Chapter 4

Information source and source coding

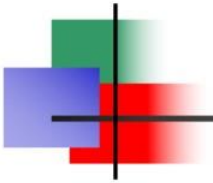
—— by Prof. XIAOFENG LI

SICE, UESTC

Problems

Ch4: Information source and source coding

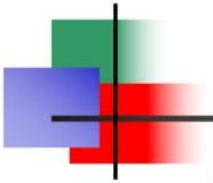
4.19, 4.20, 4.52(1)



Introduction

ADC and PCM

Modeling of digital source



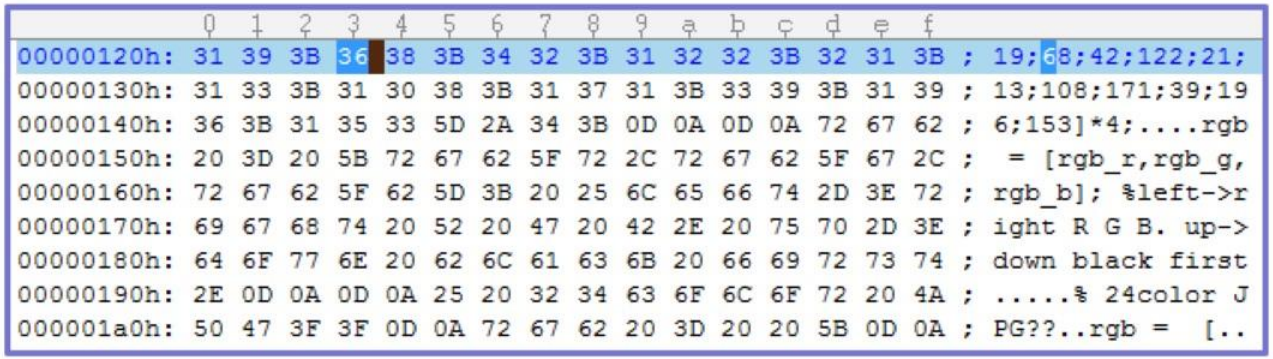
- A **analog** communication system transfers analog msg signals.
- A **digital** communication system transfers digital msg signals.

- Digital transmission have important **advantages** over analog one.

Introduction

Many info signals are naturally in the forms of sequence too, like {x0,x1,x2, ...}. For example, files on PC, such as

1. A digital picture in JPEG file;
2. A video file in H.264 format;
3. A music file in MP3 format;
4. An email



```
          0  1  2  3  4  5  6  7  8  9  a  b  c  d  e  f
00000120h: 31 39 3B 36 38 3B 34 32 3B 31 32 32 3B 32 31 3B ; 19;68;42;122;21;
00000130h: 31 33 3B 31 30 38 3B 31 37 31 3B 33 39 3B 31 39 ; 13;108;171;39;19
00000140h: 36 3B 31 35 33 5D 2A 34 3B 0D 0A 0D 0A 72 67 62 ; 6;153]*4;....rgb
00000150h: 20 3D 20 5B 72 67 62 5F 72 2C 72 67 62 5F 67 2C ;  = [rgb_r,rgb_g,
00000160h: 72 67 62 5F 62 5D 3B 20 25 6C 65 66 74 2D 3E 72 ; rgb_b]; %left->r
00000170h: 69 67 68 74 20 52 20 47 20 42 2E 20 75 70 2D 3E ; ight R G B. up->
00000180h: 64 6F 77 6E 20 62 6C 61 63 6B 20 66 69 72 73 74 ; down black first
00000190h: 2E 0D 0A 0D 0A 25 20 32 34 63 6F 6C 6F 72 20 4A ; .....% 24color J
000001a0h: 50 47 3F 3F 0D 0A 72 67 62 20 3D 20 20 5B 0D 0A ; PG??..rgb =  [..
```

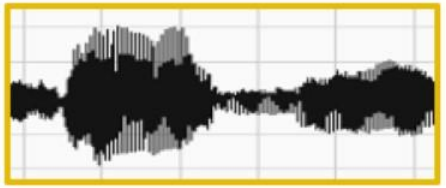
This type of info signals is called **digital signal** and They are **time-discrete** and has **finite values**.

