What Is FM Deviation, and Why It Matters for TARPN

Chuck K4RGN

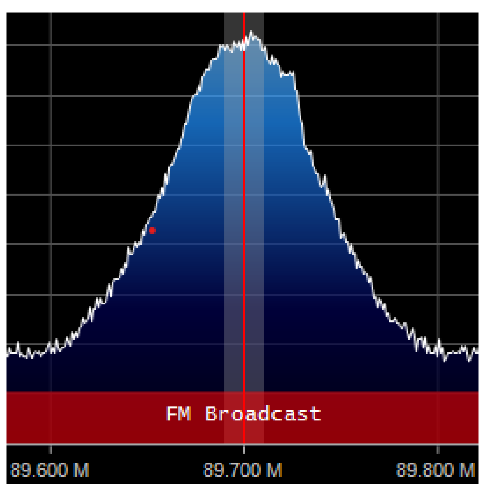
The author thanks Nino KK4HEJ, Don N2IRZ, Tadd KA2DEW, and Doug N3LTV for their contributions.

Caveat: this is written specifically in the context of TARPN, which uses analog, voice-channel, FM transmission.

Optimal performance of packet links in amateur radio depends on properly setting the deviation of the transmitter. This paper explains the underlying theory and provides practical advice on setting deviation.

First we review the basics of analog, voice-channel FM radios.

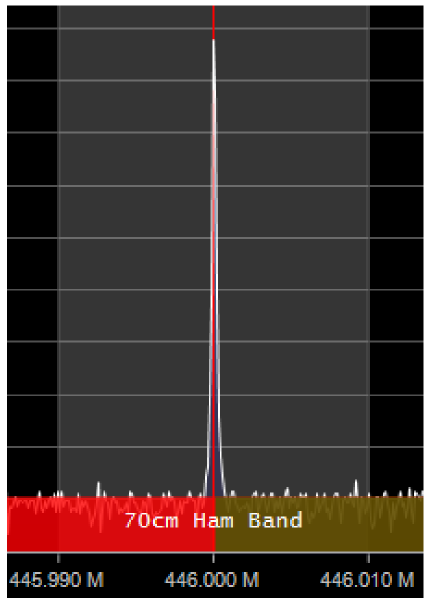
1. Let's define "spectral display" as a graph-like display of signal frequency along the X-axis and amplitude along the Y-axis. What you see on a spectral display, left to right is "transmitted bandwidth".
2. Transmitted bandwidth at any one instant is a function of several parameters. The math to calculate instantaneous transmitted bandwidth is very complex and uses a concept called Bessel functions. It turns out that theoretically, transmitted bandwidth of a modulated FM signal is infinite. However, almost all of the transmitted energy falls within a range of frequencies. We refer to that range as the transmitted bandwidth. Here is a spectral display of a commercial FM broadcast station. At the time, it was transmitting music. (Note: this is not an HD FM station.) You see that the transmitted bandwidth is roughly 200 KHz.



3. Carson's Rule provides a simplified way to calculate maximum transmitted bandwidth as a function of only two parameters: the "deviation" of the modulator and the highest audio frequency that is provided to the modulator.

Carson's Rule is not 100% accurate, but it is good enough for our purposes. We will discuss deviation a bit later.

4. An FM transmitter has constant RF power output regardless of its audio input. Depending on the audio input to the modulator, the energy from an FM transmitter is distributed across the transmitted bandwidth in various ways.
    a. An un-modulated transmission, of course, has all the power at the carrier frequency.



    b. By injecting a single sine wave into the modulator, it is possible to illustrate how the energy of the FM transmitter is distributed across the transmitted bandwidth. See Exhibit A for examples.
    c. Notice that the distribution of power across the transmitted bandwidth changes with the amplitude of the audio sine wave.
    d. Notice that the distribution of power across the transmitted bandwidth also changes with the frequency of the sine wave.
    e. For all but the most simple audio inputs, it is virtually impossible to eyeball the spectrum of an FM transmission and deduce the audio input.
5. The FCC limits maximum transmitted bandwidth. For analog voice FM radios in commercial service under Part 90 of the FCC Rules, the limit is either 11.2 KHz ("narrowband") or 20 KHz ("wideband"). Analog voice FM radios in amateur service under Part 97 of the FCC Rules almost always use wideband mode, so their limit is 20 KHz.
6. The FCC does not require that a transmitter occupy all of its maximum allowed transmitted bandwidth, but a radio is usually designed to approach the maximum allowed transmitted bandwidth in normal operation.

