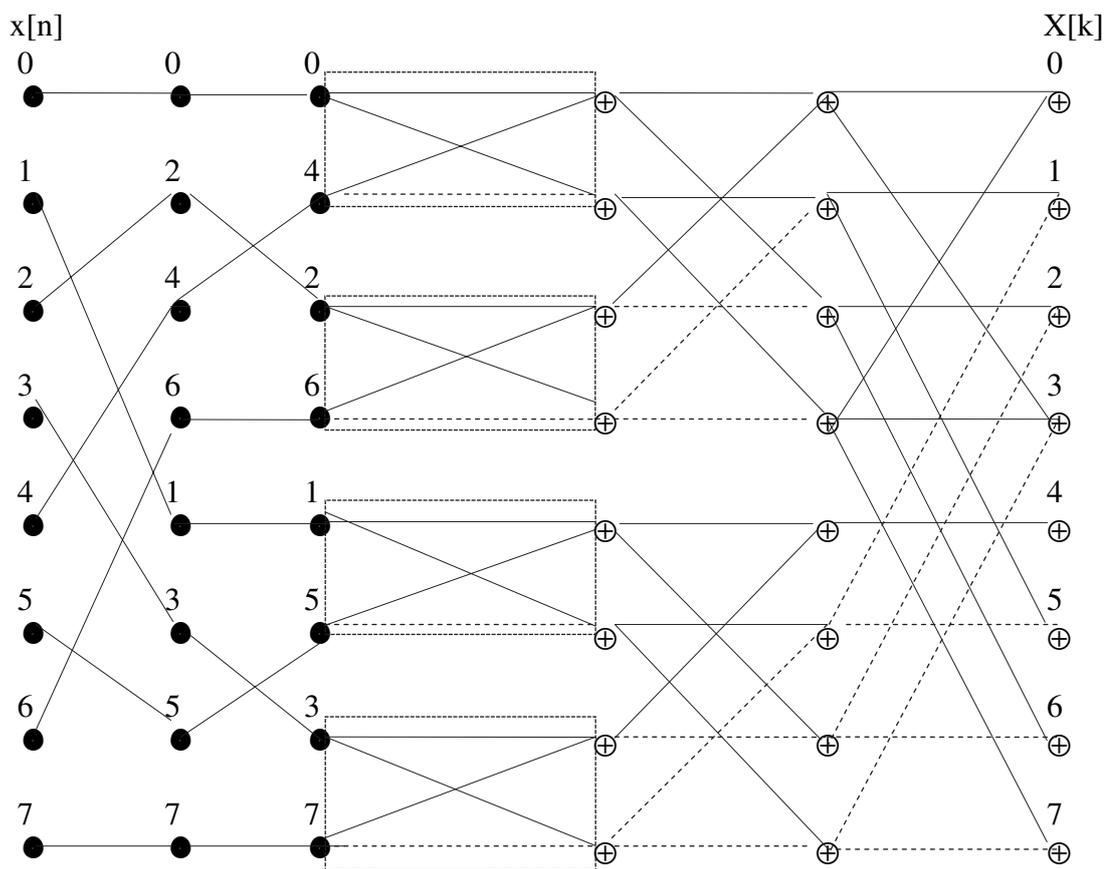

Figure 6.6 2 N=4 point DFT are changed to 4 N=2 point DFTs

It can be proved that the FFT algorithm has saved huge computing time by reducing from N^2 complex multiplications

to $N \log_2 N$, i.e. saved $\dfrac{N}{\log_2 N}$ times of computation. For example, if N=8, $\dfrac{8}{\log_2 8} = \dfrac{8}{3} = 2.67$; if N=1024,

$\dfrac{1024}{\log_2 1024} = \dfrac{1024}{10} = 102.4$. i.e. saved 102 times of multiplications. The longer the data length N, the more time can

be saved relative to the direct calculation of the DFT.

Problems

Q6.1 What are the features of the DFT coefficients X[k] of an N-sample signal which is

- a) Real,
- b) Real and even,
- c) Real and odd, and
- d) Complex?

