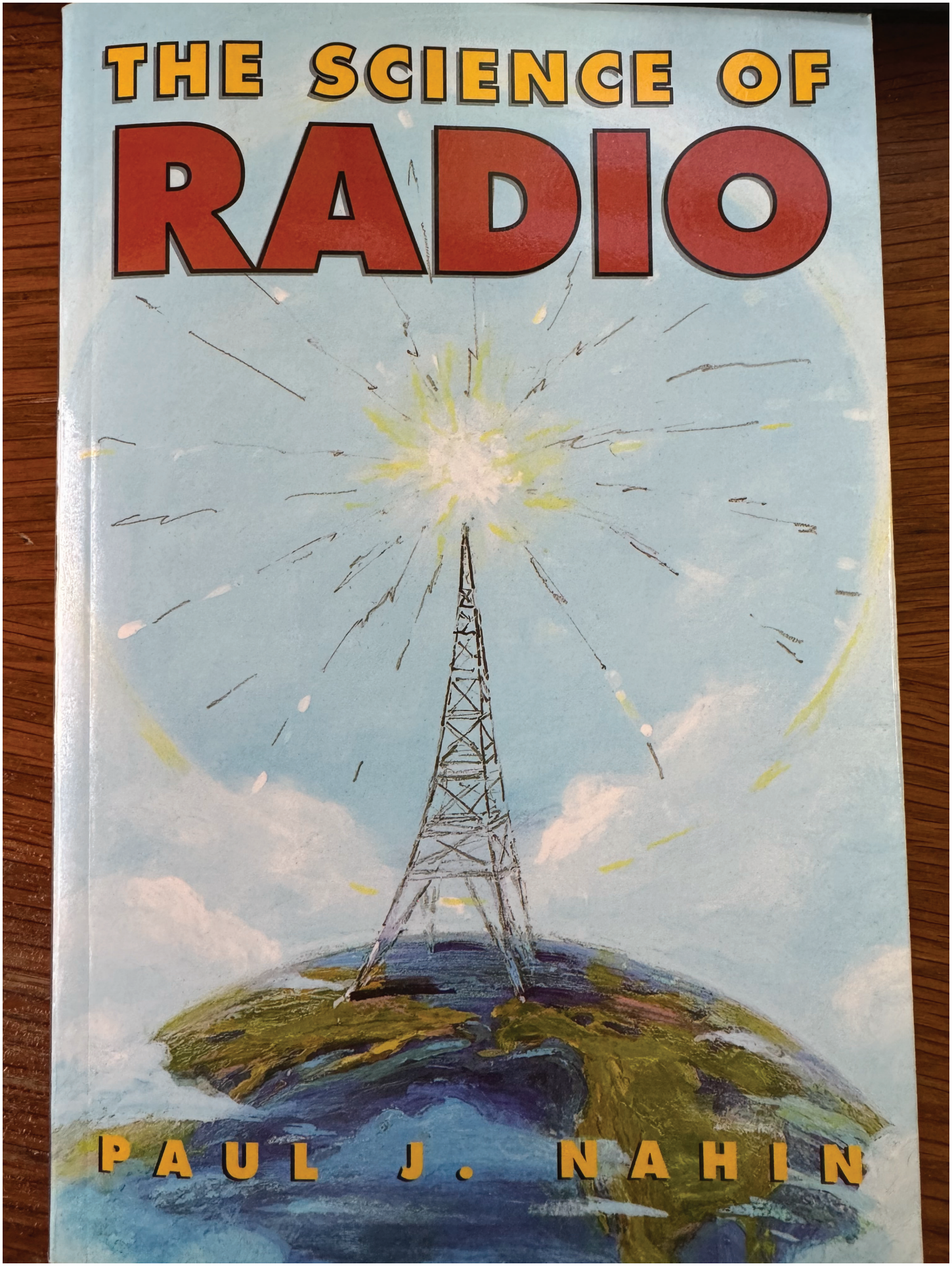
THE SCIENCE OF RADIO



PAUL J. NAHIN

MEDIATEK-1040
7,483,686

Copyright © 1990 by Raymond A. Serway. Copyright 1986, 1982 by Raymond A. Serway.
1992 Updated Printing

All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher.

Requests for permission to make copies of any part of the work should be mailed to Permissions Department, Harcourt, Brace, Jovanovich, Publishers, 8th Floor, Orlando, Florida 32887.

Text Typeface: Caledonia
Compositor: Progressive Typographers, Inc.
Acquisitions Editor: John Vondeling
Developmental Editor: Ellen Newman
Managing Editor: Carol Field
Senior Project Manager: Sally Kusch
Copy Editor: Will Eaton
Manager of Art and Design: Carol Bleistine
Art and Design Coordinator: Caroline McGowan
Text Designer: Edward A. Butler
Cover Designer: Lawrence R. Didona
Text Artwork: Rolin Graphics
Production Manager: Jay Lichty
Layout Artist: Tracy Baldwin
Director of EDP: Tim Frelick
Marketing Manager: Marjorie Waldron

Cover Credit: Beams of red, green, and blue light strike a D-shaped plastic lens and are bent toward a focus. Photography by S. Schwartzenberg/© The Exploratorium.

