Amplitude Modulation Reception

Dr. Šarūnas Paulikas
Telekomunikacijų inžinerijos katedra
Elektronikos fakultetas, VGTU

Introduction

- AM reception is the reverse process of AM modulation.
- A conventional AM receiver simply converts an amplitude-modulated wave back to the original source information.
- An AM receiver must be capable of:
  - receiving,
  - amplifying,
  - and demodulating an AM wave.
- It must also be capable of bandlimiting the total radio frequency spectrum to a specific band of frequencies.
- This process is called tuning the receiver.
To understand the demodulation process, it is necessary to have a basic understanding of the terminology used to describe the characteristics of receivers.

**FIGURE 4-1** Simplified block diagram of an AM receiver

---

The RF section is the first stage and is often called the *receiver front end*. The primary functions of the RF section are detecting, bandlimiting, and amplifying the received RF signals.

The *mixer/converter* section down-converts the received RF frequencies to *intermediate frequencies* (IF). The primary functions of the IF section are amplification and selectivity.

The *AM detector* demodulates the AM wave and converts it to the original information signal.

The *audio section* simply amplifies the recovered information.
Receiver Parameters

There are several parameters commonly used to evaluate the ability of a receiver to successfully demodulate a radio signal. The most important parameters are selectivity and sensitivity, which are often used to compare the quality of one radio receiver to another.

Selectivity

Selectivity is a receiver parameter that is used to measure the ability of the receiver to accept a given band of frequencies and reject all others.

- For example, with the commercial AM broadcast band, each station's transmitter is allocated a 10-kHz bandwidth.
- Therefore, for a receiver to select only those frequencies assigned a single channel, the receiver must limit its bandwidth to 10 kHz.
- If the passband is greater than 10 kHz, more than one channel may be received and demodulated simultaneously.
- If the passband of a receiver is less than 10 kHz, a portion of the modulating signal information for that channel is rejected or blocked from entering the demodulator and, consequently, lost.
As stated earlier, thermal noise is directly proportional to bandwidth. Therefore, if the bandwidth is reduced, noise is also reduced by the same proportion. The noise reduction ratio achieved by reducing the bandwidth is called bandwidth improvement (BI).

As a signal propagates from the antenna through the RF section, the mixer/converter section, and the IF section, the bandwidth is reduced. Effectively, this is equivalent to reducing (improving) the noise figure of the receiver.

The bandwidth improvement factor is the ratio of the RF bandwidth to the IF bandwidth:

\[
BI = \frac{B_{RF}}{B_{IF}}
\]

The corresponding reduction in the noise figure due to the reduction in bandwidth is called noise figure improvement and is expressed mathematically as

\[
NF_{IMPROVEMENT} = 10\log BI
\]
Example

Determine the improvement in the noise figure for a receiver with an RF bandwidth equal to 200 kHz and an IF bandwidth equal to 10 kHz.

**Solution.**

- Bandwidth improvement is

\[ BI = \frac{200\, \text{kHz}}{10\, \text{kHz}} = 20 \]

- Noise figure improvement is

\[ NF_{\text{IMPROVEMENT}} = 10 \log 20 = 13\, \text{dB} \]

Sensitivity

The *sensitivity* of a receiver is the minimum RF signal level that can be detected at the input to the receiver and still produce a usable demodulated information signal.

What constitutes a usable information signal is somewhat arbitrary.

Generally, the signal-to-noise ratio and the power of the signal at the output of the audio section are used to determine the quality of a received signal and whether it is usable or not.

- For commercial AM broadcast band receivers, a 10-dB or more signal-to-noise ratio with 1/2 W (27 dBm) of power at the output of the audio section is considered to be usable.
- However, for broadband microwave receivers, a 40-dB or more signal-to-noise ratio with approximately 5 mW (7 dBm) of signal power is the minimum acceptable value.
Sensitivity

- The sensitivity of a receiver is usually stated in microvolts of received signal.
  - For example, a typical sensitivity for a commercial broadcast band AM receiver is 50 μV.
  - A two-way mobile radio receiver generally has a sensitivity between 0.1 μV and 10 μV.
- Receiver sensitivity is also called receiver threshold.
- The sensitivity of an AM receiver depends on the noise power present at the input to the receiver, the receiver's noise figure, the sensitivity of the AM detector, and the bandwidth improvement factor of the receiver.
- The best way to improve the sensitivity of a receiver is to reduce the noise level.
- This can be accomplished by reducing either the temperature or the bandwidth of the receiver or improving the receiver's noise figure.

Dynamic Range

- The dynamic range of a receiver is defined as the difference in decibels between the minimum input level necessary to discern a signal and the input level that will overdrive the receiver and produce distortion.
- In simple terms, dynamic range is the input power range over which the receiver is useful.
- The minimum receive level is a function of front-end noise, noise figure, and the desired signal quality.
- The input signal level that will produce overload distortion is a function of the net gain of the receiver (the total gain of all the stages in the receiver).
Dynamic Range

- The high-power limit of a receiver depends on whether it will operate with a single- or multiple-frequency input signal.
- A dynamic range of 100 dB is considered about the highest possible.
- A low dynamic range can cause a desensitizing of the RF amplifiers and result in severe intermodulation distortion of the weaker input signals.