Introduction

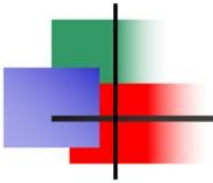
Many info signals are naturally in the forms of sequence too, like {x0,x1,x2, ...}. For example, files on PC, such as

1. A digital picture in JPEG file;
2. A video file in H.264 format;
3. A music file in MP3 format;
4. An email

There are also many analog msgs,



Analog msgs are often converted to digital form to take the advantages of digital transmission. PCM is a typical technique of this process.



Introduction

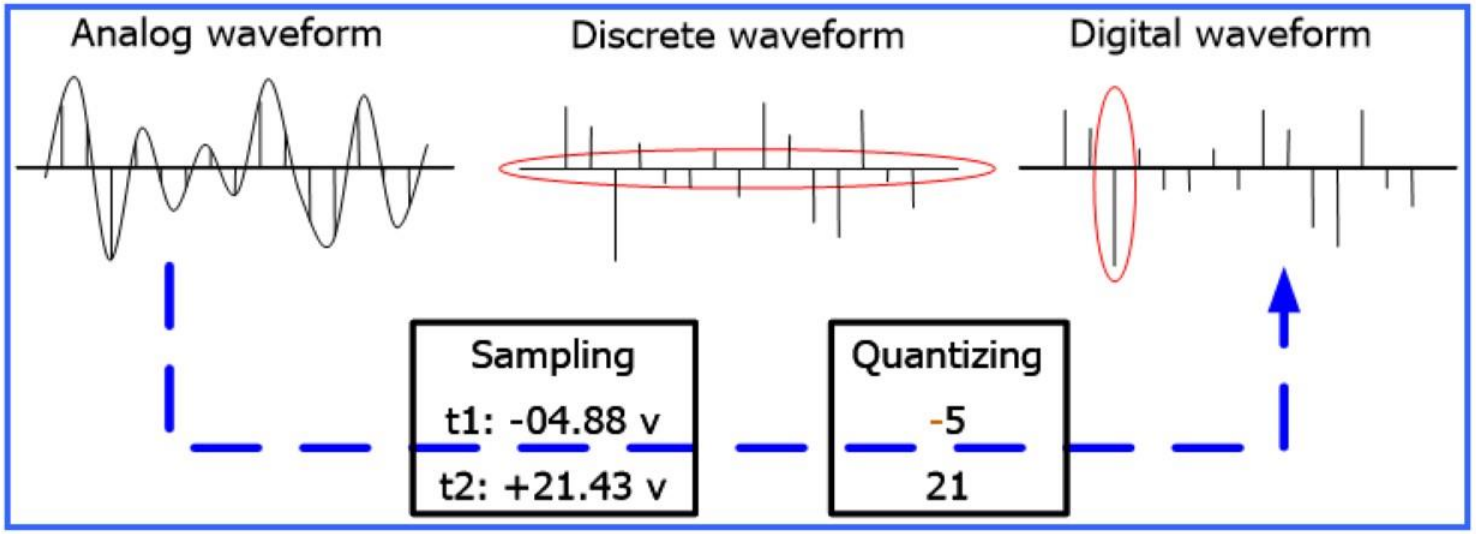
ADC and PCM

Modeling of digital source

1. ADC & PCM

The unit doing the conversion is called ADC. There are 2 important operations which make the waveform,

1. discrete in TIME; $x(t) \longrightarrow x[n]$
2. discrete in Value; $x[n] \longrightarrow x_d[n]$



Sampling: to sample the analog waveform every Ts and outputs a sequence of samples.

Quantizing: to approximating the continuous value by using finite numbers of levels