Printed in the United States of America

Physics for Scientists & Engineers, 3rd edition, Updated Printing

0-03-096026-6

Library of Congress Catalog Card Number: 89-043325

345 032 98765432

Mathematics of AM Sidebands

You can understand the theoretical reason behind Fessenden's failed attempt to send acceptable quality voice by spark-gap transmission once it is understood that *amplitude modulation* (AM) spreads energy over an interval of frequencies. If you start with a constant amplitude sine wave of a given frequency, then you have what is called a *pure tone*. But as soon as you begin to vary the amplitude of that pure tone you create *additional* tones (sidetones), both higher and lower than the original frequency. This fact immediately establishes that amplitude modulation is not the result of a linear, time-invariant system (see Appendix B). In such systems, the output frequencies are limited to the input frequencies (with amplitudes perhaps modified as a function of frequency, as discussed in Appendix C).

In AM radio, however, there are frequencies present at the output of the transmitter that are *not* present at the input. These new frequencies are created by either a nonlinearity, or a time-varying process, inside the transmitter. We can establish this fact very quickly (and I will before the end of this chapter), but it is astonishing to read the older technical literature and see how even some of the "big names" in early radio simply didn't believe it (even after being shown why they were wrong!)

Still interesting reading today is the essay by Sir John Ambrose Fleming, "The 'Wave Band' Theory of Wireless Transmission," *Nature*, January 18, 1930, pp. 92–93. Fleming (a Fellow of the Royal Society who had studied at Cambridge under the direction of the great Maxwell, himself, and was the inventor of the vacuum tube diode to be discussed in Chapter 8), called the elementary mathematical analysis in this chapter "a kind of mathematical fiction [that] does not correspond to any reality in nature." For Fleming to write this in 1930 was astonishing because the observation of the sidetone phenomenon had occurred long before in 1875. That year, the American physicist Alfred Mayer (1836–1897) *heard* them in an acoustical experiment in which he mechanically interrupted the sound of a vibrating tuning fork. Later, in 1894 when Lord Rayleigh published the second edition of his enormously influential *Theory of Sound* (for more on Rayleigh see Chapter 13), he took notice of Mayer's experiment and explained it mathematically. In light of this it isn't surprising that Fleming's prompted a flood of replies, all declaring him to be dead wrong. But he unconvinced (see *Nature*, February 8, pp. 198–199; February 22, pp. 22 1, pp. 306–307). What *really* makes Fleming's essay hard to un

that, at the time he wrote it, the first commercial single-sideband (SSB) AM radio link between New York City and London had been in operation for 3 yr! According to Fleming's position, such a radio (discussed in great detail in Chapter 20) simply couldn't work, and the fact that it did work apparently never caused him to reconsider his position. Old ideas die hard.

As mentioned in the last chapter, the problem with spark gaps is that, even before you attempt to modulate their transmissions with voice amplitude variations, their signals are *created* with AM inherently present because the antenna circuit oscillations are exponentially damped. The conclusion then (if you accept the premise of the first paragraph), is that spark-gap radio transmitters are *sloppy*, literally spilling energy into frequencies that can be considerably different from the nominal resonant frequency of the antenna circuit. Later, in Chapter 11, I'll work through the precise mathematical details of the sloppy radiation behavior so characteristic of spark-gap transmitters, which was a flaw sufficiently awful as to be a fatal one for their future.

Suppose, then, contrary to the state of technology even as the twentieth century began, that we have a stable source of constant amplitude, radio frequency oscillations (at the frequency $\omega_c=2\pi f_c$ radians/second). The frequency f_c (in hertz) is called the *carrier frequency*. For commercial AM broadcast radio, f_c is in the interval 540–KHz 1,600 KHz. I'll write the carrier as simply

$$c(t)=\cos(\omega_c t),$$

with the assumption that the carrier has unit peak amplitude. That is, whatever the actual peak amplitude of the carrier may be, it will serve as our amplitude reference.

Next, suppose we have a very simple signal we wish to transmit — the sound of a person whistling at the frequency $\omega_m=2\pi f_m$. By experience, we know f_m is at most on the order of a few kilohertz, and so $f_m \ll f_c$. Such signals as a person whistling (or talking), with frequency content varying from dc (hertz) to some maximum upper frequency (in the low kilohertz range for human speech) are called *baseband* signals. If we denote the amplitude of this "whistle signal" by A_m, then the message signal is

$$m(t)=A_m\cos(\omega_m t).$$

In fact, because $c(t)$ and $m(t)$ are certainly not synchronized, we should actually toss in an arbitrary phase angle, ϕ, to allow for the (almost certain) possibility that $m(t)$ and $c(t)$ are not simultaneously at their peaks at $t=0$. That is, let's more generally write

$$c(t)=\cos(\omega_c t+\phi) \quad \text{and} \quad m(t)=A_m\cos(\omega_m t).$$

Now, as shown in Figure 6.1, imagine that we apply the carrier signal to a multiplier, along with the message signal (added to a dc shift of unity) on the other multiplier input. Then writing the multiplier's output signal as $r(t)$, i.e., the signal to the radiated,

that, at the time he wrote it, the first commercial single-sideband (SSB) AM radio link between New York City and London had been in operation for 3 yr! According to Fleming's position, such a radio (discussed in great detail in Chapter 20) simply couldn't work, and the fact that it did work apparently never caused him to reconsider his position. Old ideas die hard.

As mentioned in the last chapter, the problem with spark gaps is that, even before you attempt to modulate their transmissions with voice amplitude variations, their signals are *created* with AM inherently present because the antenna circuit oscillations are exponentially damped. The conclusion then (if you accept the premise of the first paragraph), is that spark-gap radio transmitters are *sloppy*, literally spilling energy into frequencies that can be considerably different from the nominal resonant frequency of the antenna circuit. Later, in Chapter 11, I'll work through the precise mathematical details of the sloppy radiation behavior so characteristic of spark-gap transmitters, which was a flaw sufficiently awful as to be a fatal one for their future.

