Image Rejection and Direct-Conversion Receivers

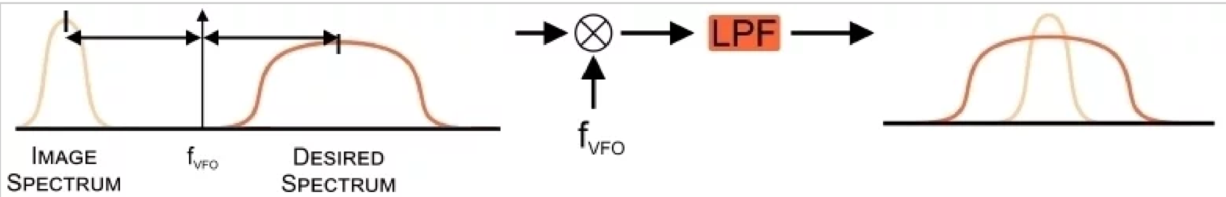
Chapter - Selected Topics

In this page we'll discuss the "image" problem in IF-based receivers, and we'll also look at an alternative approach that eliminates this complication.

In the previous page we explored benefits associated with the use of an intermediate frequency (IF). However, IF architectures entail a serious disadvantage, and in fact this disadvantage is a significant motivating factor in the development of direct-conversion-based alternatives.

The Image

An IF-based receiver uses a variable-frequency oscillator (VFO) signal to shift a received spectrum down to an equivalent spectrum centered around the intermediate frequency; the shifting is accomplished via multiplication. However, this multiplication operation affects not only the received spectrum but also whatever spectrum is located symmetrically with respect to the VFO frequency. In other words, multiplication will shift one spectrum *down* to the IF and another spectrum *up* to the IF.



As you can see, the image spectrum and the desired spectrum are both present in the IF signal that will be demodulated. In this diagram we can easily distinguish one from the other, but such is not the case in a real-life circuit—the frequency information in the desired spectrum is corrupted by the frequency information in the image spectrum.

This symmetrically located image spectrum is a serious impediment to reliable IF-based reception. Why? Because the image spectrum is (presumably) not under the control of the wireless system that you’re designing, and consequently it could be anything, including a signal that is much stronger than the desired signal. Thus, if we don’t do something to mitigate the effects of the image, the quality of the system’s reception will be dependent upon the unpredictable behavior of the signals near the image frequency.

Image Rejection

To mitigate the effects of the image spectrum, heterodyne receivers use image-reject filters. These are placed before the mixer, such that the image spectrum is suppressed before the mixer shifts it to the intermediate frequency. This is an effective solution, but there are two complications.

The Trade-Off

The image-reject filter won’t be very useful if it attenuates the desired spectrum and the image spectrum. Thus, the filter’s response must transition from low attenuation at the desired band to high attenuation at the image band. As with any filter, rapid transitions from passband to stopband are challenging, and thus it will be easier to design an image-reject filter if there is a large frequency separation between the desired band and the image band.

However, the separation between the desired band and the image band is proportional to the intermediate frequency (more specifically, it is twice the intermediate frequency). This means that more separation corresponds to a higher IF. This is not catastrophic, but we have to remember that we want an intermediate frequency to be significantly more convenient, from a signal-processing perspective, than the high frequency used for RF transmission. If we increase the intermediate frequency too much, the difficulties created by the higher IF may outweigh the benefits of improved image rejection. Thus, image-rejection filtering entails a fundamental trade-off between image suppression and the desire to maintain a lower intermediate frequency.