Q6.2 For the digital sequence

- a) x[n] = [1 -1],
- b) x[n]= [3 -2],
- c) x[n]=[1 -1 0 0],
- d) x[n] = [1 0 0 1],
- e) x[n]=[1 2 1 3].

Calculate the Discrete Fourier Transform (DFT) .

Q6.3 Explain how the Fast Fourier Transform (FFT) algorithm can be faster than direct calculation of the Discrete Fourier Transform (DFT).

Q6.4 Answer the following questions:

- With reference to the Fast Fourier Transform (FFT), why is the length, *N,* normally chosen as an integer power of 2?
- In brief, what is the reason that the FFT algorithm can be faster than direct calculation of the Discrete Fourier Transform (DFT)?
- If the length of a sequence is not yet an integer power of 2, how is it possible to take advantage of the FFT algorithm?

7 Spectral Analysis by DFT

7.4 Digital spectral analysis

Spectral analysis for digital signals is referred to a decomposition of a digital signal into its continent components in the frequency domain. It is a useful technique in many branches of engineering, natural and social sciences, and information technology. The DFT defined in Eq. 6.1 is used to obtain Fourier spectrum for a signal, and the FFT algorithm is widely accepted as the choice in implementing the algorithm because of its fast speed.

The basic assumption behind the digital spectral analysis is that a frequency-domain display is easier to reveal important information which is not apparent in the time domain. Unlike digital filtering, it is primarily investigative, not concerned with changing the original signal. The information obtained through the digital spectral analysis often leads to important insights of associated physical process.

7.5 Spectra of harmonics

Assume the sampling rate is 1024Hz. The maximum frequency at this sampling rate is 512Hz and the minimum frequency in the DFT frequency domain is also 1Hz. Any integer frequency components ranging from 1 to 512 Hz are called harmonics. A signal

$$x[n] = \sin\left(2\pi\,\frac{n}{1024}\right)$$

(7.1)

represents a 1Hz sine wave. The 32th harmonics is
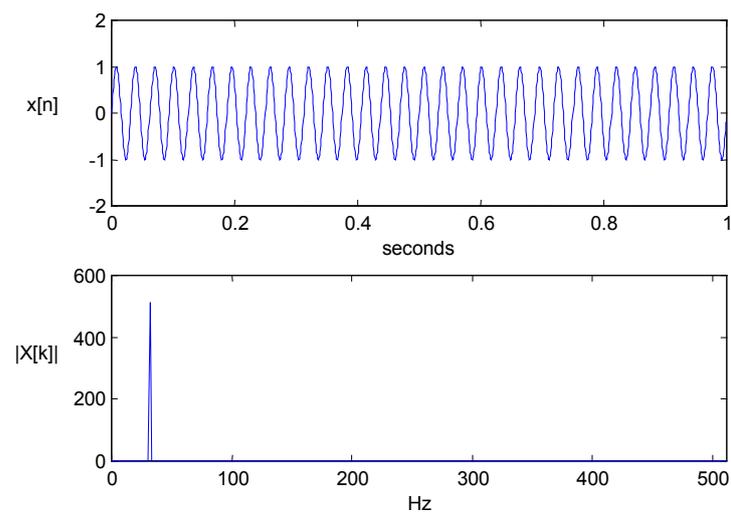
$$x[n] = \sin\left(2\pi\,32\,\frac{n}{1024}\right)$$

(7.2)

which will present a single spectral line in the DFT at k = 32. See Figure 7.1.

If more harmonics are present, like

$$x[n] = \sin\left(2\pi\,32\,\frac{n}{1024}\right) + \sin\left(2\pi\,137\,\frac{n}{1024}\right) + \sin\left(2\pi\,467\,\frac{n}{1024}\right)$$

(7.3)

Three spectral lines will be present, shown in Figure 7.2.