Fidelity

- Fidelity is a measure of the ability of a communications system to produce, at the output of the receiver, an exact replica of the original source information.
- Any frequency, phase, or amplitude variations that are present in the demodulated waveform that were not in the original information signal are considered distortion.
- Essentially, there are three forms of distortion that can deteriorate the fidelity of a communications system:
  - amplitude,
  - frequency,
  - and phase.
Phase Distortion

Phase distortion is not particularly important for voice transmission because the human ear is relatively insensitive to phase variations. However, phase distortion can be devastating to data transmission. The predominant cause of phase distortion is filtering (both wanted and unwanted). Frequencies at or near the break frequency of a filter undergo varying values of phase shift. Consequently, the cutoff frequency of a filter is often set beyond the minimum value necessary to pass the highest-frequency information signals (typically the upper cutoff frequency of a low-pass filter is approximately 1.3 times the minimum value).

Phase Distortion

Absolute phase shift is the total phase shift encountered by a signal and can generally be tolerated as long as all frequencies undergo the same amount of phase delay. Differential phase shift occurs when different frequencies undergo different phase shifts and may have a detrimental effect on a complex waveform, especially if the information is encoded into the phase of the carrier as it is with phase-shift keying modulation. If phase shift versus frequency is linear, delay is constant with frequency. If all frequencies are not delayed by the same amount of time, the frequency-versus-phase relationship of the received waveform is not consistent with the original source information and the recovered information is distorted.
Amplitude and Freq. Distortion

Amplitude distortion occurs when the amplitude-versus-frequency characteristics of a signal at the output of a receiver differ from those of the original information signal.

Frequency distortion occurs when frequencies are present in a received signal that were not present in the original source information.

Frequency distortion is a result of harmonic and intermodulation distortion and is caused by nonlinear amplification.

Second-order products \(2f_1, 2f_2, f_2 \pm f_1\), and so on) are usually only a problem in broadband systems because they generally fall outside the bandwidth of a narrowband system.

Freq. Distortion

However, third-order products often fall within the system bandwidth and produce a distortion called third-order intercept distortion.

Third-order intercept distortion is a special case of intermodulation distortion and the predominant form of frequency distortion.

Third-order intermodulation components are the cross-product frequencies produced when the second harmonic of one signal is added to the fundamental frequency of another signal \(2f_1 \pm 2f_2, 2f_2 \pm 2f_1\), and so on).

Frequency distortion can be reduced by using a square-law device, such as a FET, in the front end of a receiver.

Square-law devices have a unique advantage over BJTs in that they produce only second-order harmonic and intermodulation components.
Insertion Loss

Insertion loss (IL) is a parameter that is associated with the frequencies that fall within the passband of a filter and is generally defined as the ratio of the power transferred to a load with a filter in the circuit to the power transferred to a load without the filter.

Because filters are generally constructed from lossy components, such as resistors and imperfect capacitors, even signals that fall within the passband of a filter are attenuated (reduced in magnitude).

Typical filter insertion losses are between a few tenths of a decibel to several decibels.

In essence, insertion loss is simply the ratio of the output power of a filter to the input power for frequencies that fall within the filter's passband and is stated mathematically in decibels as

\[ IL_{\text{dB}} = 10 \log \frac{P_{\text{out}}}{P_{\text{in}}} \]
Because thermal noise is directly proportional to temperature, it stands to reason that noise can be expressed in degrees as well as watts or volts.

_Equivalent noise temperature_ $T_e$ is a hypothetical value that cannot be directly measured.

\[ T = \frac{N}{kB} \]

\[ T_e = T(NR - 1) \]

$T_e$ is a parameter that is often used in low-noise, sophisticated radio receivers rather than noise figure. $T_e$ is an indication of the reduction in the signal-to-noise ratio as a signal propagates through a receiver. The lower the equivalent noise temperature, the better the quality of the receiver. Typical values for $T_e$ range from 20° for _cool_ receivers to 1000° for _noisy_ receivers.
There are two basic types of radio receivers: 

- coherent and
- noncoherent.

With a coherent or synchronous receiver, the frequencies generated in the receiver and used for demodulation are synchronized to oscillator frequencies generated in the transmitter (the receiver must have some means of recovering the received carrier and synchronizing to it).

With noncoherent or asynchronous receivers, either no frequencies are generated in the receiver or the frequencies used for demodulation are completely independent from the transmitter's carrier frequency.

Noncoherent detection is often called envelope detection because the information is recovered from the received waveform by detecting the shape of the modulated envelope.