Suppose, then, contrary to the state of technology even as the twentieth century began, that we have a stable source of constant amplitude, radio frequency oscillations (at the frequency $\omega_c=2\pi f_c$ radians/second). The frequency f_c (in hertz) is called the *carrier frequency*. For commercial AM broadcast radio, f_c is in the interval 540–KHz 1,600 KHz. I'll write the carrier as simply

$$c(t)=\cos(\omega_c t),$$

with the assumption that the carrier has unit peak amplitude. That is, whatever the actual peak amplitude of the carrier may be, it will serve as our amplitude reference.

Next, suppose we have a very simple signal we wish to transmit — the sound of a person whistling at the frequency $\omega_m=2\pi f_m$. By experience, we know f_m is at most on the order of a few kilohertz, and so $f_m \ll f_c$. Such signals as a person whistling (or talking), with frequency content varying from dc (hertz) to some maximum upper frequency (in the low kilohertz range for human speech) are called *baseband* signals. If we denote the amplitude of this "whistle signal" by A_m, then the message signal is

$$m(t)=A_m\cos(\omega_m t).$$

In fact, because $c(t)$ and $m(t)$ are certainly not synchronized, we should actually toss in an arbitrary phase angle, ϕ, to allow for the (almost certain) possibility that $m(t)$ and $c(t)$ are not simultaneously at their peaks at $t=0$. That is, let's more generally write

$$c(t)=\cos(\omega_c t+\phi) \quad \text{and} \quad m(t)=A_m\cos(\omega_m t).$$

Now, as shown in Figure 6.1, imagine that we apply the carrier signal to a multiplier, along with the message signal (added to a dc shift of unity) on the other multiplier input. Then writing the multiplier's output signal as $r(t)$, i.e., the signal to the radiated, we have

CHAPTER 16

Multiplying by Squaring and Filtering

For radio transmitters to work, we need to shift baseband energy up to rf. Our previous work has shown you how to do that—simply multiply the baseband signal $m(t)$ by $\cos(\omega_c t)$. The spectrum of $m(t)$, which describes how the energy of $m(t)$ is distributed in frequency (Rayleigh's energy theorem), then shifts both up and down by ω_c, the so-called *carrier frequency*. Going next in the opposite direction, for radio receivers to work we obviously need to shift the up-shifted spectrum of $m(t)$ back down to baseband where we can hear it. Again, multiplication is the way to do it (I'll discuss this process in more detail in Section Four). Accurate analog multipliers that work at radio frequencies are expensive, difficult-to-build devices, however, and some great ingenuity has gone into developing *indirect* ways of performing multiplication (and, hence, spectrum shifting). You will be able to understand how these clever circuits work because you have (haven't you?) worked your way through the mathematics of the Fourier transform in the previous section. The Fourier transform is the key to unlocking the physics of what is happening in these circuits.

To begin, let's drop our sights a bit, and instead of directly building a multiplier let's study the behavior of the summer/squarer/filter shown in Figure 16.1. The output of the squarer is $m^2(t)+2m(t)\cos(\omega_c t)+\cos^2(\omega_c t)$, which includes the desired product term $m(t)\cos(\omega_c t)$. It also includes, seemingly to our misfortune, two other terms. The astonishing fact is, however, that it is possible to arrange matters so that the spectrum of the product term is distinct from the spectra of the other terms. Thus, we can apply the output of the squarer to a suitably designed bandpass filter which will pass only the energy of the product term; the *total* circuitry in Figure 16.1 (including the filter) is therefore a multiplier.

This solution for how to build a multiplier is not quite complete, of course, because it leaves us with the obvious problem of how to build a *squarer*. Intuitively, however, we might expect this to be a simpler problem, as squaring is a special, less complicated process than multiplying. After all, a squarer has just *one* input while a multiplier has two, i.e., a squarer can be built from a single multiplier (simply apply the same input signal to both multiplier inputs). In the next chapter I'll discuss the summing/squaring operation in some detail, but for now let's verify that the filtered output of the circuit

CHAPTER 19

Synchronous Demodulation and Its Problems

The previous section has shown us how, given a baseband signal $m(t)$, we can shift its spectrum up to rf by multiplying $m(t)$ by $\cos(\omega_c t)$. To form the AM signal that can be detected by the envelope detector circuit discussed in Chapter 6, we must next add to this product a constant carrier term (see Figure 6.1). This gives what is called a "double sideband, large carrier," or DSB-LC, AM signal. I discussed in Chapter 6 why a *strong* carrier term is necessary for the proper operation of the envelope detector —the output of the detector is proportional to $m(t)$ if a sufficiently strong carrier is present, but if there is no carrier or even if it is present but too weak, then the detector produces a signal proportional to $|m(t)|$ (see Figure 6.2) which is unintelligible. However, I also discussed how the insertion of a carrier at the transmitter wastes energy. A natural question, then, is to ask if there isn't *some* way we can extract (detect) $m(t)$ from $m(t)\cos(\omega_c t)$ without having to assume the transmitter has added in a carrier term? Such a signal is called "double sideband, suppressed carrier," or DSB-SC, and the answer is *yes*. The following discussion shows how, and also why commercial AM radio does *not* use DSB-SC.