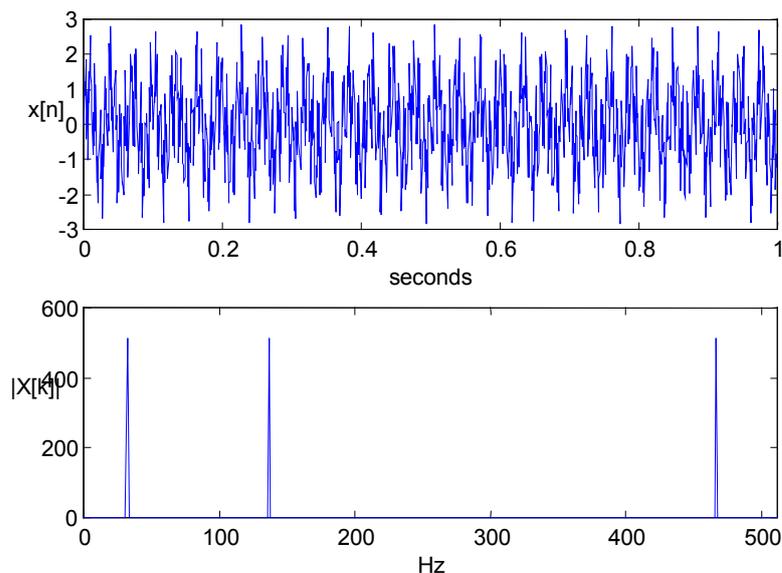**Figure 7.1** Single sine wave and its single spectral line.

Figure 7.2 3 harmonic waves

7.6   Spectral leakage

In the DFT, the signal of length N is being treated as an exact one period of a periodic signal. If a component is an integer multiple or harmonic of the basic frequency, the wave will smoothly continue from one period to the next in the time domain. On contrary, if the component is not an integer harmonic, discontinuity will occur from one period to another in the time domain. This is particularly the case in practice in which a signal usually contains many components of different frequencies. Those frequencies can take any fractional numbers and are rarely exact harmonics of the basic frequency. The spectrum will not appear as a single line but a peak with side-lobs on both sides. This can be explained as that the discontinuity between periods causes a disturbance or modulation in the magnitude and phase of the component, generating a set of new harmonics whose frequencies are close to the main harmonic. For example, a signal with three fractional frequencies of 32.5, 137.5 and 467.5 Hz, i.e.

$$x[n] = \sin\left(\frac{2\pi 32.5 n}{1024}\right) + \sin\left(\frac{2\pi 137.5 n}{1024}\right) + \sin\left(\frac{2\pi 467.5 n}{1024}\right) \tag{7.4}$$

The three components give no longer a single spectral coefficient, but three high coefficients surrounded by side-lobs, which represent the spectral leakage, shown in Figure 7.3. As a result, relative to integer harmonic cases, the leakage reduces the magnitude of the main spectral line, giving an inaccurate indication of the spectral strength. This phenomenon can be improved by windowing, described in the following section.
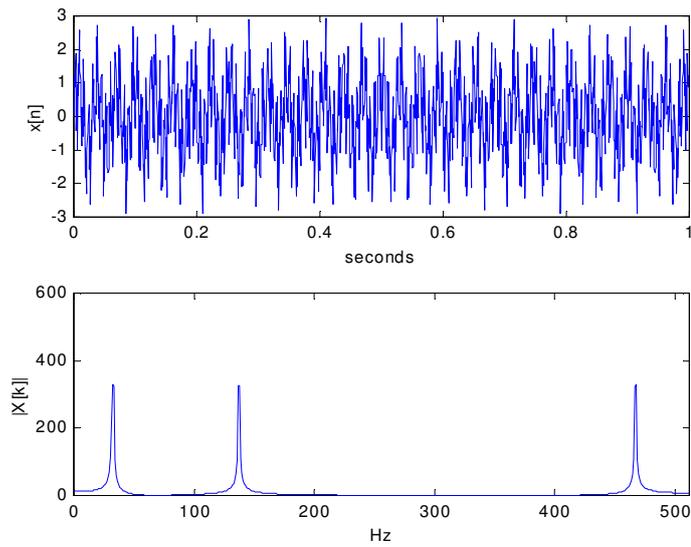
Figure 7.3 Side lobes around main spectral lines.