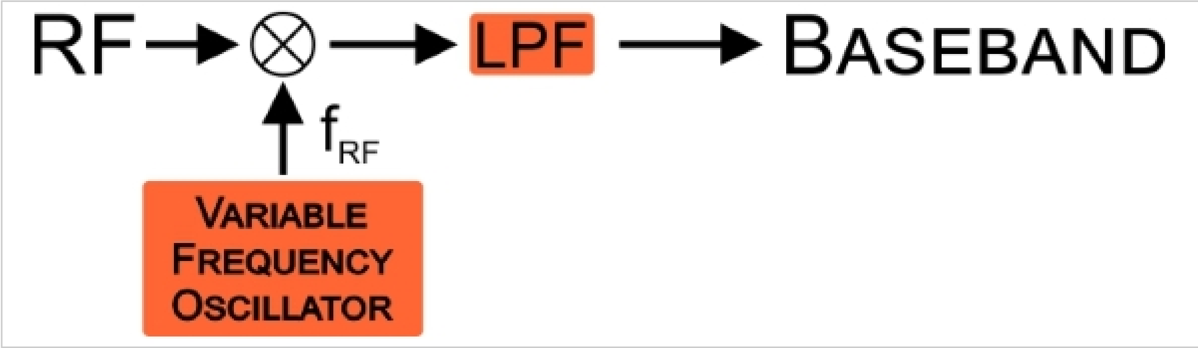
Integration, or Lack Thereof

Image rejection is typically accomplished by means of a filter that is not incorporated into an integrated circuit. In other words, image-reject filters consume PCB area and design time, and in the context of modern electronics, both of these resources are valuable and in short supply.

Companies often attempt to minimize the time involved in bringing a new product to the production phase, and an important way to reduce development time is to avoid custom design whenever possible—in other words, to use tested, characterized, proven integrated circuits instead of newly designed external circuits. Regarding PCB area, it should come as no surprise that miniaturization is a major goal in various electronics industries, and the only way to achieve extreme size reductions is through integrated-circuit technology. Thus, heterodyne receivers that rely on image-reject filters are fundamentally problematic with respect to the inescapable realities of modern electronic design.

A Possible Solution: Direct Conversion

As mentioned in the previous page, a direct-conversion receiver shifts the received signal all the way to baseband instead of to an intermediate frequency. In other words, the frequency of the VFO is always equal to the center frequency of the desired spectrum:



This approach includes one very important benefit—it eliminates the image problem. In the direct-conversion scheme, there is no image spectrum: the desired spectrum is centered around the VFO frequency, and no spectrum can be symmetrical with respect to the VFO frequency when the desired center frequency and the VFO frequency are equal.