> It is an easy and educational experiment to show that $|m(t)|$ is unintelligible. $|m(t)|$ is a full-wave rectification (see any book on electronics for the elementary four diode circuit of such a rectifier) of $m(t)$. So, simply take the wires leading to the loudspeaker of a radio and insert the rectifier. When you hear the result, you'll agree it's unintelligible!

To recover $m(t)$ from a DSB-SC signal is, on paper, actually a trivial problem. To *demodulate* the DSB-SC signal $m(t)\cos(\omega_c t)$ we'll just use the modulation or heterodyne theorem again, the same theorem we used to shift the spectrum of $m(t)$ up to rf in the first place. Thus, if we multiply $m(t)\cos(\omega_c t)$ by $\cos(\omega_c t)$ we'll shift the spectrum of the DSB-SC signal both *up* (to be centered around $\pm 2\omega_c$) and *down* (to be centered around $\omega=0$). But a spectrum centered around $\omega=0$ is just the original baseband

spectrum and is just what we want. The energy at and around $\pm 2\omega_c$ is, on the other hand, at a very high frequency, over at least a megahertz for AM radio, and can be safely ignored (if applied to the input terminals of a loudspeaker, for example, the speaker would reproduce the baseband spectrum—which is at *audio* frequencies—but, because of mechanical inertia, would be completely unresponsive to energy at rf, i.e., the speaker is a mechanical low-pass filter).

The problem with this approach is that it assumes that the *receiver*, which may be separated from the transmitter by tens, hundreds, even thousands of miles, has the signal $\cos(\omega_c t)$ available. This is a problem because for this approach to work, it is necessary for the receiver's carrier frequency sinusoid to be *precisely* at frequency ω_c *and* to be *phase coherent* with the transmitter's sinusoid, i.e., the receiver must *locally* generate $\cos(\omega_c t)$ and not $\cos\{(\omega_c + \Delta\omega)t + \theta\}$. That is, the receiver's local oscillator must be synchronized with the transmitter's carrier, and so this approach is called *synchronous demodulation*. To see why synchronization is required, I'll calculate what happens if there are phase or frequency mismatches between transmitter and receiver.

Consider first a phase mismatch, only. Receiving $m(t)\cos(\omega_c t)$ the receiver then multiplies by $\cos(\omega_c t + \theta)$ to give

$$m(t)\cos(\omega_c t)\cos(\omega_c t + \theta) = \frac{1}{2}\, m(t)[\cos(\theta) + \cos(2\omega_c t)],$$

from which a low-pass filter can recover a term proportional to $m(t)\cos(\theta)$. Phase mismatch appears therefore as the amplitude attenuation factor $\cos(\theta)$, which is not serious until (if) θ approaches 90°. At $\theta = 90°$, however, the filter output signal vanishes! Even a phase mismatch as big as $\theta = 90°$ represents precise timing, i.e., in the middle of the AM radio band, at one megahertz, each period of 1 μsec represents 360° and so a 90° mismatch means the transmitter and the receiver are only 250 *nano*seconds out of time alignment! Well, you might counter, except for the $\theta = 90°$ case we can simply negate the amplitude attenuation by increasing the receiver's amplification. The problem with that is that θ is generally an unknown *function of time*, because the received signal has made its way through space via a time-varying path (e.g., consider the case of a receiver on an airplane that isn't flying in a circle around the transmitter, or radio signals scattered back down to earth from an altitude-fluctuating ionosphere).

As bad as phase mismatch is for synchronous demodulation, frequency mismatch is arguably (far) worse. To see this, suppose after receiving $m(t)\cos(\omega_c t)$ the receiver multiplies by $\cos\{(\omega_c + \Delta\omega)\}$ to give

$$m(t)\cos(\omega_c t)\cos\{(\omega_c + \Delta\omega)t\} = \frac{1}{2}\, m(t)[\cos\{(\Delta\omega)t\} + \cos\{(2\omega_c + \Delta\omega)t\}],$$

from which a low-pass filter removes a term proportional to $m(t)\cos\{(\Delta\omega)t\}$. That is, a frequency mismatch appears as a *time-varying* amplitude factor, and this factor is significant for even very small mismatches. For example, if $f_c = 1$ MHz and Δf

1 Hz (an error of just one part in a *million*!) then the amplitude of the low-pass filter output will vary from zero to maximum and back to zero *twice a second*. This would be a catastrophic effect, rendering human speech unintelligible (as you can simulate for yourself simply by rolling the volume control on a radio back and forth twice a second).

The conclusion from all this is that to successfully demodulate DSB-SC we need some way for the receiver to *accurately* reconstruct the carrier. And there *is* a way, as shown in Figure 19.1. The received signal $m(t)\cos(\omega_c t)$ is immediately squared, giving $(1/2)\ m^2(t) + (1/2)\ m^2(t)\cos(2\omega_c t)$. The first term is a baseband signal (limited to $|\omega| < 2\omega_m$), and the second term is that same baseband signal shifted up to rf with its spectrum centered on $\omega = \pm\ 2\omega_c$. The crucial observation at this point is that this *ensures* us there will be energy *at* $\omega = 2\omega_c$, because $m^2(t)$ is *certain* to have a non-zero d-c value over its duration (because a squared quantity is never negative). For more on this, see Problem 19.1.