7.4   Windowing

In the DFT, only a component at an exact harmonic frequency gives rise to a single and well-defined spectral line. In fact, practical digital signals normally contain a majority of fractional frequencies and few of them exact harmonics. This means that the spectral leakage is generally present, and it may lead to inaccuracy in analysis and interpretation. It is therefore common practice to taper two ends of the original signal before applying the DFT, reducing or removing any discontinuities at its two ends. This can be achieved by multiplying the signal with a suitable window function. For a signal $x[n]$, $n = 0,1,2,...N-1$, applying an equal length window function $w[n]$, $n = 0,1,2,...N-1$, a windowed signal is given by

$$x_w[n] = x[n]\,w[n]$$

(7.5)

-   **Rectangular window (No Window)**

$$w[n] = 1 \qquad\qquad 0 \le n < N$$

(7.6)

The windowed signal $x_w[n]$ is not tapered by this rectangular window. The spectral leakage is fully present.

-**Triangular window**

$$w[n] = 1 - \frac{|\,2n - N + 1\,|}{N} \qquad 0 \le n < N$$

(7.7)

The windowed signal $x_w[n]$ will be tapered by the straight slopes of the triangle.

-**Hamming window**

$$w[n] = 0.54 + 0.46\cos\!\left(\frac{(2n - N + 1)\pi}{N}\right) \qquad 0 \le n < N$$

(7.8)

The windowed signal $x_w[n]$ will be tapered by the cosine function.

9.8   Performance of windows

In the time domain, applying a window is to multiply by the window function $w[n]$:

$$x_w[n] = x[n]\,w[n]$$

In the frequency domain, according to the DFT modulation property in Eq.(6.8), two spectra are in convolution:

$$X_w[k] = X[k] * W[k]$$

(7.9)

