A Gentle Introduction to Ultra-wide Band (UWB) Radio Technology

by

Kazimierz “Kai” Siwiak, Ph.D., P.E.
TimeDerivative Inc.

Introduction
Ultra-wide band (UWB) radio is a technology is a method of spectrum access that can provide very high speed data rate communications over the personal area network space. It can readily trade data rate for range. In some implementations, UWB technology provides accurate ranging and radar information. In this course we will discover the origins of UWB technology by investigating basic narrow band and wide band wireless concepts and how technology over time and regulations have shape the electromagnetic spectrum. The earliest wireless spark gap transmitters were UWB by nature, however, the corresponding receivers made inefficient use of the wideband signals. That and interference considerations drove wireless towards narrow band radio technology. With the advent of modern digital processing capabilities UWB technology again rises as an efficient means of communications. Regulations now allow the technology under the FCC Part 15 umbrella, and the ITU-R has issued a Recommendation on the use of UWB technology.

We will trace the historical development of wireless on its march to conventional narrow radio, and then to increasing signal bandwidths to gain performance improvements as wireless technology matures, culminating in today’s UWB technology. We will see how UWB signals can coexist with other systems. A simple UWB transmitter will be shown, and system performance will be explained. Finally we will investigate the potential application space for UWB technology.

Separating Signals by Frequency
All signals, regardless of origin, simultaneously share the same ‘transmission medium’ - the near-vacuum of space; the atmosphere enveloping the earth and the many materials surrounding us. Yet we can selectively choose the signal we want, such as the radio station to which we wish to listen, the television program we want to watch, or the call intended for our mobile telephone. Conventional radio signals can be discerned one from the other because they occupy unique locations in the radio spectrum for instance, unique audio tones or discrete colors in the rainbow spectrum. They are distinct, narrow-band places on a radio dial, indicated by channel numbers. They are all crafted that way due to a century-old historical interplay between the technological development of radio, and the regulations that brought order to the radio spectrum.

---

1 RECOMMENDATION ITU-R SM.1757, Impact of devices using ultra-wideband technology on systems operating within radiocommunication services (Question ITU R 227/1) 2006.
Figure 1 – The various radio services and radio signals appear at unique locations in the electromagnetic spectrum.

In Figure 1 we can identify electromagnetic emission at frequencies from just a few Hertz (25 to and 60 Hz power lines) up through the radio frequencies, then light, X-rays and Gamma Rays. The realm of radio waves is generally accepted to span the spectrum from 9 kHz to 275 GHz which the range of frequencies that have been currently allocated by the FCC or NTIA. Radio signals share this limited spectrum by occupying slivers of spectrum that are as narrow as permitted by the bandwidth of the information content of the signal.

A signal without information has zero information bandwidth. Modulating information on that signal spreads its bandwidth in proportion to the information bandwidth. For example, a music signal with tonal content up to 15 kHz requires at least 15 kHz of information bandwidth. The “ideal” in radio spectrum usage has been to use the smallest bandwidth compared with the bandwidth of the signal information. Narrowband signals are often represented by their zero bandwidth ideal, the sine and cosine functions, also known as circular functions or harmonic waves. They are the narrowest possible representation of signals in the spectrum at distinct
frequencies. Tuning radios to a particular frequency allows us to select the desired narrow band signal. So, the dogma of the circular functions\(^2\), sines and cosines, began to dominate radio development.

Figure 2 portrays the separation of signals by operating frequency, analogous to pure tones on a musical scale; the notes that might be played on a flute or piano. Each tone comprises an indefinitely large number of sinusoidal cycles, and appears as an impulse in the frequency spectrum.

We see a “spectrum management” method emerge. Signals can be separated one from the other on the basis of their location in the spectrum. They are differentiated by their unique location in spectrum, and

- “Ideal” – smallest bandwidth occupancy as needed by the modulated information
- “Best capacity” is when each user occupies the smallest sliver of bandwidth.

Thus the march was on to populate the spectrum one sliver at a time, while “managing” interference and “unintentional” signals the best way that technology will permit.

---

Figure 3 – Narrow band signal – separation by frequency.