7. As mentioned, deviation is one of the input variables in Carson’s Rule to calculate transmitted bandwidth. Deviation is determined by the design of the modulator. Maximum deviation is attained when the audio power provided to the modulator reaches the maximum figure intended by the designers of the radio. Lesser audio power produces less deviation than the maximum.
8. Typically, the modulator will clip audio power that is above the intended maximum amplitude. The modulator also cuts audio higher than a certain frequency. These measures ensure compliance with FCC limits on maximum transmitted bandwidth. Clipping is to be absolutely avoided, of course, because it introduces distortion into the transmitted signal.
9. Transmitters are designed for use with one or more specific microphones, usually from the same manufacturer as the radio. Output voltage can vary considerably from one microphone to another. Transmitters may have a manually adjustable microphone gain factor or a circuit that tries to adjust microphone gain automatically. In either case, the designers of the radio have a target in mind for maximum audio level that corresponds to maximum deviation. This target varies considerably from one radio to another.
10. 3-4 KHz for the maximum audio frequency is typical in Part 90 and Part 97 radios.
11. 4-5 KHz for maximum deviation is typical for wideband mode.
12. In terms of receiver performance, all other things being equal, the greater the deviation of the transmitted signal, the higher the signal-to-noise ratio of the demodulated signal. For Part 90 radios, reduced signal-to-noise ratios were the inevitable consequence of changing from wideband to narrowband. The FCC believed that increasing the number of available channels in a fixed amount of spectrum was more important than preserving high signal-to-noise ratios. This phenomenon also applies to Part 97 radios that have the option to transmit in narrowband mode.
13. Receiving a narrowband signal on a wideband receiver produces low audio output. Receiving a wideband signal on a narrowband receiver produces distortion. Be sure that both radios of a link are using the same wideband vs narrowband settings.
14. Neither microphones nor speakers of Part 90 and Part 97 radios are high-fidelity. Therefore, in a price-competitive market for radios, designers tend to be satisfied with a level of fidelity from the radio electronics that is reasonable for speech in noisy backgrounds. In general, these are not high-fidelity radios even if the modem attached to them has near-perfect audio characteristics. This inherent imperfect fidelity can have significant implications in the packet radio context.

With this background, we can look at TARPN specifically.

15. A TNC contains a modem that is strictly an audio device. The modem has no direct knowledge of what's happening on the radio link.

16. The role of the radio link is to reproduce, as faithfully as possible, the audio output of the sending TNC at the audio input of the receiving TNC – so long as the audio from the sending TNC is within the limits of power and highest frequency that the radios were designed to accommodate. In other words, at its best the radio link is a "virtual wire".
17. In very general terms, the signal-to-noise ratio of a received FM transmission is highly correlated with the bit error rate in the receiving TNC. For this reason, packet radio usually performs better across wideband links than narrowband links. Of course, if the distance between radios is 10 feet, you won't notice the difference.

Let's illustrate with the example of a 1200 bit/second TNC using the Bell 202 modem standard, which is an AFSK device (audio frequency-shift keying). In AFSK, a modem generates tones of specific audio frequencies, each of which corresponds to a different data value. The Bell 202, like most AFSK devices, is binary. It uses two tones and switches between them rapidly as determined by the data stream being transmitted.