The receivers described in lecture are noncoherent. Coherent receivers are described in next lecture.
The tuned radio frequency (TRF) receiver was one of the earliest types of AM receivers. TRF receivers are probably the simplest designed radio receiver available today; however, they have several shortcomings that limit their use to special applications.

FIGURE 4-3 Noncoherent tuned radio frequency receiver block diagram
Three-Stage TRF Receiver

- It includes:
  - an RF stage,
  - a detector stage, and
  - an audio stage.
- Generally, two or three RF amplifiers are required to filter and amplify the received signal to a level sufficient to drive the detector stage.
- The detector converts RF signals directly to information.
- Audio stage amplifies the information signals to a usable level.
- Although TRF receivers are simple and have a relatively high sensitivity, they have three distinct disadvantages that limit their usefulness to single-channel, low-frequency applications.

TRF Receiver Disadvantages

- The primary disadvantage is their bandwidth is inconsistent and varies with center frequency when tuned over a wide range of input frequencies.
- This is caused by a phenomenon called the skin effect. At radio frequencies, current flow is limited to the outermost area of a conductor; thus, the higher the frequency, the smaller the effective area and the greater the resistance.
- Consequently, the quality factor \((Q = \frac{XL}{R})\) of the tank circuits remains relatively constant over a wide range of frequencies causing the bandwidth \((f/Q)\) to increase with frequency.
- As a result, the selectivity of the input filter changes over any appreciable range of input frequencies. If the bandwidth is set to the desired value for low-frequency RF signals, it will be excessive for high-frequency signals.
Example

For an AM commercial broadcast-band receiver (535 kHz to 1605 kHz) with an input filter ($Q$-factor of 54), determine the bandwidth at the low and high ends of the RF spectrum.

Solution  The bandwidth at the low-frequency end of the AM spectrum is centered around a carrier frequency of 540 kHz and is

$$B = \frac{f}{Q} = \frac{540\text{ kHz}}{54} = 10\text{ kHz}$$

Example

The bandwidth at the high-frequency end of the AM spectrum is centered around a carrier frequency of 1600 kHz and is

$$B = \frac{f}{Q} = \frac{1600\text{ kHz}}{54} = 29.63\text{ kHz}$$

To achieve a bandwidth of 10 kHz at the high-frequency end of the spectrum, a $Q$ of 160 is required ($1600\text{ kHz}/10\text{ kHz}$). With a $Q$ of 160, the bandwidth at the low-frequency end is

$$B = \frac{f}{Q} = \frac{540\text{ kHz}}{160} = 33.75\text{ kHz}$$
TRF Receiver Disadvantages

- The second disadvantage of TRF receivers is instability due to the large number of RF amplifiers all tuned to the same center frequency.
- High-frequency, multistage amplifiers are susceptible to breaking into oscillations.
- This problem can be reduced somewhat by tuning each amplifier to a slightly different frequency, slightly above or, below the desired center frequency.
- This technique is called stagger tuning.
- The third disadvantage of TRF receivers is their gains are not uniform over a very wide frequency range.
- This is due to the nonuniform $L/C$ ratios of the transformer-coupled tank circuits in the RF amplifiers.

Superheterodyne Receiver

- The superheterodyne receiver has gain, selectivity, and sensitivity characteristics that are superior to those of other receiver configurations.
- *Heterodyne* means to mix two frequencies together in a nonlinear device or to translate one frequency to another using nonlinear mixing.
Essentially, there are five sections to a superheterodyne receiver:

- the RF section,
- the mixer/converter section,
- the IF section,
- the audio detector section, and
- the audio amplifier section.
The RF section generally consists of a preselector and an amplifier stage.

- They can be separate circuits or a single combined circuit.

The preselector is a broad-tuned bandpass filter with an adjustable center frequency that is tuned to the desired carrier frequency.

- The primary purpose of the preselector is to provide enough initial bandlimiting to prevent a specific unwanted radio frequency, called the image frequency, from entering the receiver (image frequency is explained later in this section).
- The preselector also reduces the noise bandwidth of the receiver and provides the initial step toward reducing the overall receiver bandwidth to the minimum bandwidth required to pass the information signals.

The RF amplifier determines the sensitivity of the receiver (that is, sets the signal threshold).

Also, because the RF amplifier is the first active device encountered by a received signal, it is the primary contributor of noise and, therefore, a predominant factor in determining the noise figure for the receiver.

A receiver can have one or more RF amplifiers or it may not have any, depending on the desired sensitivity.
RF Section

Several advantages of including RF amplifiers in a receiver are as follows:
• Greater gain, thus better sensitivity
• Improved image-frequency rejection
• Better signal-to-noise ratio
• Better selectivity

Mixer/Converter Section

The mixer/converter section includes:
• a radio-frequency oscillator stage (commonly called a local oscillator) and
• a mixer/converter stage (commonly called the first detector).

The local oscillator can be any of the oscillator circuits already discussed, depending on the stability and accuracy desired.