Narrow band signals used for communications as portrayed in Figure 3 are not pure tones in the frequency spectrum. They are modulated, which broadens the signal. They also inevitably contain noise and “un-intentional radio emission” that are well below the level of the intended signal plus modulation. These unintentional emissions are subject to regulations because they can cause interference to other signals outside their own bandwidth, which then limits the spectral capacity in a frequency band. Traditional spectrum management is the art of grouping frequencies into bands of signals with similar characteristics as close together as technology allows for the purpose of improving overall spectral capacity.

Our representative narrow band communications signal can have the very general form

\[
S(t) = I(t) \cos(2\pi f_0 t) + jQ(t) \sin(2\pi f_0 t)
\]

where \(I(t)\) and \(Q(t)\) represent the information signals that modulate the sinusoidal signal that appears at frequency \(f_0\) in the frequency spectrum. It can be shown, but is beyond scope here, that Equation (1) can represent conventional modulated narrow band signals like AM, FM, PM, and
so on. The key factor is that the signal bandwidth is controlled by the modulation bandwidth; that is, the bandwidth of \( I(t) \) and \( Q(t) \). For our pure tone example, \( Q(t)=0 \) and \( I(t)=1 \). That signal occupies an infinitesimally small bandwidth at frequency \( f_0 \), and occupies that spectrum for an indefinite time \( t \). The time properties and the frequency spectral properties are related by a Fourier transform. The signal can be represented equally well in time and in frequency.

**Separating Signals by Time**

Separation of signals by bands, by channels, and by frequencies is not the only way to share the radio spectrum. Information bearing signals can also be separated in time, especially in tiny slivers of time. These signals occupy wide bandwidths, ultra-wide bandwidths – but short, and ultra-short slivers of time.

![Figure 4](image)

**Figure 4**  
A finite length signal in time, occupies a definite spectrum width in frequency, for that finite time.

The tinier the sliver of time, the wider the bandwidth of the signal in the radio spectrum as seen in Figure 4. When confined to just four cycles of a sine wave, the signal occupies significant bandwidth, on the order of 50%.
The “main lobe” or unfiltered bandwidth of a truncated sine wave is inversely proportional to $N$, the number of sine wave cycles,

$$BW = 2 \cdot \frac{f_{\text{upper}} - f_{\text{lower}}}{f_{\text{upper}} + f_{\text{lower}}} \cdot 100\%$$

(2)

Using the notation of Figure 4,

$$BW = 2 \cdot \frac{\left(\frac{1}{T} + \frac{1}{NT}\right) - \left(\frac{1}{T} - \frac{1}{NT}\right)}{\left(\frac{1}{T} + \frac{1}{NT}\right) + \left(\frac{1}{T} - \frac{1}{NT}\right)} \cdot 100\% = \frac{200\%}{N}$$

(3)

Thus 4 sine cycles have a fractional bandwidth of 50%. A single cycle has a 200% bandwidth.

Figure 5 – The shorter a signal is in time, the wider its spectrum is in frequency, for that duration in time.

One cycle of a sine wave occupies a 200% bandwidth between the first two zero crossings in the spectrum. However, an infinitesimally short impulse in time (like a drum beat) occupies an indefinitely wide spectrum, but just for that infinitesimally short time. Figure 6 shows a sequence in time of impulse, like drum beats.
Figure 6 – Impulse in time, like a drum hits, occupies an indefinitely wide spectrum in frequency, for *an instant in time*.

It is intuitively easy to grasp that “drum beats” and “flute tones” can occupy the same spectrum, and that each tone and each drum beat can be readily discerned.

We see now that there are two ways of sharing the electromagnetic spectrum among many users. The spectrum can be divided in frequency and each user can be assigned a small sliver of the spectrum – a channel. Alternatively the many users can each occupy a very large spectrum, but for a short sliver of time each. There is, of course, a wide range of intermediate possibilities of separating signals by combinations of time and frequency. *Most UWB approaches will separate signals in a combination of both Time and in Spectrum.*

We see now that signals can be separated in *time*.
- “Ideal” now is the “shortest time duration”
- “Best capacity” is when each user occupies the *smallest sliver of time* - impulse radio!

