Chapter 6  Synchronization and Equalization

6.1 TDMA

6.2 Synchronization : Introduction

6.3 Carrier Synchronization
   6.3.1 Phase - Locked Loop
   6.3.2 ML Estimator of Unmodulated Carrier Phase
   6.3.3 Suppressed Carrier Loops
   6.3.4 Decision-Directed Carrier Recovery

6.4 Frame Synchronization

6.5 Symbol Synchronization
   6.5.1 Maximum Likelihood Timing Estimation
   6.5.2 Timing Estimator with Error-tracking Loop
   6.5.3 Early-Late Gate Synchronizer
   6.5.4 Mueller and Muller Synchronizer
   6.5.5 Gardner Loop

6.6 Equalization : Introduction

6.7 Zero-Forcing Equalizer

6.8 Linear MMSE Equalizer

6.9 Adaptive Transversal Equalizer

6.10 Decision Feedback
References:


Source rate:

\[ R_m = \frac{1}{T_m} \log_2 q \ \text{bits/second} \]
6.1 TDMA Systems

- In a time-division multiple-access (TDMA) system a number of users, say, $N$, are permitted to access a wireless channel of bandwidth $B$ on a time-shared basis.

- Two apparent features that distinguish TDMA from FDMA are:
  1. Each user has access to the full bandwidth $B$ of the channel, whereas in FDMA each user is assigned a fraction of the channel bandwidth, namely, $B/N$.
  2. Each user accesses the channel for only a fraction of the time that is in use and on a periodic and orderly basis, with the transmission rate being $N$ times the user’s required rate. By contrast, in FDMA, each user accesses the channel on a continuous-time basis.
In TDMA system, access to the full bandwidth of the wireless channel makes the system vulnerable to frequency-selective fading.

In TDMA system, the information-bearing data is usually transmitted in the form of bursts, which complicates the requirement of synchronizing the receiver to the transmitter.

Each frame of the TDMA structure contains $N$ time slots of equal duration. It is in the detailed structure of each slot and in the way in which the transmitting and receiving slots assigned in time that TDMA systems differ from one another.

Typically, the bits constituting each slot of a TDMA frame are divided into two groups:
1. **traffic** data bits
2. overhead bits
The function of overhead bits is to assist the receiver in performing some auxiliary functions such as synchronization and channel estimation.

Example: GSM

- Uplink frequencies: 890 – 915 MHz
- Downlink frequencies: 935 – 960 MHz
- Carrier separation: 200 kHz
- Modulation: GMSK
- Multiple access: TDMA
- Compressed speech rate: 13 kHz
- Channel data rate: 270.833 kbps
- Frame period: 4.615 ms
- Error-correction code: Rate ½ convolutional code, constraint length = 5

Fig. 5.1 shows a typical frame structure in TDMA system.
Fig. 6.1 Frame Structure of GSM system

- Frame = 4.6155 ms

- TS: Time slot
- T: Tail (bits)
- F: Flag (bit)
- Train: Training interval for equalizer
- Guard: Guard time interval

- Time slot = 156.25 bits = 577 µs
6.2 Synchronization Techniques-Introduction

- In a digital communication system, the output of the demodulator must be sampled periodically, at least once per symbol interval, in order to recover the transmitted information. Since the propagation delay from the transmitter to the receiver is generally unknown at the receiver, symbol timing must be derived from the received signal in order to synchronously sample the output of the demodulator.

In addition to the effect of local oscillator at the receiver, the propagation delay also results in a carrier offset, which must be estimated at the receiver if the detector is phase-coherent.
Three major synchronization requirements must be employed in digital communication receiver.

a. Carrier frequency and phase synchronization

b. Frame (or symbol) synchronization, by which the correct frame (or symbol) start position is determined.

c. Sampling clock synchronization.
Synchronization can be implemented in one of the two fundamentally different ways:

1. Data-aided synchronization.

   In data-aided synchronization systems, a preamble is transmitted along with the data-bearing signal in a time-multiplexed manner on a periodic basis. The preamble contains information about the carrier and symbol timing, which is extracted by appropriate processing of the channel output at the receiver. By this approach, the time to synchronize the receiver to the transmitter may be reduced. However, data-throughput efficiency and power efficiency will be reduced.
(2) Nondata-aided synchronization.