The mixer stage is a nonlinear device and its purpose is to convert radio frequencies to intermediate frequencies (RF-to-IF frequency translation).
Mixer/Converter Section

- Heterodyning takes place in the mixer stage, and radio frequencies are down-converted to intermediate frequencies.
- Although the carrier and sideband frequencies are translated from RF to IF, the shape of the envelope remains the same and, therefore, the original information contained in the envelope remains unchanged.
- It is important to note that, although the carrier and upper and lower side frequencies change frequency, the bandwidth is unchanged by the heterodyning process.
- The most common intermediate frequency used in AM broadcast-band receivers is 455 kHz.

IF Section

- The IF section consists of a series of IF amplifiers and bandpass filters and is often called the IF strip.
- Most of the receiver gain and selectivity is achieved in the IF section.
- The IF center frequency and bandwidth are constant for all stations and are chosen so that their frequency is less than any of the RF signals to be received.
- The IF is always lower in frequency than the RF because it is easier and less expensive to construct high-gain, stable amplifiers for the low-frequency signals.
- Also, low-frequency IF amplifiers are less likely to oscillate than their RF counterparts.
- Therefore, it is not uncommon to see a receiver with five or six IF amplifiers and a single RF amplifier or possibly no RF amplification.
Detector Section

The purpose of the detector section is to convert the IF signals back to the original source information.

The detector is generally called an audio detector or the second detector in a broadcast-band receiver because the information signals are audio frequencies.

The detector can be as simple as a single diode or as complex as a phase-locked loop or balanced demodulator.

Audio Section

The audio section comprises several cascaded audio amplifiers and one or more speakers.

The number of amplifiers used depends on the audio signal power desired.
Receiver Operation

- During the demodulation process in a superheterodyne receiver, the received signals undergo two or more frequency translations:
  - first, the RF is converted to IF;
  - then the IF is converted to the source information.
- The terms RF and IF are system dependent and are often misleading because they do not necessarily indicate a specific range of frequencies.
- For example, RF for the commercial AM broadcast band are frequencies between 535 kHz and 1605 kHz, and IF signals are frequencies between 450 kHz and 460 kHz.
- In commercial broadcast-band FM receivers, intermediate frequencies as high as 10.7 MHz are used, which are considerably higher than AM broadcast-band RF signals.
- Intermediate frequencies simply refer to frequencies that are used within a transmitter or receiver that fall somewhere between the radio frequencies and the original source information frequencies.

Frequency Conversion

- *Frequency conversion* in the mixer/converter stage is identical to frequency conversion in the modulator stage of a transmitter except that in the receiver the frequencies are down-converted rather than up-converted.
- In the mixer/converter, RF signals are combined with the local oscillator frequency in a nonlinear device.
- The output of the mixer contains an infinite number of harmonic and cross-product frequencies, which include the sum and difference frequencies between the desired RF carrier and local oscillator frequencies.
The IF filters are tuned to the difference frequencies.
The local oscillator is designed such that its frequency of oscillation is always above or below the desired RF carrier by an amount equal to the IF center frequency. Therefore, the difference between the RF and the local oscillator frequency is always equal to the IF. The adjustment for the center frequency of the preselector and the adjustment for the local oscillator frequency are gang tuned.
- Gang tuning means that the two adjustments are mechanically tied together so that a single adjustment will change the center frequency of the preselector and, at the same time, change the local oscillator frequency.

When the local oscillator frequency is tuned above the RF, it is called high-side injection or high-beat injection. When the local oscillator is tuned below the RF, it is called low-side injection or low-beat injection. In AM broadcast-band receivers, high-side injection is always used (the reason for this is explained later in this section).
Mathematically, the local oscillator frequency is:

- High-side injection
  \[ f_{LO} = f_{RF} + f_{IF} \]

- Low-side injection
  \[ f_{LO} = f_{RF} - f_{IF} \]

---

**Local Oscillator Tracking**

*Tracking* is the ability of the local oscillator in a receiver to oscillate either above or below the selected radio frequency carrier by an amount equal to the intermediate frequency throughout the entire radio frequency band.

- With high-side injection, the local oscillator should track above the incoming RF carrier by a fixed frequency equal to \( f_{RF} + f_{IF} \).
- With low-side injection, the local oscillator should track below the RF carrier by a fixed frequency equal \( f_{RF} - f_{IF} \).
The tuned circuit in the preselector is tunable from a center frequency of 540 kHz to 1600 kHz (a ratio of 2.96 to 1), and the local oscillator is tunable from 995 kHz to 2055 kHz (a ratio of 2.06 to 1).

Because the resonant frequency of a tuned circuit is inversely proportional to the square root of the capacitance, the capacitance in the preselector must change by a factor of 8.8, whereas, at the same time, the capacitance in the local oscillator must change by a factor of only 4.26.

The local oscillator should oscillate 455 kHz above the preselector center frequency over the entire AM frequency band, and there should be a single tuning control.
Fabricating such a circuit is difficult if not impossible. Therefore, perfect tracking over the entire AM band is unlikely to occur.