*So narrow bandwidth is not the only way to go! – narrow time slices can work also!*
A Time-line of Radio Development
Radio, called wireless at its inception, started out as “ultra-wideband.” This was a consequence of the first transmitters being electro-mechanical contraptions that generated signals using sparks flying between gaps. The early history of wireless is the story of invention, engineering and legislation in a relentless march towards narrowband channelized usage of the radio spectrum.

In 1864, James Clerk Maxwell, formulated the concept of electricity and magnetism using the language of mathematics in his equations of electromagnetism. His theory predicted that energy can be transported through materials and through space at a finite velocity by the action of electric and magnetic waves moving through time and space. Maxwell never validated his theory by experiment, and his results were controversial at the time.

The task of validating Maxwell’s mathematics fell to Heinrich Rudolf Hertz, who in the 1870s embarked on a remarkable set of historical experiments spanning the years 1886 to 1891. In 1886, using a spark gap apparatus to generate radio energy, Hertz created and detected such oscillations over a distance of several meters in his lab. The radiated waves were dubbed Hertzian Waves at the time. The era of “wireless” had begun. By the turn of the century, the radio arts were developed to practical usability by pioneer inventors and scientists.

---

About the Technology of the Time

Early wireless communications relied on Morse Code signaling which was generated by hand and copied by ear. Morse signaling consists of keying a carrier signal on and off in combinations of dots (“dits”) and dashes (“dahs”) that represent alphabetic characters. A moderate messaging rate was about 25 words per minute – which in today’s measure, is equivalent to 10 bits per second. So the information bandwidth of the early wireless signals was relatively small, 10’s of Hz, yet the crude transmitting apparatus emitted very wide band signals, often 100’s of kHz wide at center frequencies in the few hundreds of kHz. The consequences were that:

1. Signals occupied significantly more spectrum than necessary for communications, hence causing significant interference among stations.
2. Receivers were likewise wide band and relatively “deaf” (inefficient). Thus they collected excess background noise compared to the information bandwidth and could “hear” only the strongest signals. Consequently the signal to noise ratio (SNR) was poor.

With the combination of spectral inefficiency and receiver inefficiency, interference among wireless communicators was a serious issue. Wireless needed to become narrow band for survival. Here’s why.

We can understand the reception problem of wide band signals with small information bandwidths by reference to Figure 7. A signal $s(t)$ is delivered by the antenna to a band-pass filter $h(t)$. That is, $s(t)$ and $h(t)$ are convolved. This signal is multiplied in the mixer by a template signal $p(t)$. We also note that noise at the antenna is very wide band and it persists in time. Our objective in optimizing the receiver is to maximize the energy per information bit to noise density ratio at the receiver output.

For the classical narrow band signal case $s(t)$ persists in time, and $h(t)$ is a narrow band filter. Template signal $p(t)$ is simply an oscillator at the reference carrier frequency sine wave, and its product with the filtered $s(t)$ signal lets us recover the modulation in the integrator. The output energy per bit to noise density ratio is maximized when the filter response $h(t)$ is matched to the modulation bandwidth of $s(t)$. We have a classic ‘matched filter’ receiver where the filter response dominates the signal to noise performance improvement.
For UWB signals $s(t)$ is a sequence of impulses, unlike the noise it has a short duration in time. The signal is again convolved with the filter $h(t)$, but this time there is a mismatch since $s(t)$ does not persist in time. For the impulse signal we need some way to ‘gate’ the signal plus the added noise in time. That is done by crafting a template signal $p(t)$ which is matched to the impulsive UWB signal $s(t)$. Now we have the architecture of a ‘matched template’ receiver where the output energy per bit to noise density ratio is maximized when $p(t)$ equals the shape of $s(t)$. We now have a ‘matched template’ receiver where the template response dominates the signal to noise performance improvement.

Generating an accurate template and synchronizing it with the arrival of an impulsive signal was beyond the technology of the time. However, cleaning up transmitter signals with filters and using receiving filters, that is, narrow banding was comparatively easy to do.