18. The AFSK waveform generated is a constant-power waveform. That is, the amplitudes of the audio frequencies corresponding to bit values "0" and "1" (1200 Hz and 2200 Hz, respectively) are identical.
19. When viewing the audio spectrum of the AFSK waveform for a data stream, you will see spreading around the 1200 and 2200 Hz frequencies. This is the inevitable consequence of near-infinite frequency components as the data stream into the TNC shifts between a "0" and a "1", causing the modem to switch between the two frequencies. You can think of this phenomenon as similar to key clicks in CW. This is why a Bell 202 modem cannot be used at speeds such as 9600 bits/second. The spectra of the "0" and the "1" would continue to widen until they overlap, rendering the modem useless.
20. Given that the 1200 and 2200 Hz frequencies are standardized and that the waveform has constant power, the only variable of the waveform is its amplitude.
21. Other than switching between the two audio frequencies, the only way a TNC using the Bell 202 modem can affect the RF output of a radio is by changing the amplitude of the AFSK waveform that is passed to the radio's modulator.
22. A relatively low AFSK amplitude from the TNC into the radio's modulator will result in relatively low deviation from the radio. A relatively high amplitude will result in relatively high deviation, up to the maximum deviation that the radio was designed for.
23. We have seen that all other things being equal, high deviation is desirable. A simple but naïve approach to optimizing the performance of a 1200 bit/second AFSK link is to increase the amplitude of the AFSK waveform until the maximum deviation of the radio's modulator is reached.

24. But in reality, this procedure is deceptively dangerous.
    a. Unless you are making measurements with appropriate instruments, you don't know what AFSK amplitude corresponds to maximum deviation. Experts can determine deviation by eyeballing a spectrum display or listening to the RF signal with a receiver, but many amateur operators do not have that expertise.
    b. Nor should you rely entirely on information from the manufacturer of the radio. Specifications are often approximations. Even when specifications are accurate, performance of the radio may have been affected by extreme values of component tolerances or by drifts in component values as they aged.
    c. There is no guarantee that the receiving radio and the transmitting radio are symmetrical in terms of their RF, IF, and audio bandwidths. The transmitting radio may support a larger deviation than the receiving radio can handle.
25. Therefore it is customary to initialize 1200 bit/second AFSK at a conservative deviation of 3 KHz. In may be possible to reduce the bit error rate of a link by increasing AFSK amplitude so that deviation exceeds 3 KHz, but beware of overdriving the FM modulator. If the bit error rate of the link begins to rise rapidly as AFSK amplitude is increased, there is evidence of overdrive.
26. The discussion above applies to terrestrial radio links. When using the Bell 202 modem to communicate with the International Space Station, for example, Doppler effect arising from the satellite's orbital velocity (roughly 15,500 miles per hour) becomes significant at VHF and UHF frequencies. For ISS, peak deviation must be strictly limited to 3 KHz. Otherwise the RF signal may be Doppler-shifted partially outside the bandpass of the receiver. This 3 KHz figure is sometimes quoted as a maximum for deviation in a packet environment. That's not entirely true for terrestrial links, although it happens to be a good starting point.
27. But remember: insufficient deviation, such as "narrowband" as required by the FCC for Part 90 radios since 1997, will result in a lower signal-to-noise ratio over the air and therefore a higher bit error rate after demodulation by the receiving TNC.
28. To summarize, audio amplitude from the TNC must not go too low or too high. We suggest that you begin with the 3 KHz recommendation for deviation. If the bit error rate is significant, you may cautiously see if performance is improved by increasing audio amplitude from the TNC. If the bit error rate is acceptable at 3 KHz deviation, leave it there.

So what about data rates faster than 1200 bits/second that do not use the Bell 202 modem?

29. Because modems in TNCs are fundamentally audio devices, they must always work within the audio bandpass that Part 90 and Part 97 radios implement.

30. Although the worst-case of the audio bandpass is generally considered to be 3 KHz, we have found that some radios will support considerably higher audio frequencies if we bypass the radios' audio filters and emphasis/deemphasis circuitry to use "raw audio". Audio bandpass of 4 KHz or more may be possible with acceptable parameters of amplitude, frequency, and phase response.