1. ADC & PCM

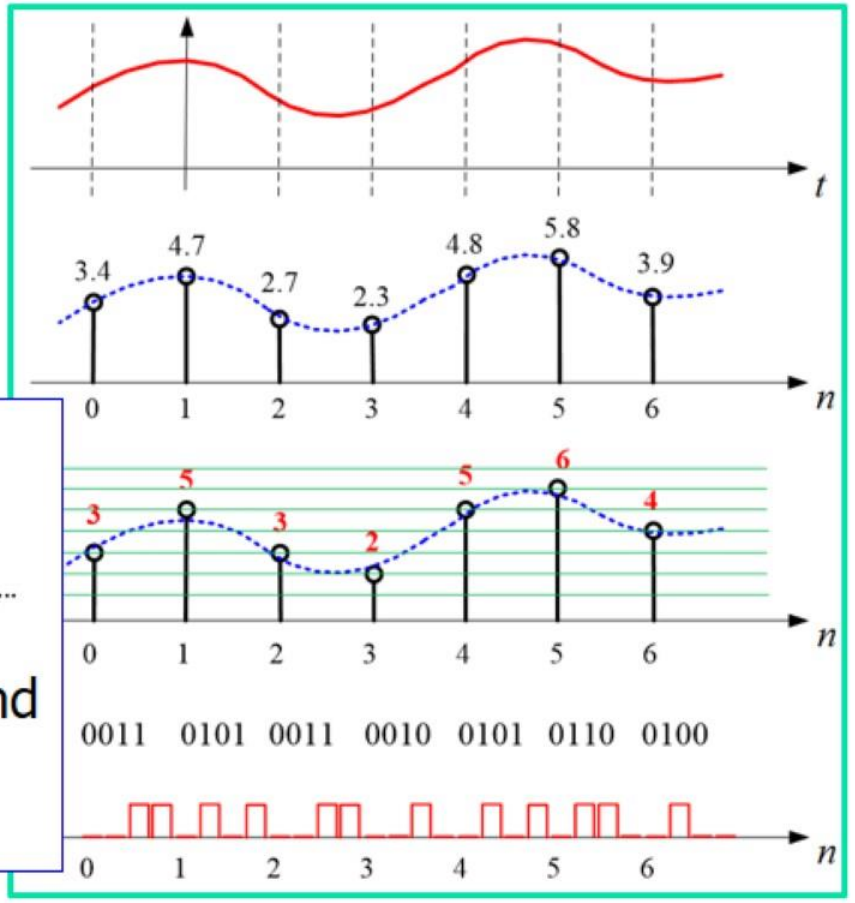
The **procedure** of ADC consists of 3 steps, sampling, quantizing and encoding.

Sampling: to sample the analog waveform