The difference between the actual local oscillator frequency and the desired frequency is called tracking error.

Typically, the tracking error is not uniform over the entire RF spectrum.

A maximum tracking error of ±3 kHz is about the best that can be expected from a domestic AM broadcast-band receiver with a 455-kHz intermediate frequency.

A tracking error of +3 kHz corresponds to an IF center frequency of 458 kHz, and a tracking error of -3 kHz corresponds to an IF center frequency of 452 kHz.
Three-Point Tracking

The tracking error is reduced by a technique, called three-point tracking.

The preselector and local oscillator each have a trimmer capacitor \( (C_t) \) in parallel with the primary tuning capacitor \( (C_0) \) that compensates for minor tracking errors at the high end of the AM spectrum.

The local oscillator has an additional padder capacitor \( (C_p) \) in series with the tuning coil that compensates for minor tracking errors at the low end of the AM spectrum.

With three-point tracking, the tracking error is adjusted to 0 Hz at approximately 600 kHz, 950 kHz, and 1500 kHz.
Low-Side Injection

- With low-side injection, the local oscillator would have to be tunable from 85 kHz to 1145 kHz (a ratio of 13.5 to 1).
- Consequently, the capacitance must change by a factor of 182.
- Standard variable capacitors seldom tune over more than a 10 to 1 range.
- This is why low-side injection is impractical for commercial AM broadcast-band receivers.
- With high-side injection, the local oscillator must be tunable from 995 kHz to 2055 kHz, which corresponds to a capacitance ratio of only 4.63 to 1.

Image Frequency

- An image frequency is any frequency other than the selected radio frequency carrier that, if allowed to enter a receiver and mix with the local oscillator, will produce a cross-product frequency that is equal to the intermediate frequency.
- An image frequency is equivalent to a second radio frequency that will produce an IF that will interfere with the IF from the desired radio frequency.
- Once an image frequency has been mixed down to IF, it cannot be filtered out or suppressed.
- If the selected RF carrier and its image frequency enter a receiver at the same time, they both mix with the local oscillator frequency and produce difference frequencies that are equal to the IF.
- Consequently, two different stations are received and demodulated simultaneously, producing two sets of information frequencies.
**Image Frequency**

For a radio frequency to produce a cross product equal to the IF, it must be displaced from the local oscillator frequency by a value equal to the IF.

With high-side injection, the selected RF is below the local oscillator by an amount equal to the IF.

Therefore, the image frequency is the radio frequency that is located in the IF frequency above the local oscillator.

\[ f_{IM} = f_{LO} + f_{IF} = f_{RF} + 2f_{IF} \]

Here we see that the higher the IF, the farther away in the frequency spectrum the image frequency is from the desired RF.

Therefore, for better image-frequency rejection, a high intermediate frequency is preferred.

However, the higher the IF, the more difficult it is to build stable amplifiers with high gain.

Therefore, there is a trade-off when selecting the IF for a radio receiver between image-frequency rejection and IF gain and stability.
The image-frequency rejection ratio (IFRR) is a numerical measure of the ability of a preselector to reject the image frequency.

For a single-tuned preselector, the ratio of its gain at the desired RF to the gain at the image frequency is IFRR

\[ IFRR = \sqrt{1 + Q^2 \rho^2}, \quad \rho = \frac{f_{IM}}{f_{RF}} - \frac{f_{RF}}{f_{IM}} \]
Example

For a citizens-band receiver using high-side injection with an RF carrier of 27 MHz and an IF center frequency of 455 kHz, determine:

- Local oscillator frequency.
- Image frequency.
- IFRR for a preselector Q of 100.
- Preselector Q required to achieve the same IFRR (~23dB) as that achieved for an RF carrier of 600 kHz.

Solution

\[ f_{LO} = 27\, \text{MHz} + 455\, \text{kHz} = 27.455\, \text{MHz} \]

\[ f_{IM} = 27.455\, \text{MHz} + 455\, \text{kHz} = 27.91\, \text{MHz} \]

\[ IFRR = 6.7 \text{ or } 16.5\, \text{dB} \]

\[ Q = \frac{\sqrt{IFRR^2} - 1}{\rho} = 3187 \]
Conclusion

It can be seen that the higher the RF carrier, the more difficult it is to prevent the image frequency from entering the receiver.

For the same IFRR, the higher RF carriers require a much higher-quality preselector filter.

![Frequency spectrum for Example 4-8](image)

Double Spotting

*Double spotting* occurs when a receiver picks up the same station at two nearby points on the receiver tuning dial.

One point is the desired location and the other point is called the *spurious point*.

- Double spotting is caused by poor front-end selectivity or inadequate image-frequency rejection.
- Double spotting is harmful because weak stations can be overshadowed by the reception of a nearby strong station.
- Double spotting may be used to determine the intermediate frequency of an unknown receiver because the spurious point on the dial is precisely two times the IF center frequency below the correct receive frequency.
**RF Amplifier Circuits**

An RF amplifier is a high-gain, low-noise, tuned amplifier that, when used, is the first active stage encountered by the received signal.