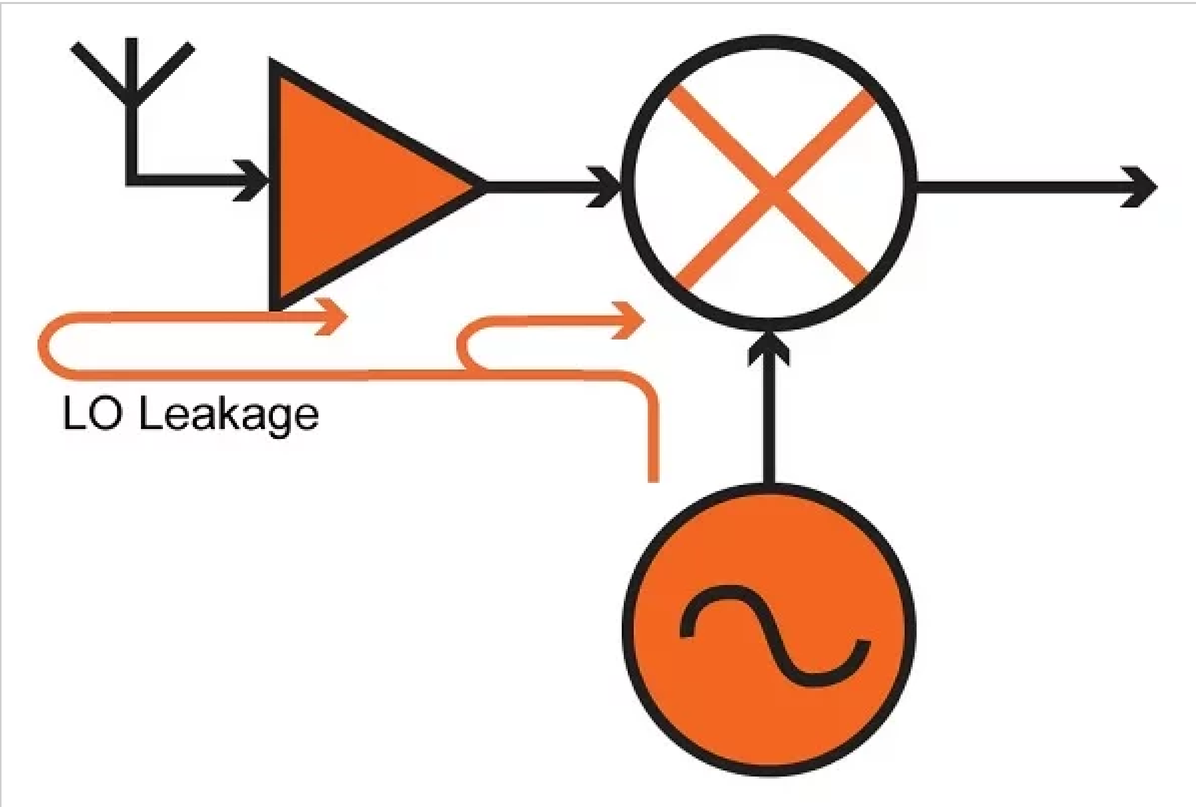
Another benefit of direct conversion is simply an extension of the benefits associated with IF-based architectures. An intermediate frequency facilitates signal processing because it is significantly lower than the transmission frequency, but signal processing can be even easier when the "intermediate" frequency is 0 Hz—i.e., when the received spectrum is shifted directly to baseband.

Direct-conversion immediately seems like a superior alternative: it's conceptually simpler, there's no image to corrupt the received spectrum, low-frequency signal processing replaces intermediate-frequency signal processing, and the absence of an image-reject filter allows for expanded use of IC technology. Why, then, does anyone even consider an IF-based architecture? Well, it turns out that there are several significant disadvantages associated with direct conversion. Here we will discuss only the disadvantage that is perhaps the most serious.

DC Offset

RF receivers are sensitive to DC signal components because the amplitude of received signals is often extremely small. These small-amplitude signals create the need for high-gain amplification, but high-gain amplification can rapidly break down when a signal has a nontrivial DC offset, because multiplication of the offset saturates the amplifier.

Mixers readily create DC offsets, because multiplying a sinusoid by another sinusoid with the same frequency and phase creates a non-varying signal component. Back in Chapter 3 (https://www.allaboutcircuits.com/textbook/radio-frequency-analysis-design/real-life-rf-signals/coupling-and-leakage-in-rf-systems/), we discussed the complications caused by the fact that RF signals do not stay in their intended conduction paths. Rather, their high frequency allows them to “leak” into portions of the circuit where we don’t want them. The problem of DC offset creation is a perfect example of this difficulty: the local oscillator signal leaks into other portions of the circuit in such a way that it is present in both of the mixer’s inputs, and the result is a DC offset in the output signal.



A direct-conversion receiver must implement some sort of DC offset cancellation, and this is not a particularly easy task; filtering is generally not feasible because the filter would suppress portions of the desired spectrum, which has been shifted down to the band around DC. A heterodyne receiver, on the other hand, can easily remove DC offsets via filtering because there is plenty of frequency separation between DC and the IF band.

Summary

- IF downconversion causes an image spectrum to corrupt the desired spectrum.
- The image spectrum can be suppressed via filtering, but the filtering approach to image rejection involves an important design trade-off and prevents monolithic integration.
- Direct conversion eliminates the image problem, but it has various disadvantages.
- A particularly challenging characteristic of the direct-conversion architecture is its susceptibility to DC offsets generated by the mixer.