In this approach, the use of preamble is avoided, the receiver must establish synchronization by extracting the necessary information from the modulated signal. Both data throughput and power efficiencies are thereby improved but at the expense of an increase in the time to establish synchronization.
6.3 Carrier synchronization

- Generation of a carrier reference at the receiver is necessary, especially for coherent communication systems that are phase-coherent with the received carrier.
- The passband received signal is downconverted to complex baseband using a local oscillator which is typically synthesized from a crystal oscillator.
- Crystal oscillators typically have tolerances of the order of 10 parts per million (ppm) so that the frequency of the local oscillator at the receiver typically has some difference $\Delta f$. 
The baseband equivalent model at the receiver can be expressed by

\[ y(t) = A \ e^{j\theta} \ x(t-\tau) \ e^{j2\pi \Delta f \ t} + n(t) \]
\[ = A \ e^{j2\pi \Delta f \ t + \theta} \ x(t-\tau) + n(t) \]  \hspace{1cm} (6.1)

where \( x(t) = \sum b_k p(t-kT) \) is the transmitted complex baseband signal, \( \tau \) is the transmission delay, \( \theta = -2\pi f_c \tau \) and \( f_c \) is the carrier frequency at the transmitter.
Carrier synchronization corresponds to estimation of $\Delta f$ and $\theta$.

Carrier synchronization typically involves two stages:

1. obtaining an initial estimate of the unknown parameters, and
2. tracking these parameters as they vary slowly over time (typically after the training phase, so that the $b_k$’s are unknown).

For packetized communication systems, the variations of the synchronization parameters over a packet are often negligible, and the tracking mode can often be eliminated.
Two approaches can be taken to carrier synchronization depending on whether there is a carrier component present in the modulated signal. If a carrier component is present, a phase-lock loop (PLL) circuit can be used to track it and the PLL output then is used as a reference for coherent demodulation. If a carrier component is not present, as in BPSK, a nonlinear operation must be performed on the modulated signal to generate a spectral component at the carrier frequency, and then to use phase-lock loop (PLL) for carrier and phase acquisition.

Initial estimation of the frequency and phase offsets are often obtained using a training sequence, with subsequent tracking in decision-directed mode, before coherent demodulation.
6.3.1 Phase - Locked Loop

- A phase-locked loop (PLL) consists of three components:
  1. a phase detector,
  2. a low-pass filter, and
  3. a voltage-controlled oscillator (VCO), as shown in Fig. 6.2.

- The VCO is an oscillator that produces a periodic waveform with a frequency that may be varied about some free-running frequency $f_0$, according to the applied voltage $v_2(t)$. When the applied voltage $v_2(t)$ is zero, the free running frequency $f_0$ is the frequency of the VCO.

- The phase detector produces an output signal $v_1(t)$ that is a function of the phase difference between the incoming signal $v_{in}(t)$ and the oscillator signal $v_o(t)$. 

- The filtered signal $v_2(t)$ is the control signal that is used to change the frequency of the VCO output.
- The PLL configuration may be designed so that it acts as a narrowband tracking filter when the low-pass filter (LPF) is a narrowband filter. In this operation mode, the frequency of the VCO will become that of one of the line components of the input signal spectrum, so that, in effect, the VCO output signal is a periodic signal with a frequency equal to the average frequency of this input signal component.
- Once the VCO has acquired the frequency component, the frequency of the VCO will track the input signal component if it changes slightly in frequency.
In another mode of operation, the bandwidth of the LPF is wider so that the VCO can track the instantaneous frequency of the whole input signal.

When the PLL tracks the input signal in either of these ways, the PLL is said to be “locked”.

If the PLL is built using analog circuits, it is said to be an ‘analog PLL‘. Conversely, if digital circuits and signals are used, the PLL is said to be a digital PLL.
Fig. 6.2 Basic PLL
6.3.2 ML Estimator of Unmodulated Carrier Phase

- Likelihood function

We express the received signal as

$$r(t) = s(t;\varphi,\tau) + n(t) \quad (6.2)$$

where $\varphi$ and $\tau$ represent the signal parameters to be estimated.

For simplicity, we let $\theta$ denote the parameter vector $\{\varphi,\tau\}$ so that $s(t;\varphi,\tau)$ is simply denoted as $s(t;\theta)$. Since $n(t)$ is additive white zero-mean Gaussian noise, the joint probability distribution function (PDF) $p(r,\theta)$ may be expressed as

$$p(r,\theta) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^N \exp\left\{-\sum_{n=1}^{\infty} \frac{[r_n - s_n(\theta)]^2}{2\sigma^2}\right\}$$

$$\quad (6.3)$$
where \( r_n = \int_0^T r(t) \psi_n(t) \, dt \) and
\[
s_n(\theta) = \int_0^T s(t;\theta) \psi_n(t) \, dt
\]
and \( T \) represents the integration interval in the expansion of \( r(t) \) and \( s(t;\theta) \).