Thus, a (*very*) narrow bandwidth bandpass filter centered at $\omega = 2\omega_c$ will have an output at twice the carrier frequency. The *narrower* the bandwidth of this filter the better. It is beyond the level of this book, but an exotic circuit called a *phase locked loop* (PLL) can be used for the task of reconstructing the carrier.[1] Such a circuit can lock onto both the frequency and the phase of the transmitter's carrier, and then *track* it if either one (or both) of the carrier frequency and phase change with time (which they both will surely do in real life). We can drop the filter output frequency down to carrier frequency using the regenerative frequency divider circuit discussed in Chapter 16. This reconstructed carrier signal is then used to heterodyne the original $m(t)\cos(\omega_c t)$ signal, and in particular to reproduce the baseband signal $m(t)$. There is an obvious problem with this receiver, however, as the circuit of Figure 19.1 cannot be said to be "simple." The use of DSB-SC is not attractive for commercial AM broadcast radio, with its need for *tens of millions* (or even more) of *cheap* receivers to make listening attractive to the potential customers of the advertisers whose checkbooks make the whole business possible. This, quite simply, is the technical reason for why synchronous demodulation isn't the way broadcast AM radio historically evolved.

At this point, with only its deficiencies discussed, you might well be wondering if synchronous demodulation is *ever* used? The answer is yes, under certain circumstances. I've already discussed the energy inefficiency of DSB-LC in Chapter 6, and in situations where it is important to put every watt of available power into the informa-

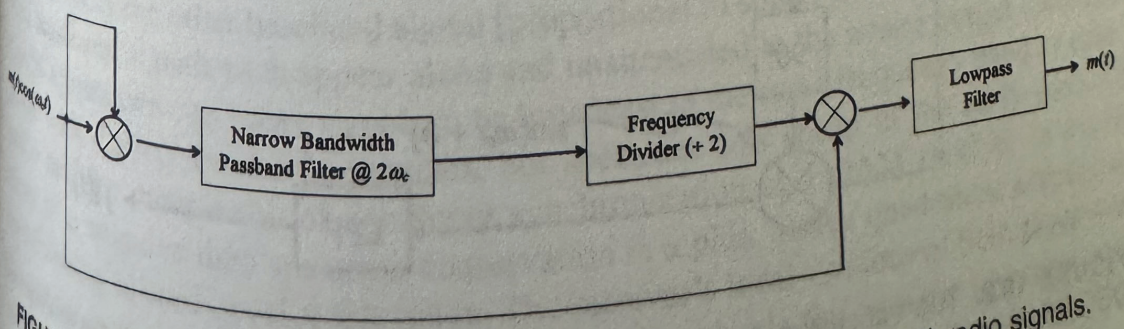


FIGURE 19.1. A receiver for double sideband, suppressed carrier (DSB-SC) radio signals.

tion signal (in the sidebands, not the carrier), e.g., in a *portable* transmitter, then DSB-SC may be worth the extra cost of the receivers. A second reason that we can't really pursue in this book because it requires knowledge of probabilistic signals, is that DSB-SC is much better able to withstand the effects of noise ("static") than is DSB-LC. And finally, I should point out that DSB-LC is, really, demodulated in an AM radio with a *synchronous* detector in a bit of disguise! This is so because the process of envelope detection requires the use of a diode that conducts only during one-half of each cycle of the input signal—and the timing of this condition is controlled by the carrier signal (which is *not* locally generated as it would be a true synchronous receiver).

Synchronous communication systems do have one extraordinary property that should be mentioned before we leave them. Consider the transmitter and receiver circuits shown in Figure 19.2, where it is assumed that the receiver has a perfect replica of the carrier available. As I'll show next, these circuits allow us to do the seemingly impossible—to transmit two entirely different baseband signals, $m_1(t)$ and $m_2(t)$, at the same carrier frequency *at the same time* without mutual interference! The key is to