The primary purposes of an RF stage are selectivity, amplification, and sensitivity.

Therefore, the following characteristics are desirable for RF amplifiers:
- Low thermal noise
- Low noise figure
- Moderate to high gain
- Low intermodulation and harmonic distortion (that is, linear operation)
- Moderate selectivity
- High image-frequency rejection ratio

---

**RF Amplifier Circuits**

Two of the most important parameters for a receiver are *amplification* and *noise figure*, which both depend on the RF stage.

An AM demodulator (or detector as it is sometimes called) detects amplitude variations in the modulated wave and converts them to amplitude changes in its output.

Consequently, amplitude variations that were caused by noise are converted to erroneous fluctuations in the demodulator output, and the quality of the demodulated signal is degraded.
RF Amplifier Circuits

The more gain that a signal experiences as it passes through a receiver, the more pronounced are the amplitude variations at the demodulator input, and the less noticeable are the variations caused by noise.

The narrower the bandwidth is the less noise propagated through the receiver and, consequently, the less noise demodulated by the detector.

Noise voltage is directly proportional to the square root of the temperature, bandwidth, and equivalent noise resistance.

Therefore, if these three parameters are minimized, the thermal noise is reduced.

The temperature of an RF stage can be reduced by artificially cooling the front end of the receiver with air fans or even liquid helium in the more expensive receivers.

The bandwidth is reduced by using tuned amplifiers and filters, and the equivalent noise resistance is reduced by using specially constructed solid-state components for the active devices.

Noise figure is essentially a measure of the noise added by an amplifier.

Therefore, the noise figure is improved (reduced) by reducing the amplifier's internal noise.
Intermodulation and harmonic distortion are both forms of nonlinear distortion that increase the magnitude of the noise figure by adding correlated noise to the total noise spectrum.

The more linear an amplifier's operation is, the less nonlinear distortion produced, and the better the receiver's noise figure.

The image-frequency reduction by the RF amplifier combines with the image frequency reduction of the preselector to reduce the receiver input bandwidth sufficiently to help prevent the image frequency from entering the mixer/converter stage.

Consequently, moderate selectivity is all that is required from the RF stage.

---

**Bipolar Transistor RFA**

![Bipolar Transistor RFA Circuit Diagram]
A cascoded amplifier offers higher gain and less noise than conventional cascaded amplifiers.
The purpose of the mixer/converter stage is to down-convert the incoming radio frequencies to intermediate frequencies. This is accomplished by mixing the RF signals with the local oscillator frequency in a nonlinear device. In essence, this is heterodyning.

A mixer is a nonlinear amplifier similar to a modulator, except that the output is tuned to the difference between the RF and local oscillator frequencies.

![Mixer/Converter Block Diagram](image-url)
Self-Excited Mixer

Separately Excited Mixer
A single-diode mixer is inefficient because it has a net loss.

However, a diode mixer is commonly used for the audio detector in AM receivers and to produce the audio subcarrier in television receivers.

Balanced mixers are one of the most important circuits used in communications systems today.

Balanced mixers are also called balanced modulators, product modulators, and product detectors.

The phase detectors used in phase-locked loops are balanced modulators.

Balanced mixers are used extensively in both transmitters and receivers for AM, FM, and many of the digital modulation schemes, such as PSK and QAM.

Balanced mixers have two inherent advantages over other types of mixers: noise reduction and carrier suppression.
Intermediate frequency (IF) amplifiers are relatively high-gain tuned amplifiers that are very similar to RF amplifiers, except that IFA operate over a relatively narrow, fixed frequency band. Consequently, it is easy to design and build IFA that are stable, do not radiate, and are easily neutralized. Because IFA operate over a fixed frequency band, successive amplifiers can be inductively coupled with double-tuned circuits (with double-tuned circuits, both the primary and secondary sides of the transformer are tuned tank circuits). Therefore, it is easier to achieve an optimum (low) shape factor and good selectivity. Most of a receiver's gain and selectivity is achieved in the IFA section. An IF stage generally has between two and five IF amplifiers.

Three Stage IFA

![Three Stage IFA Circuit Diagram](image)
**Inductive Coupling**

*Inductive or transformer coupling* is the most common technique used for coupling IF amplifiers.

With inductive coupling, voltage that is applied to the primary windings of a transformer is transferred to the secondary windings.

The proportion of the primary voltage that is coupled across to the secondary depends on the number of turns in both the primary and secondary windings (the turns ratio), the amount of *magnetic flux* in the primary winding, the *coefficient of coupling*, and the speed at which the flux is changing (angular velocity).