We can prove that
\[
\lim_{N \to \infty} \left( \frac{1}{2 \sigma^2} \right) \sum_{n=1}^{N} [r_n - s_n(\theta)]^2
\]

\[
= \left( \frac{1}{N_0} \right) \int_{T_0} \left[ r(t) - s(t;\theta) \right]^2 \, dt
\]

Thus the maximization of \( p(r, \theta) \) with respect to the signal parameters \( \theta \) is equivalent to the maximization of the likelihood function
\[
\Lambda(\theta) = \exp \left\{ - \left( \frac{1}{N_0} \right) \int_{T_0} \left[ r(t) - s(t;\theta) \right]^2 \, dt \right\}
\]

(6.5)
ML Carrier Phase Estimation
For simplicity, we assume that the delay \( \tau \) is known and, in particular, we set \( \tau = 0 \).
Thus the likelihood function can be expressed as
\[
A(\phi) = \exp \left\{ - \left( \frac{1}{N_0} \right) \int_{T_0} \left[ r(t) - s(t; \phi) \right]^2 dt \right\}
\]
\[
= \exp \left\{ - \left( \frac{1}{N_0} \right) \int_{T_0} r^2(t) dt + \left( \frac{2}{N_0} \right) \int_{T_0} r(t) s(t; \phi) dt - \left( \frac{1}{N_0} \right) \int_{T_0} s^2(t; \phi) dt \right\}
\]
(6.7)

Note that the first term of the exponential factor does not involve the signal parameter \( \phi \). The third term is a constant equal to the signal energy over the time interval \( T_0 \).
Only the second term, which involves the cross correlation of the received signal \( r(t) \) with the signal \( s(t; \varphi) \), depends on the choice of \( \varphi \).

Thus we may express the likelihood function as

\[
A(\varphi) = \exp \left\{ \frac{2}{N_0} \int_{T_0} [r(t) s(t; \varphi)] dt \right\} \quad (6.8)
\]

The ML estimate \( \varphi_{ML} \) is the value of \( \varphi \) that maximize \( A(\varphi) \). Equivalently, the value \( \varphi_{ML} \) also maximizes the logarithm of \( A(\varphi) \), i.e.

\[
A_L(\varphi) = \left( \frac{2}{N_0} \right) \int_{T_0} [r(t) s(t; \varphi)] dt \quad (6.9)
\]

The following example is used to demonstrate that the PLL provides the ML estimate of the phase of an unmodulated carrier.
Consider the transmission of the unmodulated carrier $A \cos 2\pi f_c t$. The received signal is

$$r(t) = A \cos(2\pi f_c t + \phi) + n(t) \quad (6.10)$$

where $\phi$ is the unknown phase.

The log-likelihood function is

$$\Lambda_L(\phi) = \frac{2A}{N_0} \int_{T_0} [r(t) \cos(2\pi f_c t + \phi)] dt \quad (6.11)$$

A necessary condition for a maximum is that

$$d \Lambda_L(\phi) / d \phi = 0$$

This condition yields

$$\int_{T_0} [r(t) \cos(2\pi f_c t + \phi_{ML})] dt = 0$$

or, equivalently,

$$\phi_{ML} = -\tan^{-1} \left\{ \frac{\int_{T_0} [r(t) \sin(2\pi f_c t)] dt}{\int_{T_0} [r(t) \cos(2\pi f_c t)] dt} \right\} \quad (6.12)$$

The optimality condition given by the above equation implies the use of a loop to extract the estimate, as shown in Fig.6.3.
Fig. 6.3 A ML estimator of the phase of an unmodulated carrier
6.3.3 Suppressed Carrier Loops:
Costas Loop and Squaring Loop

- Two types of carrier recovery loops may be used to coherently demodulated a double side band suppressed carrier (DSB-SC) signal. Fig.6.4(a) shows the Costas loop and Fig.6.4 (b) shows the squaring loop. These loops can also be used for demodulating a BPSK signal, since the BPSK signal has similar mathematical form as a DSB-SC signal.