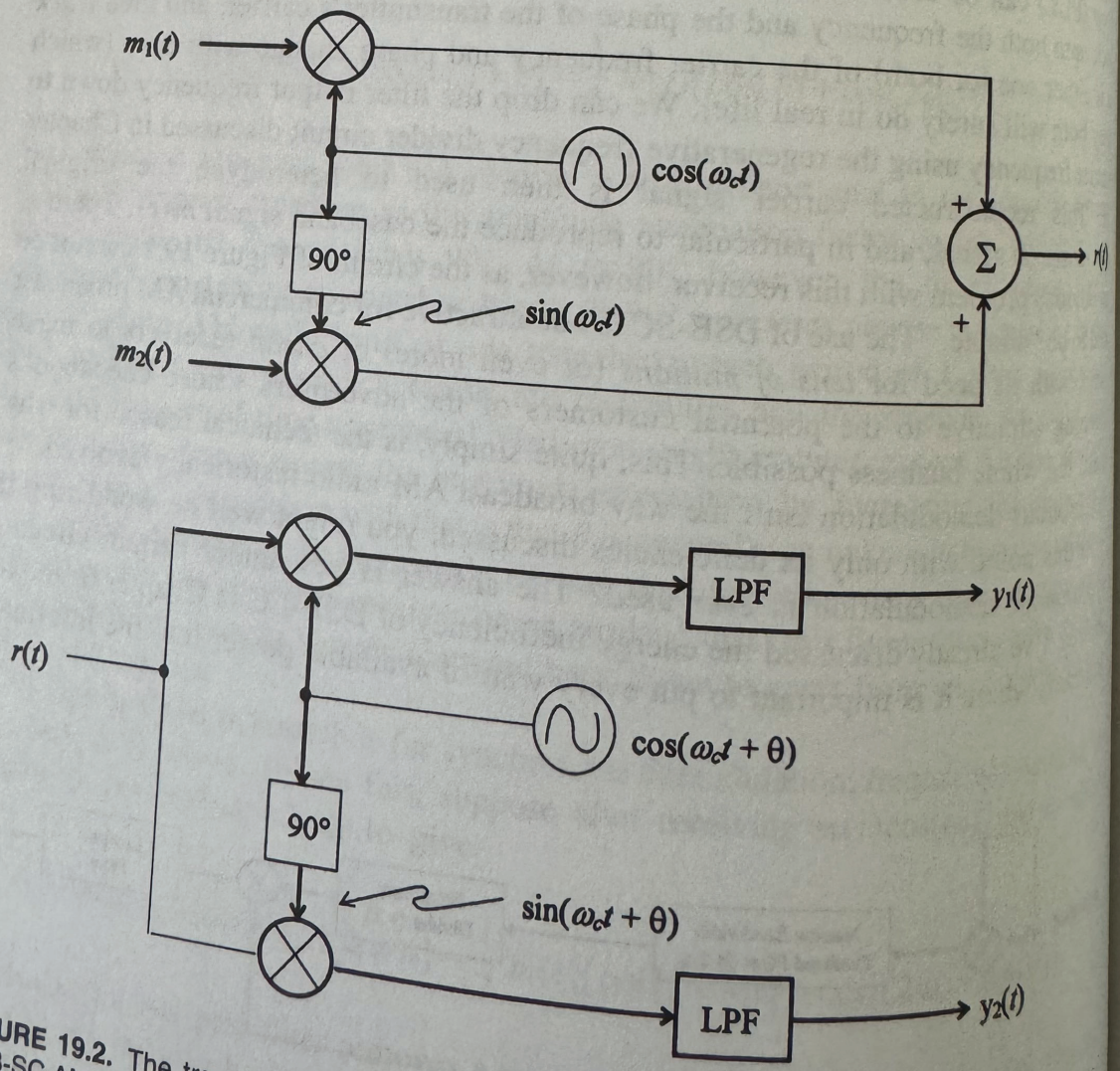


FIGURE 19.2. The transmitter (top) and receiver (bottom) for quadrature amplitude multiplexed DSB-SC AM radio, with a phase mismatch of θ at the receiver.

use two carrier signals at the same frequency but 90° out of phase, which is why this approach is called *quadrature amplitude multiplex* (QAM). From the figure we see that the transmitted signal is

$$r(t)=m_1(t)\cos(\omega_c t)+m_2(t)\sin(\omega_c t).$$

Now, following our earlier approach to DSB-SC, let's assume there is a phase mismatch of θ at the receiver. Thus,

$$s_1(t)=m_1(t)\cos(\omega_c t)\cos(\omega_c t+\theta)+m_2(t)\sin(\omega_c t)\cos(\omega_c t+\theta),$$

$$s_2(t)=m_1(t)\cos(\omega_c t)\sin(\omega_c t+\theta)+m_2(t)\sin(\omega_c t)\sin(\omega_c t+\theta).$$

Expanding these two expressions using the obvious trigonometric identities, we get

$$s_1(t)=\frac{1}{2}m_1(t)\{\cos(\theta)+\cos(2\omega_c t+\theta)\}+\frac{1}{2}m_2(t)\{-\sin(\theta)+\sin(2\omega_c t+\theta)\},$$

$$s_2(t)=\frac{1}{2}m_1(t)\{\sin(\theta)+\sin(2\omega_c t+\theta)\}+\frac{1}{2}m_2(t)\{\cos(\theta)-\cos(2\omega_c t+\theta)\},$$

or, after low-pass filtering, we find the two channels in the receiver produce output signals

$$y_1(t)=\frac{1}{2}[m_1(t)\cos(\theta)-m_2(t)\sin(\theta)],$$

$$y_2(t)=\frac{1}{2}[m_1(t)\sin(\theta)+m_2(t)\cos(\theta)].$$

If $\theta=0$ (the receiver is perfectly phase coherent with the transmitter) then we find $y_1(t)=(1/2)\,m_1(t)$ and $y_2(t)=(1/2)\,m_2(t)$, and so we have perfect separation of the two baseband signals. If $\theta \neq 0$, however, then each baseband signal is attenuated by a $\cos(\theta)$ factor in its own channel as well as suffers from "leakage" (called *cross-talk*) of the other baseband signal [proportional to $\sin(\theta)$]. If $\theta=90°$, in fact, the two baseband signals each appear alone and unattenuated in the *wrong* output channel! QAM allows twice as many signals to be broadcast in the same frequency band as does DSB with distinct carrier frequencies, but spectrum was cheap in the early days of radio and the complexity of achieving synchronization at a QAM receiver is much more of a negative than spectrum conservation is a plus. Today quadrature amplitude multiplexing is used in color television. The receiver's local versions of both phases of the carrier are kept synchronized with the transmitter's via the periodic insertion of a short burst of carrier signal (called the *color burst*) in the transmitter signal, $r(t)$.

of calling the "Rolls Royce" of radio receivers. It is the circuit found in all AM radios today, and what follows is how it works.