31. Regardless of the audio bandpass, however, the radios still have a maximum deviation. As for AFSK, adjusting the amplitude of the audio waveform from the TNC still has the effect of adjusting the deviation of the RF signal relative to the maximum deviation that the radios can provide.

32. According to Carson's Rule, an audio signal whose highest frequency is 6 KHz and a 3 KHz deviation will produce a transmitted bandwidth of 18 KHz. This is as wide a signal as any Part 90 or Part 97 radio can be expected to support. A combination of 4 KHz deviation and a 6 KHz maximum audio frequency is almost certain to exceed IF or RF bandpass limitations in one or both radios. In other words, there can be a tradeoff between audio bandwidth and deviation. A combination of 4 KHz audio bandwidth and 4 KHz deviation might work very well on some radios but not on others.

33. In short, managing the instantaneous amplitude of the audio from the TNC to the modulator of the transmitting radio remains very important. In fact, it may be even more important for higher data rates than for 1200 bits/second.

34. As with AFSK, insufficient deviation means poor signal-to-noise ratio, which means high bit error rates.

How do we apply this theory to the NinoTNC and other TNCs intended for use at 2400 bits/second or faster?

35. The NinoTNC is capable of data rates of 1200, 2400, 4800, and 9600 bits/second. At 1200 bits/second the NinoTNC implements the Bell 202 standard.

36. Whether 9600 bits/second is usable in practice, however, depends on specific radios. Many recent radios support 9600. Of the ones that do, most provide easy access to raw audio. Some older radios may also support 9600 if their circuitry is modified to use raw audio.

37. As reference: the highest audio frequencies produced by a sending NinoTNC are 2200 Hz for 1200 bits/second; 2400 Hz for 2400 or 4800 bits/second; and 4800 Hz for 9600 bits/second. Note that in the case of 2400 bits/second only, the NinoTNC's audio output has a 7200 Hz component which the transmitting radio will cut off.

38. For speeds of 1200 and 2400, we recommend an initial setting of the TX DEV potentiometer in the NinoTNC to produce 3 KHz deviation. For speeds of 4800 and 9600, we recommend an initial setting to produce 2.4 KHz deviation. It may be possible to improve performance by cautiously increasing TX DEV, which will increase deviation. But beware of the limitations discussed above.

How do I configure a TNC to set a value for amplitude that results in the desired deviation?

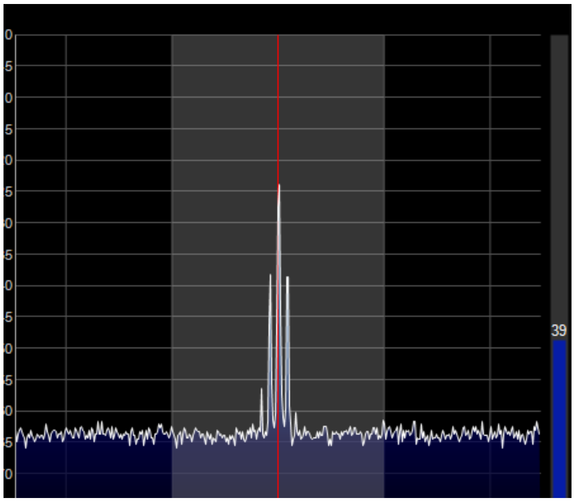
39. Use the “Bessel zero” technique described by Jim Andrews, KH6HTV at https://bit.ly/3itQh2k. The technique requires an instrument that produces a real-time spectral display. Although an expensive spectrum analyzer can be used for this purpose, a simple SDR USB dongle that sells for less than $30 is satisfactory. Such a dongle is used in conjunction with freeware such as SDR# (SDRSharp) for Windows and Linux or CubicSDR for the Macintosh.
40. The TNC must be capable of emitting an audio sine wave of known frequency upon request. The NinoTNC does this when its test button is depressed. See NinoTNC documentation for the specific frequencies available.
41. Adjust the audio amplitude from the TNC until the transmitted signal has nearly zero RF power at the carrier frequency. When looking at the transmitted spectrum, almost all the RF power will be in sidebands on either side of the carrier frequency when the carrier null has been achieved. Note that it might not be possible to reduce the carrier power to absolute zero, but you will see that it reaches a minimum.
42. As you increase audio amplitude from zero, the first point at which power at the carrier frequency is nulled corresponds to deviation that is 2.4 times the frequency of the sine wave.
    a. If the sine wave is 500 Hz, the deviation at the first carrier null will be 1200 Hz.
    b. If the sine wave is 1000 Hz, the deviation at the first carrier null will be 2400 Hz.
    c. If the sine wave is 1250 Hz, the deviation at the first carrier null will be 3 KHz.
    d. If the sine wave is 2080 Hz, the deviation at the first carrier null will be 5 KHz.
43. Beware that there in some situations there may be multiple Bessel nulls as the audio amplitude is increased further. The factor of 2.4 applies only at the lowest Bessel null. To avoid measurement error, always start at minimum amplitude and slowly increase amplitude until the first Bessel null is reached. If you inadvertently reach a second or third Bessel null, your calculation of deviation will be way off.
44. Settings of deviation are not critical, so long as you are not at the maximum deviation of the radio. Accuracy of 5% is acceptable. Some TNC potentiometers may be difficult to adjust more precisely than 5% anyway.