- Fig.6.5(a) and Fig.6.5(b) show the carrier recovery schemes in MPSK demodulator and QAM demodulator, respectively.
Fig. 6.4 (a) Costas loop for acquiring a coherent reference from a BPSK-modulated signal.
Fig. 6.4(b) Squaring loop for acquiring a coherent reference from a BPSK-modulated signal

\[ x_c(t) = A d(t) \cos(2\pi f_0 t + \theta) \]

\[ k + k \cos(4\pi f_0 t + 2\theta) \]

Coherent demodulator
Fig. 6.5 (a) $M$th power loop for carrier recovery in MPSK demodulator
Fig. 6.5(b) A QAM receiver with carrier recovery
6.3.4 Decision-Directed Carrier Recovery

- At the start of the carrier recovery process it is possible to achieve symbol synchronization prior to full carrier recovery because symbol timing can be determined without knowledge of the carrier phase or the carrier's minor frequency variation/offset.

- In decision directed carrier recovery the output of a symbol decoder is fed to a comparison circuit and the phase difference/error between the decoded symbol and the received signal is used to discipline the local oscillator.

- Decision directed methods are suited to synchronizing frequency differences that are less than the symbol rate because comparisons are performed on symbols at, or near, the symbol rate. Other frequency recovery methods may be necessary to achieve initial frequency acquisition.
6.4 Frame Synchronization

- Frame synchronization is usually accomplished by the detection of a unique sync word, which is a known pattern of 1s. This unique sync word and 0s is sent in the data stream periodically, sometimes at the start of the block of data (known as preamble) or in the middle (known as midamble). To ensure that the data itself does not contain the unique word pattern the process of bit stuffing is used. For example, if the unique sync word is fifteen 1s followed by a 0, the data sequence at the transmitter is examined, and if the sequence of fourteen 1s occurs, a 0 is “stuffed” at the transmitter and removed at the receiver. Fig. 6.6 shows a typical structure of a frame.
- A digital correlator, as shown in Fig. 6.7, can be used to detect the position of the unique sync word.
- If the unique sync word happens to be present in the input shift register, all the inputs to the correlator will be binary 1s, and the output of the correlator will go high, and the sync word ids perfectly aligned in the shift register. Thus the frame synchronization is achieved.
- Usually a pseudo random (PN) sequence is chosen as the unique sync word.
- Fig. 6.8 shows a frame synchronizer with receiver front-end.
Fig. 6.6 A typical frame structure in digital transmission
Fig. 6.7 Digital correlator

**Figure 5-9** A digital correlator for detecting a unique word in frame synchronization.
Fig. 6.8 Frame synchronizer with receiver front-end
6.5 Symbol Synchronization

- Symbol (or clock) synchronization can be processed alongside carrier recovery. Alternately, clock synchronization is accomplished first, followed by carrier recovery: sometimes, the reverse procedure is followed. The choice of a particular approach or the other is determined by the application of interest.

- In one approach, the symbol synchronization problem is solved by transmitting a clock along with the data-bearing signal, in multiplexed form. Then, at the receiver, the clock is extracted by appropriate filtering of the modulated waveforms. Such an approach minimizes the time required for carrier/clock recovery. However, a disadvantage of the approach is that a fraction of the transmitted power is allocated to the transmission of the clock.
In another approach, a good method is, first, to use a noncoherent detector to extract the clock. In this approach, we make use of the fact that clock timing is usually much more stable than carrier phase. Then the carrier is recovered by processing the noncoherent detector output in each clocked interval.

In yet another approach, when clock recovery follows carrier recovery, the clock is extracted by processing demodulated (not necessarily detected) baseband waveforms, thereby avoiding any waste of transmitted power.
6.5.1 Maximum Likelihood Timing Estimation

- We are to derive the maximum likelihood estimate of delay assuming the carrier phase $\psi$ is known, the timing recovery will not affect the carrier phase recovery loop and associated downconversion.

- We denote the in-phase and quadrature-phase components for the lowpass equivalent signal $r(t)$ as $r_I(t)$ and $r_Q(t)$, and for $s(t;\tau)$ as $s_I(t;\tau)$ and $s_Q(t;\tau)$, respectively.