every Ts and outputs a sequence of samples.

Sampling theory:

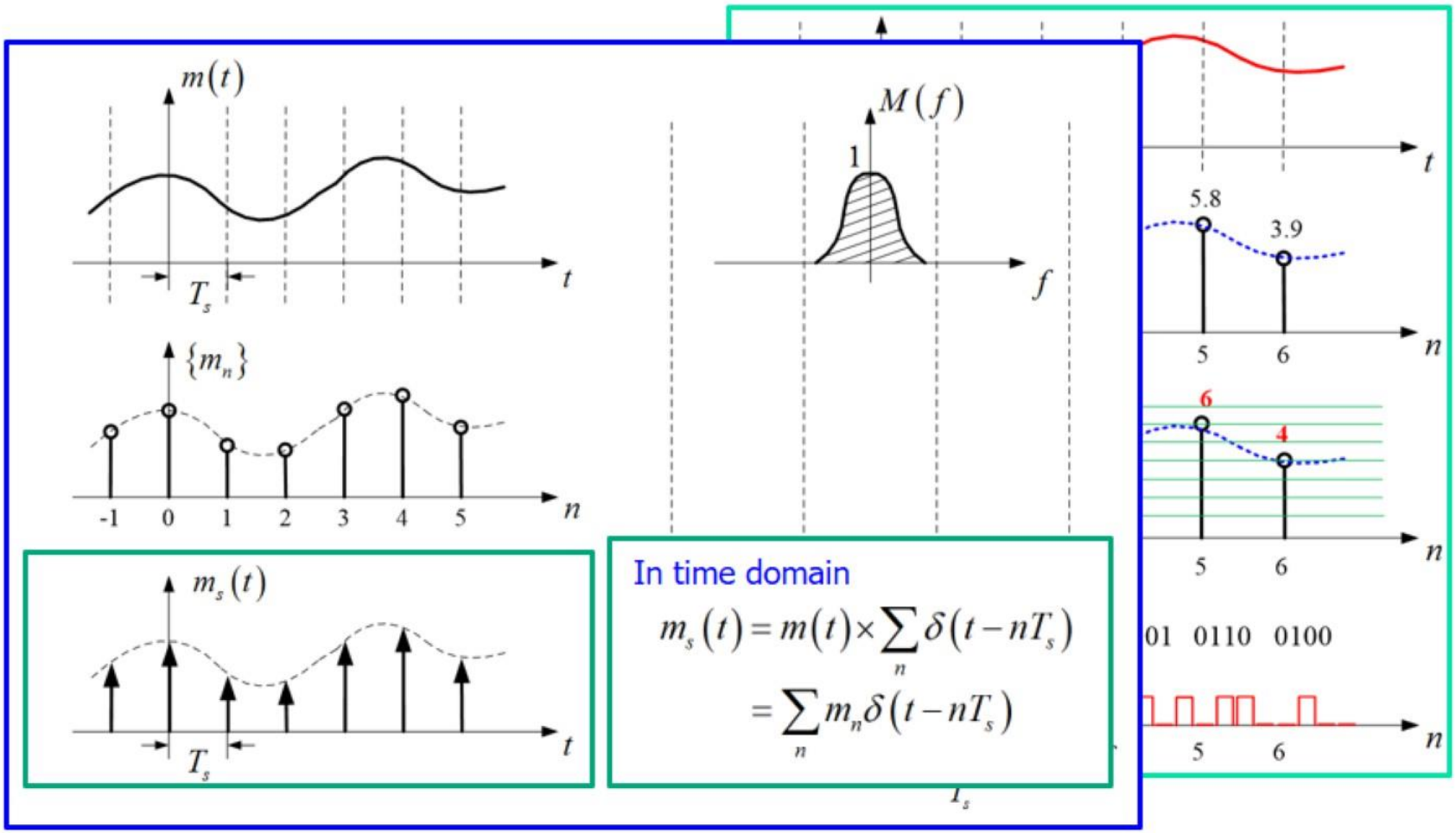
$$m(t) \xleftrightarrow{f_s > 2f_H} \{m_n\}_{n=0,\pm1,\pm2,\cdots}$$