Radio: The Era of Broadcasting and Regulations
By 1905, Reginald Fessenden invented a continuous-wave voice transmitter using a high-frequency mechanical alternator. In another 10 years voice broadcasting became commonplace. Prior to 1912, radio was largely the domain of experimenters and ship to shore communications for both Naval and commercial operations. Interference was a serious problem because the old spark transmitters emitted very wide band signals. The Communications Act of 1934 established the Federal Communications Commission (FCC) given regulatory powers in both wire-line and radio based communications. Stations were to be licensed and separated by wavelength, or
frequency, and stations were to use a “pure wave” and a “sharp wave” (sine wave carriers) in the words of the FRC. Sine wave communications and narrow band signals were now mandated. Unfiltered spark emissions, dubbed “class B damped sine wave emissions,” were prohibited. Radio signals were destined to become channelized. These rules required that radio signals be narrowband. By organizing the spectrum and controlling interference, the regulations smoothed the way for commercial AM broadcasting to grow.

Advantages in Wider Bandwidths
The information in AM is encoded by amplitude variations in a carrier. Any other natural amplitude variations, like amplitude noise, static and lightning crashes, would add to the desired amplitude modulated information and be perceived as noise and distortion. Edwin Armstrong found a way to make broadcast radio insensitive to these amplitude distortions. In 1933 he discovered the advantages of wide band frequency modulation (FM). In FM modulation the frequency rather than the amplitude of the transmitter carrier was varied in proportion to the amplitude of the voice signal. Most importantly, Armstrong realized that an FM signal did not need to have a narrow bandwidth. It could vary over a wide range, several times as wide as an AM signal, and as a result have a far better signal to noise ratio than AM. This meant that programs broadcasted using wide band FM could be made higher fidelity and less distorted than AM broadcasts.

Armstrong’s discovery laid the foundations for information theory, which quantifies how noise immunity can be achieved at the cost of signal bandwidth. Voice and music transmissions could now be static free. At the cost of intentional and controlled bandwidth spreading of a signal beyond its information bandwidth, a significant desirable benefits of noise immunity compared to AM was achieved – this was a very important challenge to the narrowband mantra.

Commercial broadcast interests developed along channelized services in the AM broadcast band, and later in wider channel bandwidths in the FM broadcast band. Throughout 1940 to 1960 radio frequencies for land-mobile communications were allocated in the 150 MHz and 450 MHz ranges. As the pressure increased for more and more radio signaling and radio services, higher and higher frequencies in the radio spectrum were being assigned, channelized and developed.

---

Traditionally, the FCC had favored narrow-band radios which concentrate all of their power in fairly narrow channels within the radio frequency spectrum. However, as the number of users sharing the spectrum was increased, the number of available channels became limited. Claude Shannon in 1948 offered a new paradigm, redefining the relationship among power density, noise and information capacity\(^5\). Shannon said that under certain specific conditions, the more an information signal is spread in bandwidth in a way that makes the signal resemble background noise, the more information it is capable of holding. Because one signal spread in this way resembles noise to another signal similarly spread, both can coexist because under some specific conditions, signal energy can be detected more efficiently than noise energy. Thus the wider the bandwidth, the more such sharing can occur and the more total information can be conveyed. Hence, an alternative to transmitting a signal with a high power density and low bandwidth would be to use a low power density and a wide bandwidth.

Shannon’s observations led to “spread spectrum” modulation where the signals are intentionally spread using a special family of digital codes to many times their information bandwidth. The special digital codes are used to distinguish multiple users that are simultaneously sharing the same band. Spread spectrum technology applied to cellular telephone system resulted in a change in spectrum regulation policy. It was the second time in the history of radio that the advantages of wide band signaling was recognized as important enough to result in changes in the

---

regulations away from the narrow band mantra. This time, the FCC had allocated a block of spectrum within which multiple users shared the entire block of spectrum by overlapping signals across the entire band, rather than by allocating narrow slivers of bandwidth per user. Spread spectrum users would be separated by coding rather than by frequency channels. Because of the increased efficiency in the use of precious (and expensive) spectrum, this led to improvements in the capacity of cellular systems that in turn reduced the cost of spread spectrum cellular services. As seen in the progression in Figure 13, one “chip” in a modern spread spectrum system, while wide band in comparison to a narrow band signal, is still narrow compared with UWB technology.

Figure 13 – UWB: Large Fractional Bandwidth.