  Here we focus on the in-phase branch because the timing recovered from this branch can be used for the quadrature branch.
The equivalent lowpass signal is given by
\[ s_I(t; \tau) = \sum_k s_I(k) g(t - kT_s - \tau) \]  
(6.13)

where \( g(t) \) is the in-phase pulse shape and \( s_I(k) \) denotes the amplitude associated with the in-phase component of the message transmitted over the \( k \)-th symbol period. The in-phase component of the received signal is
\[ r_I(t) = s_I(t; \tau) + n_I(t) \]

The likelihood function with known phase is then given by
\[
\Lambda(\tau) = \exp \left\{ - \left( \frac{1}{N_0} \right) \int_{T_0} \left[ r_I(t) - s_I(t; \tau) \right]^2 dt \right\} \\
= \exp \left\{ - \left( \frac{1}{N_0} \right) \int_{T_0} r_I^2(t) dt \\
+ \left( \frac{2}{N_0} \right) \int_{T_0} r_I(t) s_I(t; \tau) dt \\
- \left( \frac{1}{N_0} \right) \int_{T_0} s_I^2(t; \tau) dt \right\} 
\]  
(6.14)
The ML timing estimate $\tau$ must satisfy

$$\sum_k s_1(k) \frac{dz_k(\tau)}{d\tau} = 0 \quad (6.15)$$

where

$$z_k(\tau) = \int_{T_0} \{r(t) g(t - kT_s - \tau^\wedge) \} dt \quad (6.16)$$

For decision-directed estimation, Eq. (6.16) gives rise to the estimator shown in Fig. 6.9. The input to the voltage-controlled clock (VCC) is (6.15). If the input is zero, then the timing estimate $\tau^\wedge = \tau$. If not, the clock, that is the timing estimate, is adjusted to drive the VCC input to zero.
Fig. 6.9   Decision-directed maximum likelihood timing estimator
A non-decision-directed timing estimate can be obtained by averaging the likelihood ratio $\Lambda(\tau)$ over the PDF of the information symbols, to obtain $\Lambda(\tau)$, and then differentiating $\Lambda(\tau)$ to obtain the condition for the maximum-likelihood estimate $\tau_{ML}$.

In the case binary (baseband) estimator, with assumption that the information symbols have Gaussian probability distribution function (PDF), it can be shown that the maximum likelihood estimate of $\tau$ can be obtained by solving the equation

$$\sum_k z_k(\tau) \frac{d z_k(\tau)}{d\tau} = 0$$

(6.17)

An implementation of a tracking loop based on Eq. (6.17) is shown in Fig.6.10.
Fig. 6.10  Non-decision-directed ML estimation of timing for baseband PAM signal.
6.5.2 Timing Estimator with Error-tracking Loop

- A general all-digital timing estimator structure with error-tracking loop is illustrated as Fig. 6.11.

- The interpolator is able to generate samples in between those actually sampled by the A/D converter. By generating these intermediate samples as needed, the interpolator can adjust the effective sampling frequency and phase.

  Polynomial interpolators, which can be derived using Lagrange’s interpolation formula has an advantage that it can be implemented easily by FIR filters.

  The simplest interpolator is a linear interpolator.
The loop filter can be a second-order filter which has a proportional plus integral path configuration as shown in Fig.6.12.

There are several well-known algorithms for the timing-error estimation:

- early-late gate timing-error detection algorithm --- non-decision-directed
- Mueller and Muller timing-error detector --- decision-directed
- Gardner loop for timing-error detection --- decision-directed
Fig. 6.11 General timing – error tracking synchronizer
Fig. 6.12 A typical second-order loop filter
6.5.3 Early-Late Gate Synchronizer

- The early-late gate synchronizer is essentially a tracking loop and is one of the most simple and widely used bit synchronizer.

- An early-late gate synchronizer is a non-decision-directed timing estimator which exploits the symmetry properties of the signal at the output of the matched filter or correlator.

- Consider the matched filter, as shown in Fig. 6.13, which has an input of rectangular pulse \( s(t) \), \( 0 \leq t \leq T \). The output attains its maximum value at time \( t = T \) as shown in Fig, 6.13. Thus the output of the matched filter is the time autocorrelation of the pulse \( s(t) \). The proper time to sample the output of the matched filter for a maximum output is at \( t = T \), i.e., at the peak of the correction.
In the presence of noise, the identification of the peak the value of the signal is generally difficult. Instead of sampling the signal at the peak, we sample early, at \( t = T - \delta \) and late at \( t = T + \delta \).

The absolute values of the early samples \( |y[m(T - \delta)]| \) and the late samples \( |y[m(T + \delta)]| \) will be smaller than the samples of the peak value, \( |y[m(T)]| \), on the average. Under this condition, the proper sampling time is the midpoint between \( T - \delta \) and \( T + \delta \).