---

**Coefficient of Coupling**

\[ k = \frac{\phi_s}{\phi_p}; \ \phi_s \text{ and } \phi_p \text{ primary and secondary flux.} \]

- If all the flux produced in the primary windings cuts through the secondary windings, the coefficient of coupling is 1.
- If none of the primary flux cuts through the secondary windings, the coefficient of coupling is 0.
- A coefficient of coupling of 1 is nearly impossible to attain unless the two coils are wound around a common high-permeability iron core.
- Typically, the coefficient of coupling for standard IF transformers is much less than 1.
**Single-Tuned Transformers**

**Double-Tuned Transformers**

\[ k_{\text{critical}} = \frac{1}{\sqrt{Q_s Q_p}}, \quad k_{\text{optimal}} = 1.5 k_{\text{critical}} \]
The overall bandwidth of $n$ single-tuned stages is

$$B_n = B_1 \sqrt{2^{1/n} - 1}$$

The bandwidth for $n$ double-tuned stages is

$$B_{ndt} = B_{ldt} \left[2^{1/n} - 1\right]^{1/4},$$

$$B_{ldt} = k f_0$$
**Example**

Determine the overall bandwidth for
- Two single-tuned amplifiers each with a bandwidth of 10 kHz.
- Three single-tuned amplifiers each with a bandwidth of 10 kHz.
- Four single-tuned amplifiers each with a bandwidth of 10 kHz.
- A double-tuned amplifier with optimum coupling, a critical coupling of 0.02, and a resonant frequency of 1 MHz.

Repeat this for the double-tuned amplifier.

**Solution**

\[
B_2 = 10 \text{kHz} \left(\sqrt{2^{1/2} - 1}\right) = 6436 \text{ Hz}
\]
\[
B_3 = 10 \text{kHz} \left(\sqrt{2^{1/3} - 1}\right) = 5098 \text{ Hz}
\]
\[
B_4 = 10 \text{kHz} \left(\sqrt{2^{1/4} - 1}\right) = 4350 \text{ Hz}
\]

Double tuned:
\[
k_{optimal} = 1.5 \cdot 0.02 = 0.03
\]
\[
B_{dt} = 0.03 \cdot 1 \text{MHz} = 30 \text{ kHz}
\]
\[
B_{2dt} = 24.067 \text{ Hz}
\]
\[
B_{3dt} = 21.420 \text{ kHz}
\]
\[
B_{4dt} = 19.786 \text{ kHz}
\]
AM Detector Circuits

The function of an AM detector is to demodulate the AM signal and recover or reproduce the original source information.

The recovered signal should contain the same frequencies as the original information signal and have the same relative amplitude characteristics.

The AM detector is sometimes called the second detector, with the mixer/converter being the first detector because it precedes the AM detector.

AM Demodulator
AM Demodulator

It is a simple noncoherent AM demodulator, which is commonly called a peak detector.

Because a diode is a nonlinear device, nonlinear mixing occurs in $D_1$ when two or more signals are applied to its input. Therefore, the output contains the original input frequencies, their harmonics and their cross products.

If a 300-kHz carrier is amplitude modulated by a 2-kHz sine wave, the modulated wave is made up of a lower side frequency, carrier, and upper side frequency of 298 kHz, 300 kHz, and 302 kHz, respectively.

If the resultant signal is the input to the AM detector, the output will comprise the three input frequencies, the harmonics of all three frequencies, and the cross products of all possible combinations of the three frequencies and their harmonics.

Because the RC network is a low-pass filter, only the difference frequencies are passed on to the audio section. Therefore, the output is simply

\[
V_{\text{out}} = 300 - 298 = 2\, \text{kHz}
\]

\[
= 302 - 300 = 2\, \text{kHz}
\]

\[
= 302 - 298 = 4\, \text{kHz}
\]

Consequently, for practical purposes, the original modulating signal (2 kHz) is the only component that is contained in the output of the peak detector.
AM Demodulator

- The input and output waveforms for a peak detector with various percentages of modulation.
- With no modulation, a peak detector is simply a filtered half-wave rectifier and the output voltage is approximately equal to the peak input voltage minus the 0.3 V.
- As the percent modulation changes, the variations in the output voltage increase and decrease proportionately; the output waveform follows the shape of the AM envelope.
- However, regardless of whether modulation is present or not the average value of the output voltage is approximately equal to the peak value of the unmodulated carrier.

Detector Distortion

- When successive positive peaks of the detector input waveform are increasing, it is important that the capacitor hold its charge between peaks (that is a relatively long \( RC \) time constant is necessary).
- However, when the positive peaks are decreasing in amplitude, it is important that the capacitor discharge between successive peaks to a value less than the next peak (a short \( RC \) time constant is necessary).
- Obviously, a trade-off between a long- and a short-time constant is in order.
Detector Distortion

If the RC time constant is too short, the output waveform resembles a half-wave rectified signal. This is sometimes called **rectifier distortion**.

If the RC time constant is too long, the slope of the output waveform cannot follow the trailing slope of the envelope. This type of distortion is called **diagonal clipping**.

The RC network following the diode in a peak detector is a low-pass filter.