Will we always face this complexity of amplitudes, peak deviations, etc?

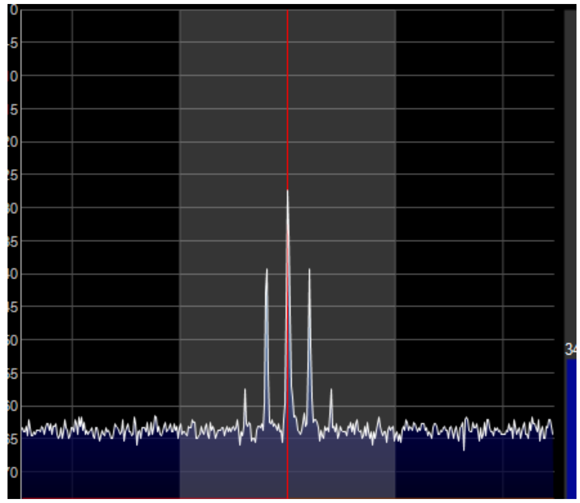
45. No. Someday we hope to have a true digital radio that TARPN can use. In such a radio, the constraints of an analog, voice-grade Part 90 radio or its Part 97 equivalent are removed. DSP can be applied directly to RF instead of audio. Data rates faster than 9600 bits/second can be achieved. At 220 MHz and higher, FCC restrictions on data rate and transmitted bandwidth are considerably relaxed.

Exhibit A: Examples of FM spectra, modulated by a single sine wave

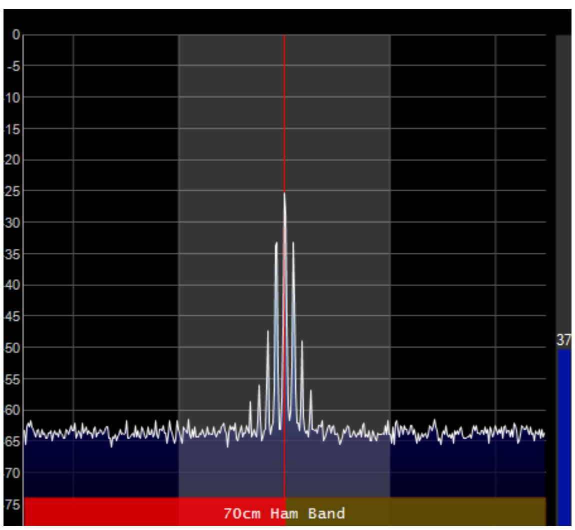
Low audio frequency, low amplitude



High audio frequency, low amplitude



Low audio frequency, high amplitude



High audio frequency, high amplitude