Fig. 6.14 shows the block diagram of an early-late gate synchronizer.
Fig. 6.13 Matched Filter

(a)\\s(t)\\n
(b)\\text{Matched filter output}\\text{Early sample}\\text{Optimum sample}\\text{Late sample}\\n
0 \quad T-\delta \quad T \quad T+\delta \quad 2T
Fig. 6.14 Early-late gate synchronizer
6.5.4 Mueller and Muller Synchronizer

- The Mueller and Muller algorithm only require one sample per symbol. The error term is calculated by using the equation

\[ x_k(\varepsilon, \varepsilon^\wedge) = a^\wedge_{k-1} y_{k+\varepsilon} - a^\wedge_k y_{k-1+\varepsilon} \]  

(6.18)

where \( a_m \) denotes the receiver’s decision about the \( m \)-th channel symbol \( a_m \).

The sampler takes the sample s at the receive filter output at a rate of \( 1/T \). For simplicity, we assume that all decisions are correct, i.e. \( a^\wedge_m = a_m \).
This algorithm is based on the assumption that Nyquist pulses with perfect symmetry. The synchronizer, as shown in Fig. 6.15, uses decision-feedback to produce a timing error detector output. In the absence of data transitions, the useful timing error detector output equals zero, therefore the transmission of long strings of identical symbols must be avoided.

This algorithm is sensitive to carrier offsets and thus carrier recovery must be performed prior to the M&M timing recovery.
Fig.6.15  M&M synchronizer
6.5.5 Gardner Loop

- The Gardner algorithm has been widely applied in many practical timing recovery loop implementations. The algorithm uses two samples per symbol and has the advantage of being insensitive to carrier offsets.

- Fig. 6.16 shows a block diagram of a timing recovery scheme using Gardner algorithm. The output of the timing-error detector is given by

\[ e_k = \text{Re} \{ [y_{k-1} - y_k] y^*_k \} \]

\[ = y^I_{k-1/2} [y^I_{k-1} - y^I_k] + y^Q_{k-1/2} [y^Q_{k-1} - y^Q_k] \]  \hspace{1cm} \text{(6.19)}

where \( y^I_k \) and \( y^Q_k \) denote the I and Q channel strobe samples of the k-th symbol, respectively.
Fig. 6.16 General all-digital timing recovery scheme
Fig. 6. General timing – error tracking synchronizer
6.6 Equalization Techniques - Introduction

- For transmitting digital signals over a communication channel, each pulse is usually shaped by a raised-cosine transfer function

\[ H_{rc}(f) = H_T(f) H_R(f) \]  as shown in Fig. 6.17.

- However, the channel may induce intersymbol interference to the pulse train due to frequency selective fading or multipath effect. The received digital signal exhibit distortions, as shown in Fig. 6.18.

  - The pulse sidelobes do not go through zeros at sampling instants adjacent to the mainlobe of each pulse. The distortion can be viewed as positive or negative echoes occurring both before and after the main lobe. To achieve an overall raised-cosine transfer function, an equalizing filter is required at the receiver end to compensate the distortion.
Fig. 6. 17 Communication system with equalization filter

Data source \( H_T(f) \) Transmitter filter \( H_C(f) \) Channel filter \( H_R(f) \) Equalization filter \( v(t_m) = v(t_m) = mT_b + t_d \)
Fig. 6.18 Receive pulse exhibiting distortion
Figure 11.2: Equalizer types, structures, and algorithms.
6.7 Zero-Forcing Equalizer

- An ideal zero-ISI equalizer is simply an inverse filter which has a frequency response that is the inverse of the channel’s frequency response. This inverse filter is often approximated by a finite-impulse response (FIR) filter or transversal filter, as shown in Fig.6.19.

- Consider a transversal filter with $2N+1$ taps, the impulse response of the filter is
  \[ h(t) = \sum_{k=-N}^{N} c_k \delta (t- kT) \tag{6.20} \]

  and the output of the filter $z(t)$ can be expressed as
  \[ z(t) = \sum_{k=-N}^{N} c_k y(t- kT) \tag{6.21} \]

  where $T$ is the time interval between adjacent taps of the filter.
Fig. 6.19 Transversal filter structure

Input: $y(t)$
The zero-forcing solution can be applied to the samples of \( z(t) \) taken at time \( t = kT \). These samples are

\[
z( mT ) = \sum_{k = -N}^{N} c_k y(mT - kT) \tag{6.22}
\]

Since there are \( 2N + 1 \) equalizer coefficients, we can control only \( 2N + 1 \) sampled values of \( z(t) \). By selecting the \( \{ c_n \} \) weights so that the equalizer output is forced to zero at \( N \) sample points on either side of the desired pulse.