i.e. a signal multiplied by a window in the time domain is equivalent to a convolution between the spectra of the signal and window in the frequency domain. The shapes of triangular and Hamming windows in the time and frequency domains are shown in Figure 7.4. Therefore, the leakage will be determined by the shape of the window's Fourier spectrum. The rectangular window (no-window) introduces significant side-lobs, which indicate the leakage seriously exists. The triangular window can reduce side-lobs but broadens spectral lines of integer harmonics. The Hamming window slightly broadens spectral lines of integer harmonics, but leakage can be dramatically reduced. Therefore, the Hamming window is a good choice for reducing leakage. In Figure 7.5, the effects of those three windows are illustrated.
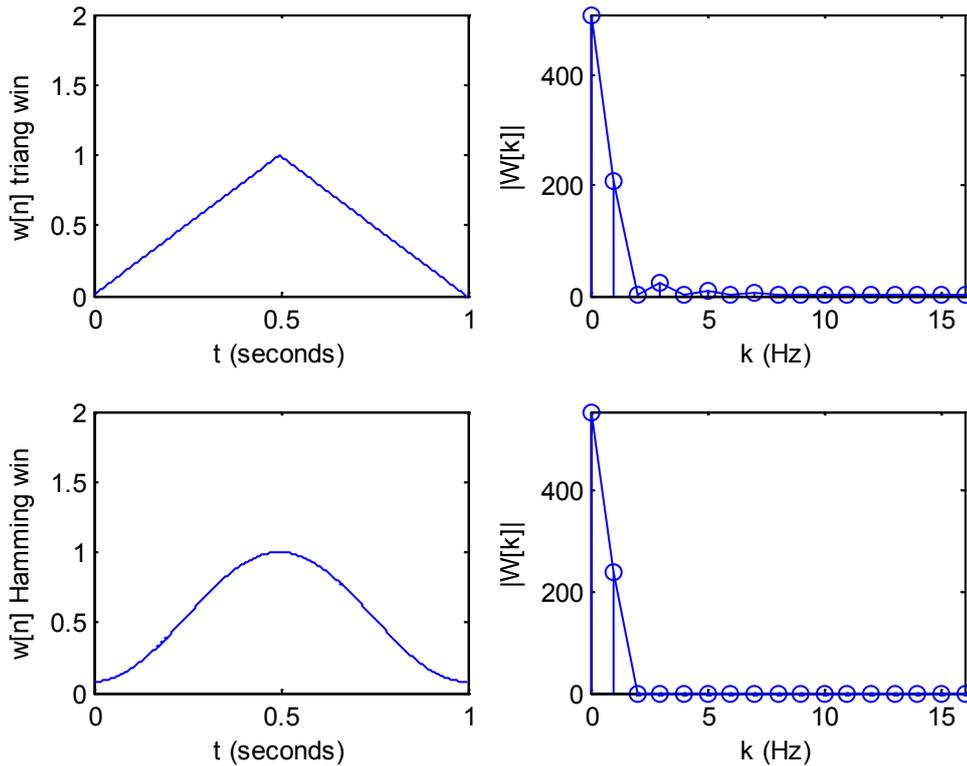
Figure 7.4 The shapes of triangular and Hamming windows in the time and frequency domains
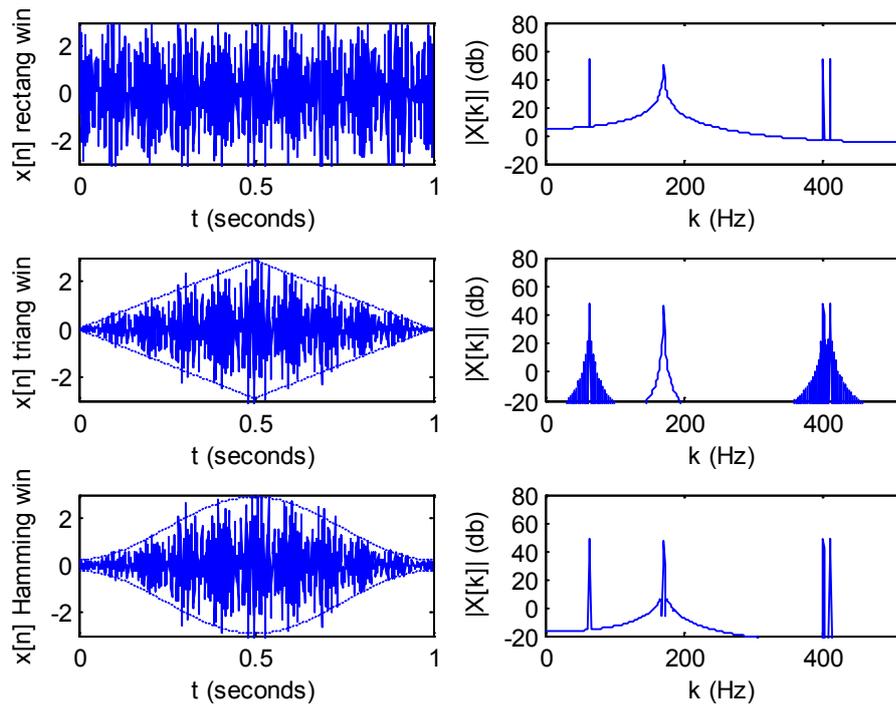
Figure 7.5 Three windows and their effects on spectral lines.

7.6   Applications of digital spectral analysis

- Detection of harmonics of interest from noisy background

A sinusoidal signal may be mixed with strong noise and hard to be observed. It is appropriate to use the Fourier spectrum to detect the sinusoidal signal from the noise. In the frequency domain, a signal may be displayed and extracted easily from noise background. In the following Figure 7.6, the upper one is the signal mixed with noise in the time domain. It is hard to observe what harmonic components are contained inside the signal. The middle curve is the Fourier spectrum where a strong harmonic appears. Therefore, it is easy to remove the recognisable noise by reducing its noise spectrum and to restore the harmonic shown in the lower plot using the inverse DFT.