Provided that m(t) is band-limited and the sampling rate fs is greater than the highest frequency fh.



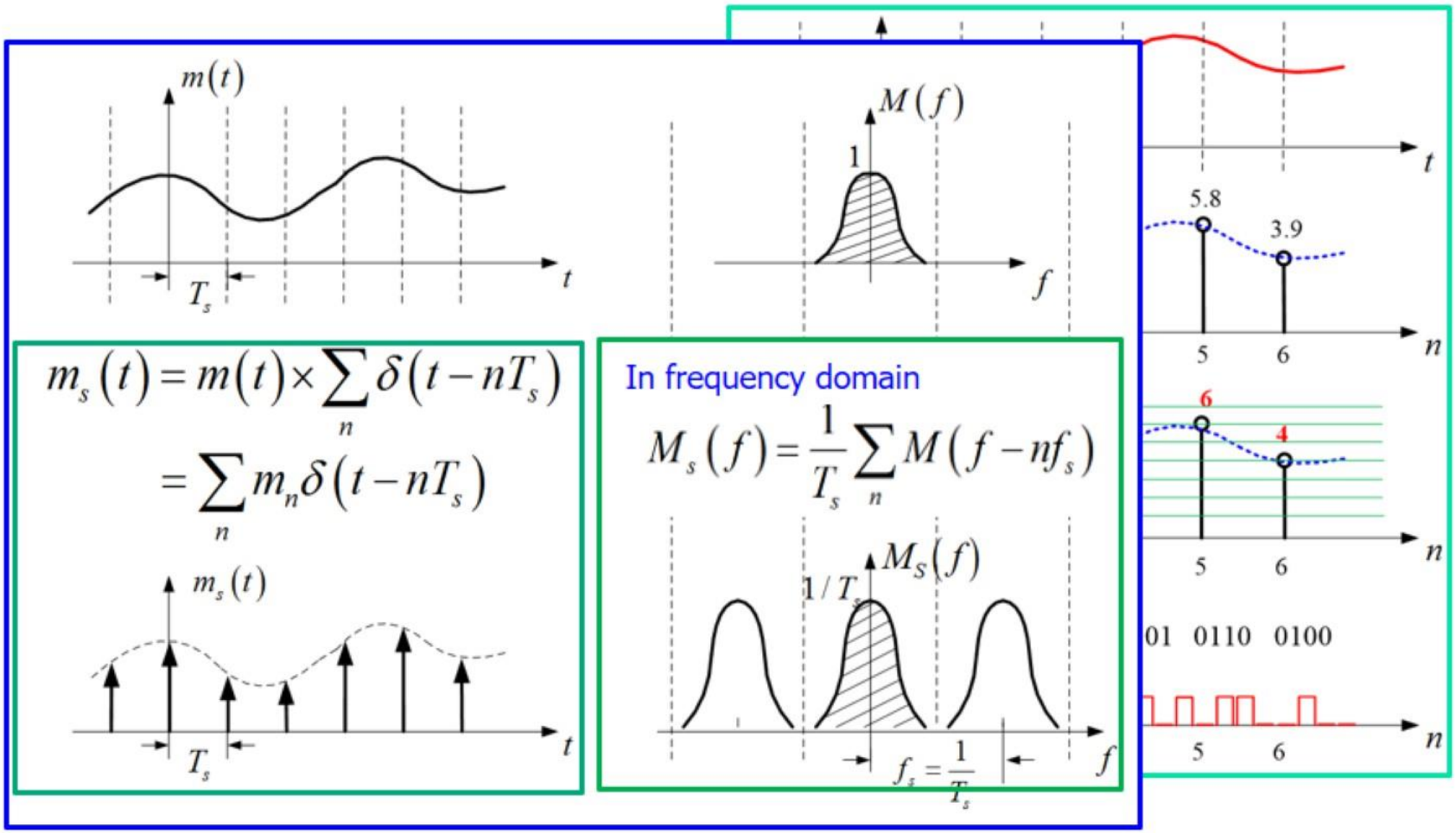
1. ADC & PCM

The **procedure** of ADC consists of 3 steps, sampling, quantizing and encoding.



1. ADC & PCM

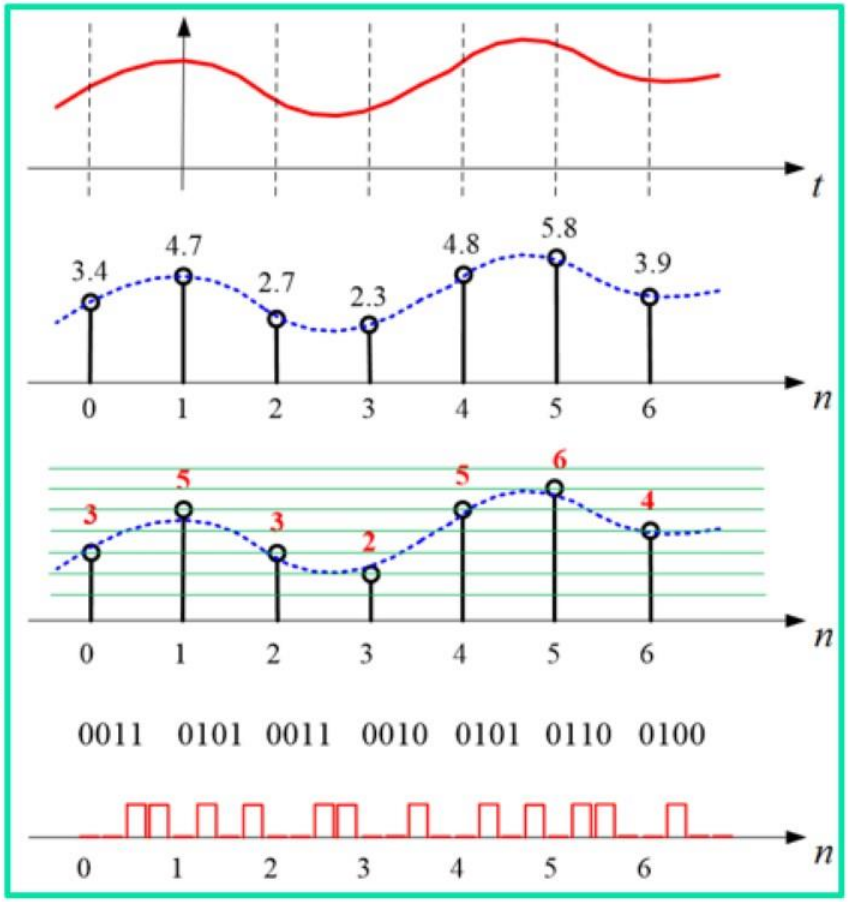
The **procedure** of ADC consists of 3 steps, sampling, quantizing and encoding.