\[
z( mT ) = \sum_{k = -N}^{N} c_k y(mT - kT)
\]

\[
\begin{align*}
1 & \quad \text{for } m = 0 \\
= & \\
0 & \quad \text{for } m = \pm 1, \pm 2, \pm N \tag{6.23}
\end{align*}
\]
Denote that

\[
Z_{eq} = \begin{bmatrix} 0 & 0 & \ldots & 0 & 1 & 0 & \ldots & 0 & 0 \end{bmatrix}^T
\]

\[
C = \begin{bmatrix} c_{-N} & c_{-N+1} & \ldots & c_N \end{bmatrix}^T
\]  \hspace{1cm} (6.24)

and

\[
y(0) \quad y(-T) \quad \ldots \quad y(-2NT) \\
y(T) \quad y(0) \quad \ldots \quad y((-2N+1)T) \\
y(2T) \quad y(T) \quad \ldots \quad y((-2N+2)T)
\]

\[
Y = \begin{bmatrix} \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots \\
y(2NT) \quad y((2N-1)T) \quad \ldots \quad y(0) \end{bmatrix}
\]  \hspace{1cm} (6.25)

Then Eq.(5.4) can be expressed as

\[
Z_{eq} = Y C
\]  \hspace{1cm} (6.26)

- Since the zero-forcing equalizer neglects the effects of noise, it is not always the best
Example: A 3-tap equalizer

Given a received distorted set of pulse samples \{0.0, 0.2, 0.9, -0.3, 0.1\}

\[ N=1, \text{we have} \]

\[
\begin{align*}
0 & \quad x(0) \quad x(-1) \quad x(-2) \quad \cdots \quad c_{-1} \\
1 & = \quad x(1) \quad x(0) \quad x(-1) \quad \cdots \quad c_0 \\
0 & \quad x(2) \quad x(1) \quad x(0) \quad \cdots \quad c_1 \\
\end{align*}
\]

\[
\begin{align*}
0.9 & \quad 0.2 & \quad 0 & \quad \cdots \quad c_{-1} \\
-0.3 & \quad 0.9 & \quad 0.2 & \quad \cdots \quad c_0 \\
0.1 & \quad -0.3 & \quad 0.9 & \quad \cdots \quad c_1 \\
\end{align*}
\]

Then we obtain

\[
\begin{align*}
c_{-1} & = -0.2140 \\
c_0 & = 0.9631 \\
c_1 & = 0.3448
\end{align*}
\]
6.8 Linear MMSE Equalizer

- Suppose that the desired output from the transversal filter equalizer is $d(t)$, which is a sequence of $\pm 1$-valued pulse of duration $T$ seconds.

- The filter tap-weights are chosen so that the mean-square error between desired output $d(t)$ and its actual output is minimized. As shown in Fig.5.7, the actual output, including noise, is denoted as $z(t)$. The minimum mean-square error (MMSE) criterion may be expressed as

$$\varepsilon = E \left[ (z(t) - d(t))^2 \right] = \text{minimum}$$
For a transversal filter input denoted by \( y(t) \), which includes AWGN, the filter output is

\[
z(t) = \sum_{k=-N}^{N} c_k y(t-kT)
\]

The mean-square error is generally a concave function of the tap-weights. The optimum weight vector \( C_o \) can be obtained by Wiener-Hopf equation;

\[
R_{yy} C_{op} = R_{yd} \quad \text{or} \quad C_{op} = R_{yy}^{-1} R_{yd}
\]

(6.27)

where

\[
R_{yy} = \begin{bmatrix}
R_{yy}(0) & R_{yy}(T) & \cdots & R_{yy}(2NT) \\
R_{yy}(T) & R_{yy}(0) & \cdots & R_{yy}((2N-1)T) \\
\vdots & \vdots & \ddots & \vdots \\
R_{yy}(2NT) & R_{yy}((2N-1)T) & \cdots & R_{yy}(0)
\end{bmatrix}
\]

(6.28)
\[
R_{yd}(-NT)
\]
\[
R_{yd}(-N+1)T)
\]
\[
R_{yd} = \ldots
\]
\[
R_{yd}(NT)
\]

where
\[
R_{yd}(mT) = \mathbb{E} \left[ y(t) d(t+mT) \right]
\]
\[
R_{yy}(mT) = \mathbb{E} \left[ y(t) y(t+mT) \right]
\]

and the MMSE is given by

\[
\varepsilon_{\min} = \mathbb{E} \left[ d^2(t) - R_{yd}^{-1} R_{yy} R_{yd} \right]
\]
Fig. 6.20 MMSE equalizer structure
6.9 Adaptive Transversal Equalizer

- Setting the tap-weights of the zero-forcing and MMSE equalizers involves the solution of a set of simultaneous equations. In these kinds of equalizers, to solve the equations, channel response or the autocorrelation of the measured data are required. In practice, it may be difficult or impossible to determine these quantities.