To begin, let me first quickly describe the spectrum environment in which the everyday AM superheterodyne operates. In commercial AM broadcast radio each station is assigned a "chunk" of spectrum 10 KHz wide. Centered on each such chunk is the station's assigned carrier frequency, which is selected from the interval 540 to 1600 KHz. Thus, a station with a carrier frequency of 640 KHz (like KFI, Los Angeles) can transmit its signals over the interval 635 to 645 KHz, and a station with a carrier frequency of 1270 KHz (like WTSN in Dover, New Hampshire) can transmit its signals over the interval 1265 to 1275 KHz. This allows each AM station to transmit a DSB-LC signal carrying a baseband signal with a bandwidth of up to 5 KHz. This is wide enough for good voice and music reproduction (but not for high-fidelity broadcasting, a market niche filled by Armstrong's development of wideband FM, which allows each FM station to use ± 75 KHz around its assigned carrier frequency—but that's another story).

By law, the holder of a commercial AM broadcast license in the United States must stay on his assigned carrier frequency with a maximum allowed deviation of ± 20 Hz. Carrier frequencies are spaced 10 KHz apart, but two stations operating in the same geographical area would not be assigned adjacent carrier frequencies. That could result in what is called *adjacent channel* interference, which is discussed later in this section. Very-high-power stations (like KFI, which transmits with 50,000 W) could conceivably be received anywhere in the entire country, and such stations originally received carrier frequency assignments duplicated *nowhere*. Such stations were called *clear channel* stations, not because they necessarily were always received clearly, but

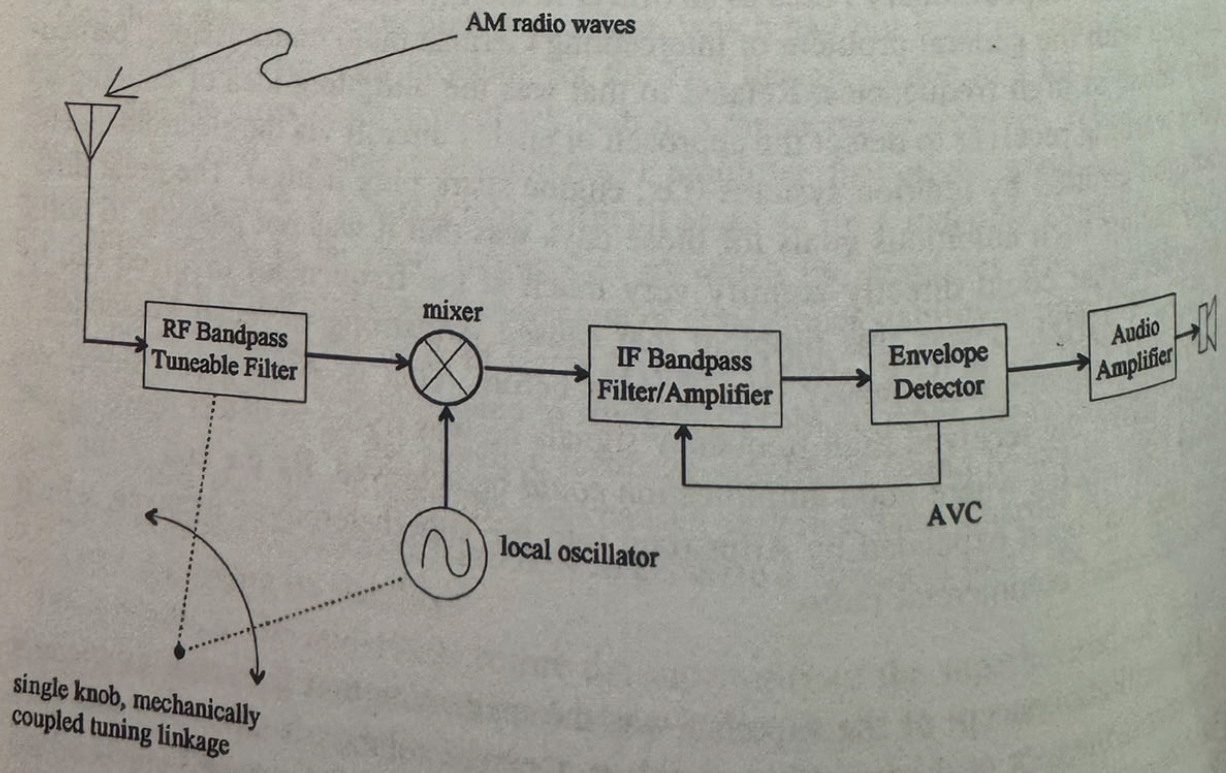


FIGURE 21.1. The AM

because their carrier frequency had been cleared for their exclusive use. Today, with thousands of AM broadcast stations, there are simply not enough frequencies available in the AM band to have a totally cleared channel for all of the so-called clear channel stations. There are, in fact, now over 90 such "clear channel" stations (all transmitting with the legal maximum of 50,000 W).

Those stations that do share a common carrier frequency are, however, at least widely separated by geography. For example, there are two clear channel stations at 1030 KHz: one is WBZ in Boston, and the other is KTOW in the rather distant town of Casper, Wyoming. In the United States, station carrier frequencies end with a "0," but outside the United States it is not uncommon to find stations with carrier frequencies ending with "5." In the United States, this is called a *split channel*. This has been troublesome in the past with stations operating in Mexico, with adjacent channel interference resulting (e.g., a high-power Mexican station operating at 995 KHz will interfere with a nearby United States station operating at either 990 or 1000 KHz—such foreign stations have been aptly called "border blasters"!). How this occurs is discussed later in this chapter.