1. ADC & PCM

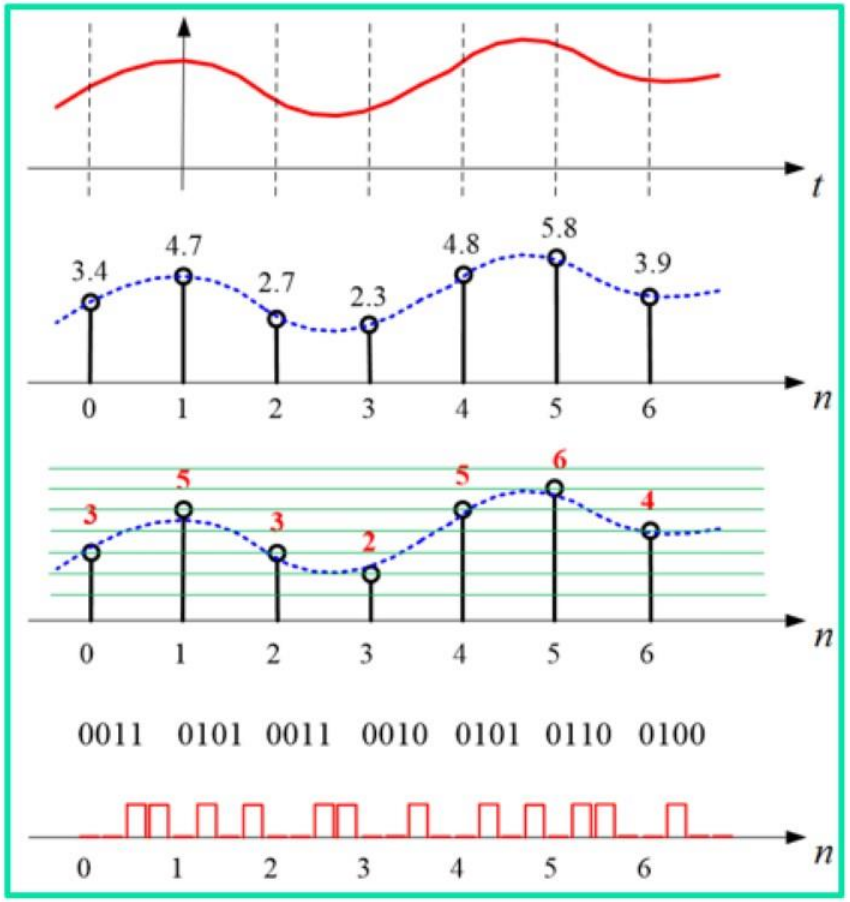
The **procedure** of ADC consists of 3 steps, sampling, quantizing and encoding.

Quantizing: to approximating the continuous value, by rounding real number into integer