Still Wider has More Advantages
Through the years, a small cadre of scientists has worked to develop various techniques of sending and receiving short impulse signals between antennas. Impulses are short time signals - the shorter the impulse, the wider its bandwidth, as depicted in Figure 13. The experiments led to “impulse radio,” later dubbed UWB radio. By the late 1960s and 70s the virtues of wide band non-sinusoidal communications were being investigated for non-government uses. Prior to that, the primary focus was on impulse radar techniques and government sponsored projects. In the late 1970s and 80s the practicality of modern low power impulse radio techniques for communications and positioning/location was demonstrated using a time coded time-modulated
approach\textsuperscript{6,7} with UWB spread spectrum pulse techniques. Digital pulse radio, the modern echo of the Hertz and Marconi century-old spark transmissions, now re-emerges as ultra-wideband radio. On February 14, 2002 the FCC\textsuperscript{8} adopted the formal rule changes officially permitting ultra-wideband operations. The ruling defines access to a 7,500 MHz wide swath of unlicensed spectrum between 3.1 and 10.6 GHz that is made available for commercial communications development in the United States. It defines UWB signals as having the lesser of 500 MHz or 20\% fractional bandwidth. Nothing is stated about the signal or modulation method as long as the bandwidth requirements were met. The implicitly allowed, for example, an aggregation of narrow band signals, for example as in OFDM to masquerade as “UWB” under the rules.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure14}
\caption{UWB in the Mainstream}
\end{figure}

The ITU-R followed in 2005 with a Recommendation\textsuperscript{9} on UWB technology. What the FCC rules did was to band limit UWB signals. Figure 15 shows what a UWB pulse could look like before the ruling.

\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{9} RECOMMENDATION ITU-R SM.1757, Impact of devices using ultra-wideband technology on systems operating within radiocommunication services (Question ITU R 227/1) 2006.
\end{itemize}
\end{footnotesize}
Jurisdiction over the radio frequency spectrum in the United States is shared by the FCC, an Executive Branch agency reporting to The President, and the National Telecommunications and Information Administration (NTIA), a Department of Commerce reporting to The Congress. Within the FCC, the Office of Engineering and Technology (OET) provides advice on technical and policy issues pertaining to spectrum allocation and use. The FCC oversees the non-federal users, and allocates spectrum to commercial, private, amateur, state and local public safety departments. Since there are no statutory “Federal” or “non-federal” bands, the frequencies administered by each agency are interspersed across the frequency spectrum. The NTIA, which
is an operating unit of the Department of Commerce, oversees spectrum for Federal government users. The two agencies follow a *Memorandum of Understanding on Joint Use Spectrum*. A reciprocal agreement provides that the NTIA and FCC will give

“notice of all proposed actions which would tend to cause interference to Non-Government and Government station operations. [This] notification must be given in time for [NTIA/FCC] to comment prior to final action...Final action by either agency will not, however, require approval by the other agency.”

Thus neither the NTIA nor the FCC, who appear in different branches of the U.S. government, can act unilaterally in spectrum regulatory matters. Because UWB emissions would span across both Federal and non-federal users, the regulations governing UWB needed coordination and agreement between the NTIA and FCC. The result of this agreement is the FCC Report & Order, which is reflected in the emission mask shown in Figure 17. There are several such emission limit masks, this one is for hand-held devices. The masks are compromises reflecting the needs of both agencies.

![Figure 17 – FCC emission mask for UWB hand-held devices.](image)

It is interesting to note that the peak emission level of -41.3 dBm/MHz between 3.1 and 10.6 GHz corresponds to the maximum “unintentional” emissions of electrical devices. That level was chosen on the basis that the unintentional emission levels were deemed an acceptable. The
Report & Order establishes different technical standards and operating restrictions for three types of UWB devices based on their potential to cause interference. These three types of UWB devices are:

1. Imaging systems including Ground Penetrating Radars (GPRs) and through-wall radar, surveillance, and medical imaging devices
2. Vehicular radar systems
3. Communications and measurement systems

UWB operation is defined as a transmission system having UWB transmitter and is defined (see in terms of):

(a) UWB Bandwidth: the frequency band bounded by the points that are 10 dB below the highest radiated emission, as based on the complete transmission system including the antenna. The upper boundary is designated \( f_H \) and the lower boundary is designated \( f_L \). The frequency at which the highest radiated emission occurs is designated \( f_M \).