When you tune your AM superheterodyne receiver, the knob you turn is adjusting two separate and distinct parts of the circuit simultaneously. First, a front-end *tuneable* radio frequency bandpass filter connected directly to the antenna is adjusted. You are adjusting the center frequency of this filter's passband until it is aligned with the carrier frequency of the station that you want to select from all others (this filter is often called the *preselector*). There is no problem with this tuneable filter possibly oscillating (as with the tuneable amplifier stages in the old TRF receivers) because it is either a relatively low-gain device, or even simply passive with *no* gain at all. In the latter case the rf filter *attenuates* the signal centered in its passband *less* than it attenuates signals that are displaced from the center frequency.

A less obvious role for the preselector rf filter is that it isolates the local oscillator (which is the second piece of circuitry adjusted by the tuning knob) from access to the antenna, i.e., the presence of the rf filter "unilateralizes" the receiver. (This is important because the local oscillator can, itself, radiate its rf signal into space and so interfere with other, nearby receivers. More will be said about the local oscillator, soon.) The passband of the rf filter does not, in fact, have to have a very sharp cutoff, i.e., its skirts do not have to even come close to being vertical. The filter's frequency response curve can actually be pretty 'sloppy' in its rate of roll-off, and just how sloppy we'll see in just a bit. (This is good, because it is *very* difficult to build a *tuneable* filter with a sharp roll-off at its passband edges!) Notice that this means if two stations have carrier frequencies that are not very far apart then *both* signals will get through the rf filter pretty much the "same for wear." This might appear to be the beginning of a bad case of adjacent channel interference, but you'll soon see how the superheterodyne neatly sidesteps this.

Now, continuing to move through Figure 21.1, the local oscillator shifts (with the aid of the multiplier, or "mixer") the selected chunk of spectrum that has gotten through the rf filter into the input of a bandpass filter with a center frequency *fixed* at 455 KHz. This is an active filter, in fact, and as it is also an amplifier it is called the *intermediate*

frequency amplifier (or 'IF amp'). The intermediate frequency (the 455-KHz center frequency) gets its name from the fact that f_{IF}=455 KHz is a frequency *below* the high frequency of any AM broadcast band carrier and *above* the low frequencies of the baseband information signal that modulates the carrier. The IF amplifier is a high-gain bandpass filter with a bandwidth of 10 KHz (the bandwidth of the standard AM DSB-LC signal, which of course is twice the bandwidth of the baseband signal) and *very* steep skirts. Since the IF amp is *not* tuneable it is relatively easy to build such a sharp cutoff bandpass filter/amplifier that is stable (that won't oscillate). The sharp cutoff of the IF amplifier eliminates any adjacent channel signal interference (the previously mentioned case of two stations with carrier frequencies close together) that has gotten through the sloppy rf filter and to the IF amp input (along with the desired channel). Only the desired channel can survive the passage *through* the IF amp.

From the spectrum shifting (or heterodyne) theorem we know that, when a signal is "mixed" (multiplied) with the local oscillator signal, the spectrum of the signal will be translated in frequency by the value of the oscillator frequency. Thus, there are *two* station carrier frequencies that will be shifted into the IF amplifier passband, i.e., if f_{LO} is the local oscillator frequency, and if $f_{LO}>f_{IF}$, then the spectrums around the carrier frequencies $f_{LO} \pm f_{IF}$ will *both* be shifted so as to be centered on f_{IF}. If, on the other hand, $f_{LO}< f_{IF}$ then the *two* station spectrums around the carrier frequencies $f_{IF} \pm f_{LO}$ will both be shifted to be centered on f_{IF}. In either case, we call the carriers of the two signals that end up in the IF amplifier passband *image frequencies*.

If there are two stations operating, in fact, on image frequencies, then that can be a catastrophic problem because they'll end up "on top" of each other (so to speak) in the IF amplifier! The sharp cutoff of the IF amplifier's passband does nothing in eliminating this problem of two image frequency stations interfering with each other (remember, it is adjacent channel interference that the sharp cutoff addresses), but now, at last, you can see why the front-end preselector rf filter is present. If the rf filter is tuned to one of the image frequencies (presumably the one you want to listen to!) then, by definition, the other image frequency is *not* tuned. We say that the untuned image is *rejected* and, while it may still put some energy into the IF amplifier passband, we hope it will not be much energy. Indeed, the further apart in frequency that the image frequencies are, the better the rf filter will reject the unwanted image.

So how far apart *are* the images? In the case of $f_{LO}>f_{IF}$ the image separation is $(f_{LO}+f_{IF})-(f_{LO}-f_{IF})=2f_{IF}$=910 KHz, a constant independent of tuning (of the particular station you are listening to). In the case of $f_{LO}<f_{IF}$ the image separation is $(f_{IF}+f_{LO})-(f_{IF}-f_{LO})=2f_{LO}$, a variable dependent on the setting of the local oscillator (which is, obviously, always less than $2f_{IF}$). Thus, for a given value of f_{IF}, the choice $f_{LO}>f_{IF}$ gives the better image frequency separation. And if f_{IF} is sufficiently high then even a sloppy, tuneable rf filter will give excellent image frequency rejection.

In the early days of radio f_{IF} was much smaller than today's value of 455 KHz because of technology limitations in building tuned IF amplifier coupling transformers that could work at high frequency. In 1924, for example, General Electric built radios for RCA using an IF frequency of just 40 KHz. The increase of the value of f_{IF} was motivated by the desire for better image rejection by the preselector rf filter. At carrier