1. ADC & PCM

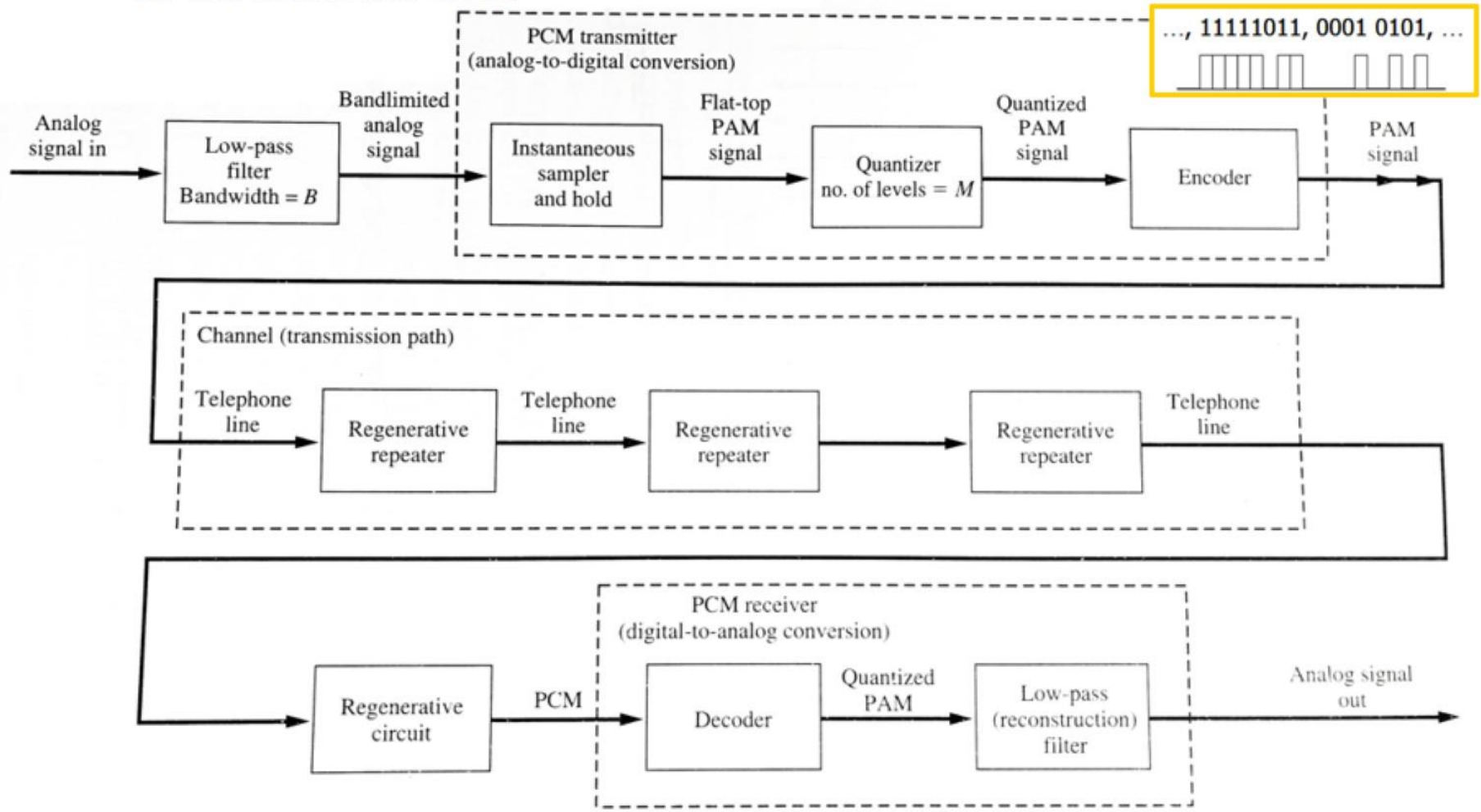
The **procedure** of ADC consists of 3 steps, sampling, quantizing and encoding.

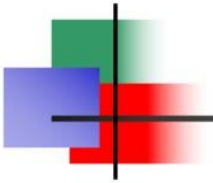


Encoding: to represent the levels into binary codes, more generally, into digital code-words.

1. ADC & PCM

PCM (Pulse Coding Modulation) is essentially an ADC used in telecomm svs.





Introduction

ADC and PCM

Modeling of digital source

2. Modeling of digital source

A simple model for the digital information source is the **DMS (discrete memoryless source)**.

$$\{x_n\} = x_0, x_1, x_2, x_3, \cdots x_n, \cdots$$

A random process

A DMS is a discrete-time, discrete- and finite amplitude random process (RP), in which all variables are generated independently and with the identical distribution. This implies that the outputs of a DMS are i.i.d. RVs (random variables) from a finite set.

To say the source is memoryless means that the RVs from the source are independent of each other.

2. Modeling of digital source

A simple model for the digital information source is the **DMS (discrete memoryless source)**.

$$\{x_n\} = x_0, x_1, x_2, x_3, \cdots x_n, \cdots$$

A random process

Let $A = \{a_1, a_2, ..., a_M\}$ denote the set;

$\{p_1, p_2, ..., p_M\}$ denote probabilities of the occurrence of the a_i's respectively.

Suppose X is a rv. from the DMS. We have,

$$P[X = a_i] = p_i \quad (i = 1, 2, ..., M)$$

A is called the alphabet of the DMS and a_i a symbol. The src is said to be a M-ary source, while M is the size of the alphabet.

2. Modeling of digital source

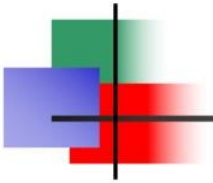
Binary DMS is an example with M=2, for binary message,

$$\{x_n\} = 1,0,0,1,1,0\cdots 1,0,1,\cdots$$

We have, $A = \{0,1\}$, where $\begin{cases} P(X=1) = p \\ P(X=0) = 1-p = q \end{cases}$

$$0 \le p \le 1$$

If $p = 0.5$, the src is called a binary symmetric src, or **BSS** for short.



End of this chapter

Thank you