- Fault detection

The distribution of spectra often characterises the operating condition of mechanical systems, and sudden rising of some components or any other changes in it may indicate an occurrence of fault. Key machinery or even the whole production lines can be monitored and early faults can be reported to prevent from unscheduled shutdown, which causes losses or even disasters. Vibration or acoustic signals are normally picked up and digital spectra are calculated regularly with an appropriate interval. A computer is programmed to make comparison with the historical data to report any abnormality.

- Identification of unknown systems

An unknown system is being deliberately excited by a suitable known signal – often an impulse or a step signal – and its response is being measured, then yielding the frequency properties of the system. See Figure 7.7.

In the time domain, the input-output relationship of a system is in a form of convolution:

$$y[n] = \sum_{k=0}^{N-1} x[k]h[h-k]$$

(7.10)

Note that this is a periodic or circular convolution, different from Eq.(3.16), as the signal $x[n]$ is assumed periodic. Take the DFT for both sides, using the convolution property of DFT:

$$Y[k] = H[k]\ X[k]$$

(7.11)

Therefore, the frequency response function of the system can be identified by

$$H[k] = \frac{Y[k]}{X[k]}$$

(7.12)

In practice, due to noise in measurement, the frequency response function must be obtained by averaging of multiple records of the input and output.
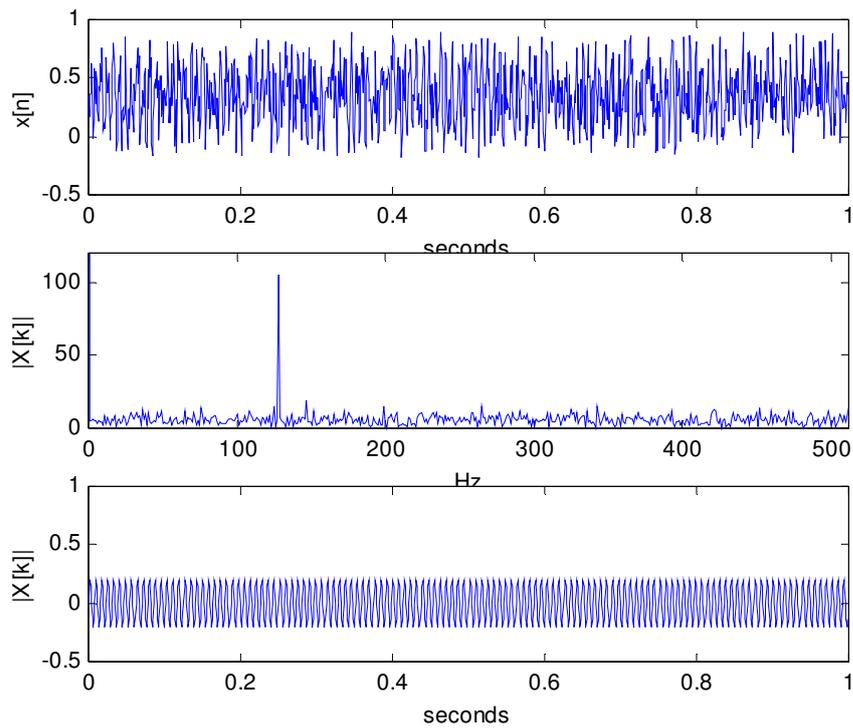
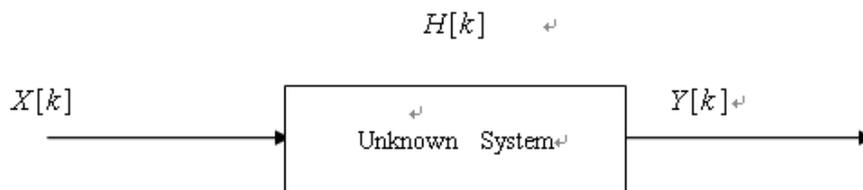**Figure 7.6** Detection of a sinusoidal signal from noisy background

Figure 7.7 System identification using input and output spectra

Problems

Q7.1 Examine sinusoidal components in a signal by spectral analysis

$$x[n] = 0.1\sin\left(\frac{2\pi 3\, n}{1024}\right) + 0.2\sin\left(\frac{2\pi 107.5 n}{1024}\right) + 0.5\sin\left(\frac{2\pi 422.5 n}{1024}\right)$$

The first component gives a single spectral coefficient. The second and the third components display side-lobes around a high coefficient, which represent the spectral leakage. Using the following MATLAB code, plot the signal in the time and frequency domains on the screen and sketch them on a paper.