The slope of the envelope depends on both the modulating signal frequency and the modulation coefficient \(m\).

Therefore, the maximum slope (fastest rate of change) occurs when the envelope is crossing its zero axis in the negative direction.
Detector Distortion

- The highest modulating signal frequency that can be demodulated by a peak detector without attenuation is given as
- For 100% modulation, the numerator goes to zero, which essentially means that all modulating signal frequencies are attenuated as they are demodulated.
- Typically, the modulating signal amplitude in a transmitter is limited or compressed such that approximately 90% modulation is the maximum that can be achieved.
- For 70.7% modulation is

\[ f_{m(max)} = \frac{\sqrt{1/m^2 - 1}}{2\pi RC} \]

Automatic Gain Control Circuit

- The automatic gain control circuit monitors the received signal level and sends a signal back to the RF and IF amplifiers to adjust their gain automatically.
- AGC is a form of degenerative or negative feedback.
- The purpose of AGC is to allow a receiver to detect and demodulate, equally well, signals that are transmitted from different stations whose output power and distance from the receiver vary.
  - For example, an AM radio in a vehicle does not receive the same signal level from all the transmitting stations in the area or, for that matter, from a single station when the automobile is moving.
- The AGC circuit produces a voltage that adjusts the receiver gain and keeps the IF carrier power at the input to the AM detector at a relatively constant level.
- The AGC circuit is not a form of automatic volume control (AVC); AGC is independent of modulation and totally unaffected by normal changes in the modulating signal amplitude.
Receiver with AGC

Simple AGC Circuit
Delayed AGC

Simple AGC is used in most inexpensive broadcast-band radio receivers.

However, with simple AGC, the AGC bias begins to increase as soon as the received signal level exceeds the thermal noise of the receiver.

Consequently, the receiver becomes less sensitive (this is sometimes called automatic desensing).

Delayed AGC prevents the AGC feedback voltage from reaching the RF or IF amplifiers until the RF level exceeds a predetermined magnitude.

Once the carrier signal has exceeded the threshold level, the delayed AGC voltage is proportional to the carrier signal strength.

With delayed AGC the receiver gain is unaffected until the AGC threshold level is exceeded, whereas with simple AGC the receiver gain is immediately affected.

Delayed AGC is used with more sophisticated communications receivers.

AGC Response Characteristics

![AGC Response Characteristics](image)

FIGURE 4-30 Automatic gain control (AGC) b) response characteristics (RF output voltage vs. RF input signal level)
An inherent problem with both simple and delayed AGC is the fact that they are both forms of post-AGC (after-the-fact) compensation. With post-AGC, the circuit that monitors the carrier level and provides the AGC correction voltage is located after the IF amplifiers; therefore, the simple fact that the AGC voltage changed indicates that it may be too late (the carrier level has already changed and propagated through the receiver). Therefore, neither simple nor delayed AGC can accurately compensate for rapid changes in the carrier amplitude. Forward AGC is similar to conventional AGC except that the receive signal is monitored closer to the front end of the receiver and the correction voltage is fed forward to the IF amplifiers. Consequently, when a change in signal level is detected, the change can be compensated for in succeeding stages.
The purpose of a squelch circuit is to quiet a receiver in the absence of a received signal. If an AM receiver is tuned to a location in the RF spectrum where there is no RF signal, the AGC circuit adjusts the receiver for maximum gain. Consequently, the receiver amplifies and demodulates its own internal noise. This is the familiar crackling and sputtering heard on the speaker in the absence of a received carrier. In domestic AM systems, each station is continuously transmitting a carrier regardless of whether there is any modulation or not. Therefore, the only time the idle receiver noise is heard is when tuning between stations.

However, in two-way radio systems, the carrier in the transmitter is generally turned off except when a modulating signal is present. Therefore, during idle transmission times, a receiver is simply amplifying and demodulating noise. A squelch circuit keeps the audio section of the receiver turned off or muted in the absence of a received signal (the receiver is squelched). A disadvantage of a squelch circuit is weak RF signals will not produce an audio output.
Double-Conversion AM Receiver

- For good image-frequency rejection, a relatively high intermediate frequency is desired.
- However, for high-gain selective amplifiers that are stable and easily neutralized, a low intermediate frequency is necessary.
- The solution is to use two intermediate frequencies.
- The first IF is a relatively high frequency for good image-frequency rejection, and the second IF is a relatively low frequency for easy amplification.
- The first IF is 10.625 MHz, which pushes the image frequency 21.25 MHz away from the desired RF.
- The first IF is immediately down-converted to 455 kHz and fed to a series of high-gain IF amplifiers.
Double-Conversion AM Receiver

FIGURE 4-34 Double-conversion AM receiver

Double-Conversion AM Receiver

FIGURE 4-35 Filtering requirements for the double-conversion AM receiver shown in Figure 4-24
Receiver gains and losses

FIGURE 4-36  Receiver gains and losses