(b) Center frequency: \( f_C = (f_H + f_L)/2 \).

(c) Fractional bandwidth: \( BW = 2(f_H - f_L)/(f_H + f_L) \).

(d) UWB transmitter is an intentional radiator that, at any point in time, has a fractional bandwidth \( BW \geq 0.20 \) or has a UWB bandwidth \( \geq 500 \text{ MHz} \), regardless of the fractional bandwidth.

(k) EIRP: Equivalent isotropically radiated power, that is, the product of the power supplied to the antenna and the antenna gain in a given direction relative to an isotropic antenna.

(m) Hand held device is a portable device, such as a lap top computer or a PDA, that is primarily hand held while being operated and that does not employ a fixed infrastructure.

The regulations permit indoor UWB systems and hand-held devices. The emission limits for hand-held systems are shown in Figure 17. UWB operation is spelled out in the Report & Order.

Signals and modulations are not defined, and specifically the UWB bandwidth need not be a consequence of a short time signal, thus a 500+ MHz wide ensemble of OFDM carriers centered at 4 GHz is “UWB” with respect to spectrum access under the rules.

A conceptual UWB transmitter is shown in Figure 18. Data pulses are shown as positive and negative pulses to represent binary 1 and 0. Variable pulse amplitude, on-off keying, and pulse time position, radiated polarization, or any combination could also be used to modulate pulses. These polarity modulated baseband data pulses are shaped in a pulse shaper then multiplied by sinusoid to heterodyne them up to the UWB band and transmitted by an antenna.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{uwb_transmitter}
\caption{A UWB transmitter.}
\end{figure}

\footnote{\textsuperscript{10} US 47 CFR Part 15 <http://www.access.gpo.gov/nara/cfr/waisidx_05/47cfr15_05.html>
Where UWB Technology Might Appear

Today, a significant growth in personal communications is taking place using spread spectrum techniques in blocks of spectrum that have been set aside for unlicensed operations shared by other users. These signals coexist with other signals transmitted in the same frequency bands. This method makes much more efficient use of the congested spectra and allows greatly expanded utilization. A specification for a pulse based UWB physical layer (PHY) is in the IEEE 4a Amendment\(^1\) to the 802.15.4 Standard, and also in the 1394TA Specification\(^2\) for UWB over coaxial networks.

Extending the musical instrument analogy introduced earlier, many different drummers can coexist when their drum beats encode special sequences. Digital techniques can efficiently differentiate among them. Furthermore, flute tones can overlay the same spectrum and convey information as well.

---

\(^1\) IEEE Std 802.15.4a™-2007 (Amendment to IEEE Std 802.15.4™-2006)Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (WPANs) Amendment 1: Add Alternate PHYs.

\(^2\) Networking IEEE 1394 Clusters via UWB over Coaxial Cable-Part 1:Continuous Pulse (C-UWB) PHY 2006019.
Wireless local area networks (WLAN), wireless personal area networks (WPANS) and mobile systems serve different range vs. data rates. UWB stands out as a solution that as a consequence of the available bandwidth, can provide up to Giga bits per second of data at WPAN ranges, and can trade data rate for range to reach out to moderate ranges with modest data rates.

Summary
Ultra-wide band (UWB) technology provides access to a large swath of spectrum that can provide very high speed data rate communications over the personal area network space. UWB changes how signals have traditionally been separated one from another. Narrow band systems separate one signal from another by channelized frequencies, while UWB use separation in time. The bandwidth of UWB signals can greatly exceed the bandwidth of the data modulation, which can result in systems that have great capacity. The matched-filter receiver of narrow band technology has a corresponding matched-template receiver for UWB technology. After agreement between the FCC and NTIA, regulations on UWB were issued by the FCC under US 47 CFR Part 15. Under the regulations UWB is defined in terms a transmitter minimum bandwidth. UWB provides a path to very high data rates at personal area ranges and smaller data rates at extended ranges in the same technology.