%Examine signal components

%sampling rate 1024Hz

N=1024;

dt=1/N;

t=0:dt:1-dt;

%****define the signal**

x=0.1*sin(2*pi*32 .*t) + ...

 0.2*sin(2*pi*107.5 .*t)+ ...

 0.15*sin(2*pi* 422.5 .*t) ;

%**********************

subplot(2,1,1);

plot(t,x);

axis([0 1 -1 1]);

xlabel(‚seconds‘);

ylabel(‚x[n]‘);

X=fft(x);

df=1;

f=0:df:N-1;

subplot(2,1,2);

plot(f,abs(X));

axis([0 N/2 0 120]);

xlabel(‚Hz‘);

ylabel(‚|X[k]|‘);

Q7.2 Use MATLAB code to apply rectangular, triangular and Hamming windows to the signal

$$x[n] = 0.1\sin\left(\frac{2\pi 2\ n}{1024}\right) + 0.2\sin\left(\frac{2\pi 196.5 n}{1024}\right) + 0.15\ \sin\left(\frac{2\pi 408 n}{1024}\right)$$

Observe the level of magnitude and the distribution of side-lobs carefully to find any improvement of leakage reduction. Sketch the screen display and make your comments. Make sure the shape of the envelope in the time domain is drawn according to the window function.

8 Summary

After digesting the content of this book, I hope beginners may have gained an overview of the topic of digital signal and system analysis, have understood the basic methods and have known how to personally deal with digital signals and digital systems. No matter the incentive is curiosity, interest or acquiring needed knowledge for one's profession, the content selected in this book should be well suited. The standards of learning outcome are equivalent to university year two which lays a good foundation for higher level studies or moving on to specialised topics, such as digital filters, digital communications, discrete time-frequency and time-scale analysis.

Bibliography

[1]   Bracewell, R.N., The Fourier Transform and its Applications, McGraw-Hill Book Company, ISBN 0-07-066454-4

[2]   Brigham E O, *The Fast Fourier Transform*, Prentice Hall, 1974

[3]   Chen, C-T, Digital Signal Processing, Oxford University Press 2001, ISBN 0-19-513638-1

[4]   Cooley J W & Tukey J W. An algorithm for the machine calculation of complex Fourier series. *Math. Comput.* 19:297-301, 1965.

[5]   Cristi, R, Modern Digital Signal Processing, Thomson Brooks/Cole, ISBN 0-534-40095-7

[6]   Denbigh P, *System Analysis and Signal Processing*, Addision-Wesley, ISBN:0-201-17860-5, Mulgrew B and Grant P, Digital Signal Processing, PALGRAVE, ISBN 0-333-74531-0

[7]   Ifeachor E C and Jerivs B W, Digital Signal Processing- A Practical Approach, Prentice Hall, ISBN 0-201-59619-9. 2002.

[8]   Kuo S M and Gan W-S, Digital Signal Processors, ISBN 0-13-127766-9. 2005.

[9]   Lynn PA, Fuerst W, *Introductory Digital Signal Processing with Computer Applications*, John Wiley & Sons, ISBN: 0-471-97631-8

[10]   Mulgrew B, Grant P and Thompson J, Digital Signal Processing Concepts and Applications, ISBN 0-333-96356-3. 2003.

[11]   Proakis JD and DG Manolakis DG, *Digital Signal Processing*, Prentice Hall, 4ed,

ISBN: 0-13-187374-1.

[12]   Stranneby, D and Walker W, Digital Signal Processing and Applications, Elsevier, ISBN 0-7506-6344-8. 2004.