- An adaptive equalizer scheme is illustrated in Fig. 5.6. The tap-weights of the transversal filter can be adjusted iteratively, by the LMS algorithm, or other adaptive algorithms.

The LMS adaptive algorithm is given by

\[ c_k(n+1) = c_k(n) - \alpha y(n) \varepsilon(n) \]  \hspace{1cm} (6.31)
In adaptive equalization, there are two modes of operation: training and tracking.

During the training period, the coefficients of the equalizer are updated at time k based on a known training sequence that has sent over the channel.

During the tracking mode, the known training sequence is replaced by the output of the decision device.
Fig. 5.6 Adaptive Equalizer Scheme

Model of an adaptive equalizer in a data transmission system.
6.10 Decision-Feedback Equalizer

- The basic limitation of a linear equalizer, such as the transversal filter, is the poor performance on channels having spectral nulls.
- A decision-feedback equalizer (DFE), as shown in Fig. 5.7, consists of a feedforward filter $W(z)$ with the received sequence as input (similar to the linear equalizer) followed by a feedback filter $V(z)$ with the previously detected sequence as input.
- Assuming that $W(z)$ has $N_1 + 1$ taps and $V(z)$ has $N_2$ taps, we can express the DFE output as

$$z(k) = \sum_{i=-N_1}^{0} c_i y(k-i) - \sum_{j=1}^{N_2} c_j I_{k-j} \quad (6.32)$$

where the above equation, $\{c_k\}$ are the tap coefficients of the filter, $I_k$ is an estimate of the $k$-th information symbol, and $\{I_{k-1}, I_k, \ldots, I_{k-N_2}\}$ are previously detected symbols.
• Qualitative impulse of a discrete channel
Fig. 5.7 Decision feedback equalizer structure
A decision–feedback equalizer is a nonlinear equalizer that uses previous detector decisions to eliminate the ISI on pulses that are currently being demodulated. The ISI being removed was caused by the tails of the previous pulses.

The forward filter whitens the noise and produces a response with post-cursor ISI only. Its task may be viewed as the elimination of the precursors.

Since the feedback filter $V(z)$ sits in a feedback loop, it must be strictly causal or else the system is stable. The feedback filter of the DFE does not suffer from noise enhancement because it estimates the channel frequency response rather than its inverse. For channels with deep spectral nulls, DFEs generally perform much better than linear equalizer.

Both zero-forcing criterion and MMSE criterion can be applied to determine the filter coefficients.
Procedure for MMSE-DFE filter design:
1. Design linear filter $H_R(f)$ so that noise is minimized
2. Design feedback $FI_R$ filter so that ISI = zero

- In the ideal case (infinite-length feedback filter), all ISI can be completely eliminated!
- In practice, only *postcursor* ISI from a finite number of previous decisions can be eliminated. Precursor ISI can be reduced by linear (precursor) filter and adding delay in the system.

Error propagation:
One decision error in DFE causes a *burst of new errors*. The errors only stop after $M$ (= order of feedback filter) consecutive correct decisions.
Adaptive DFE Equalizer

In the case of adaptive DFE, the LMS algorithm for filter-coefficients updating is given by

\[ c_k(n+1) = c_k(n) + \mu \varepsilon(n) \ v(n-k) \]  \hspace{1cm} (6.33)

where \( \varepsilon(n) = I_\sim(n) - I_\wedge(n) \),

\[ v(n-k) = y(n-k) \] for feedforward section, i.e. \( k = -N_1, \ldots, 0 \)

\[ I_\sim(n-k) \] for feedback section, i.e., \( k = 1,2,\ldots, N_2 \)

\hspace{1cm} (6.34)
Note:

Maximum Likelihood Sequence Detection (MLSE) is the optimum method of suppressing ISI, but has complexity in the order of $O(M^v)$ where $M$ is the constellation size and $v$ is the channel delay. Therefore, MLSE is generally impractical on channels with relatively long delay spread or high data rate.

Reduced-state sequence estimation (RSSE) is a reasonable suboptimal approximation for MLSE in practical applications.