For publication in the Boston Audio Society (BAS) Speaker, spring 2014

Digital Audio and Video Fidelity

Ken Wacks, Ph.D.
www.kenwacks.com

Communicating through the noise

For most of history, communications was based on face-to-face talking or written messages sent by courier and post. Long distance communications in real-time was limited to shouting and pre-arranged bon-fires, smoke signals, or drumbeats. Distances were extended by relaying signals via intermediary messengers (usually located on strategic hilltops).

With the advent of electricity in the 19th century, major advances in communications were introduced. Messages could now be represented by electrical parameters (voltage, current, the frequency of alternating current, etc.), which enabled greater distances to be spanned while maintaining intelligible signals.

A fundamental tenet of communications is that perfect signal transmission is nearly impossible. The smoke or flare might be obscured by haze, and electrical signals may be confused with electrical noise. Electrical noise in wires or radio signals is fundamental to the physics of materials because of the random vibrations of molecules. The universe is filled with radio signals from stars (especially the sun). Communication signals compete with this background noise to be detected.

Achieving perfect communications

Much research in communications has focused on the signal-to-noise problem: how to overcome electrical or radio noise so a communications signal can be received over a long distance. In 1948, Professor Claude Shannon of MIT proved the Coding Theorem showing that perfect communications over an imperfect (noisy) channel is possible if the characteristics of the signals are managed properly. He did not develop the tools, but proved that they could exist.

Since the introduction of the Coding Theorem, techniques have been developed that come close to perfect communications. For example, we can transmit huge data files representing computer programs, web pages, audio, and video with no errors. Actually, errors may occur during transmission, but clever techniques locate the errors and correct them before the data file is delivered to the recipient.

What is digital communications?

With analog transmission, a single electrical parameter (such as voltage, current, or frequency) usually represents a single physical measurement like temperature or sound pressure. A digital representation of the value of a physical measurement consists of a sequence of electrical signals.

A simple analog transmission of a temperature ranging from 0 to 100 degrees might represent the temperature reading as an electrical current linearly in the range to 0 to 100 ma. The value of the current is an analog or representative of the real temperature. If there were an average of 20 ma of inherent electrical noise on the communications channel, the recipient would know the temperature at the distant location within an expected error of ±10 ma. If the equipment allowed a stronger signal, say 0 to 400 ma, the same noise would result in a smaller error at the receiver. Figure 1 illustrates a simple analog transmission system.

In a simple digital transmission system the measured signal is converted mathematically into a binary number and each binary digit (bit) is sent as a separate signal. For example, a temperature of 50 is sent as a sequence of six signals representing 110010 (the binary encoding of 50). The two binary digits 0 and 1 are transmitted as electric signals. A 0 might be sent using a 0 ma electric current, while a 1 is 100 ma. The receiver is designed so any signal below 50 ma is interpreted as a binary 0 and any signal above 50 ma is considered a binary 1. Thus, with less than ±50 ma of noise, an error never occurs! This is illustrated in Figure 2 with a noise level of ±10 ma.

Digital can deliver better accuracy, but there are costs compared to analog transmission:

- More complex signal coding.
- More signals sent.
- More complex receiver decoding.

Because more signals are sent, the communications equipment must operate faster to transmit the same number of values as the analog system. Digital transmission can improve message accuracy, but there is a cost in transmission efficiency. Also, the digital system is not perfect. For example, the scheme above for temperature transmission cannot represent fractions of a degree. If we want more precision, we need to change the binary encoding.

Source coding and channel coding

Communications equipment performs two important and distinct functions in processing a signal at the sender and at the receiver, as shown in Figure 3:

- Source encoding / decoding
- Channel encoding / decoding

A good analogy is a letter that is mailed. The process of writing the letter and placing it in an addressed envelope represent source encoding. The shipping and sorting done by the post office constitutes channel processing. With proper channel encoding and decoding, the recipient gets the letter with no damage.

For voice, music, video, etc., source encoding at the sender coverts the source signal into a format suitable for transmitting. This format may be an analog signal with appropriate limits on amplitude and frequency so it can be transmitted. At the receiver the signal is converted back into a form where it can be heard, viewed, or stored, as appropriate.

Channel encoding takes the digital or analog representation of the source and sends it down the communications channel. A channel encoder might be a radio or a modem connected to a telephone line. Most channel encoders represent the transmitted data, whether digital or analog, as physical parameters of an electrical waveform, such as amplitude, frequency, or phase.

All channel transmissions introduce errors because of electrical noise that permeates the universe. Nevertheless, as Shannon proved, it is possible to achieve perfect transmission. Techniques ranging from automatic gain control for analog signals to error check codes for digital signals help achieve this goal.

Channel decoding is responsible for minimizing errors in the received signal. In general, an analog signal may be degraded by transmission errors, but remains useful. Most of us have watched an analog TV show in spite of “snow” or “ghosts” on the screen.

A digital channel decoder attempts to correct transmission errors using error-checking codes that were added to the bit stream at the transmitter. Error check codes can correct only a limited number of digital errors. If this fails to correct the signal, the digital data are corrupted and may be useless. The result on a TV screen may be the loss of a data block (visible as a back square within a TV image) or a frame (the entire picture) for a moment or longer. Thus, digital channel transmissions tend be all or nothing, while analog transmissions gradually degrade.

Digital = Reproducible and Predictable

When digital channel encoding and decoding work properly, the signal delivered to the source decoder input matches the output of the source encoder. Errors are usually not introduced by the source decoder. However, communications between the sender and recipient many still not be perfect because the source encoder may have degraded the sound, the music or the picture. This degrading may have been done deliberately in order to use a transmission channel with limited capacity or to combine many sources onto one channel.

When digital channel encoding and decoding are working properly, the recipient is delivered what the sender planned. Thus, a broadcaster knows that almost all customers receiving a digital signal will get exactly the same content. Except for variations in TVs and radio sets, the broadcaster has more control over the quality of the content with digital transmission than with analog.

How good is digital?

Digital is as good or as bad as the broadcaster or CD producer chooses. Perfect data transmission does not imply that the content will have perfect quality.

For digital TV, the broadcaster may choose to cram multiple TV programs onto one channel, rather than use the entire channel for an HDTV program. This is done at the source encoder where the degree of data compression is selected by the broadcaster. If too many programs are crammed into one channel using digital techniques, the resulting picture may actually look worse than a former analog transmission in a strong signal area (without noise and ghosts).

MP3 is an audio source encoding technique for audio that requires much less data than a CD through the use of clever data compression. However, even though MP3 data can be transmitted digitally with no errors, the MP3 reproduction may not sound as good as a CD. When the CD was converted to MP3, data were deliberately removed. Most BAS members can hear distortion in some MP3 music as a result of data compression.

Digital is different

Digital offers the producer flexibility in controlling what the customer sees or hears. Most digital equipment depends on electronics with processors and memory like a computer. Therefore, it is feasible to design encoders and decoders that are programmable for field upgrades. This makes digital technology more flexible and extensible than analog technology.

Even though the cost of electronics keeps falling, the amount of memory and processing power embedded in consumer electronics keeps increasing. Therefore, digital components are not necessarily less expensive than the former analog devices.

For sure, digital transmission offers benefits, but it is not a panacea. Thus, buying a product labeled “digital” does not imply the highest audio fidelity, a perfect picture, a high-definition picture (HDTV), or a product that costs less than analog.

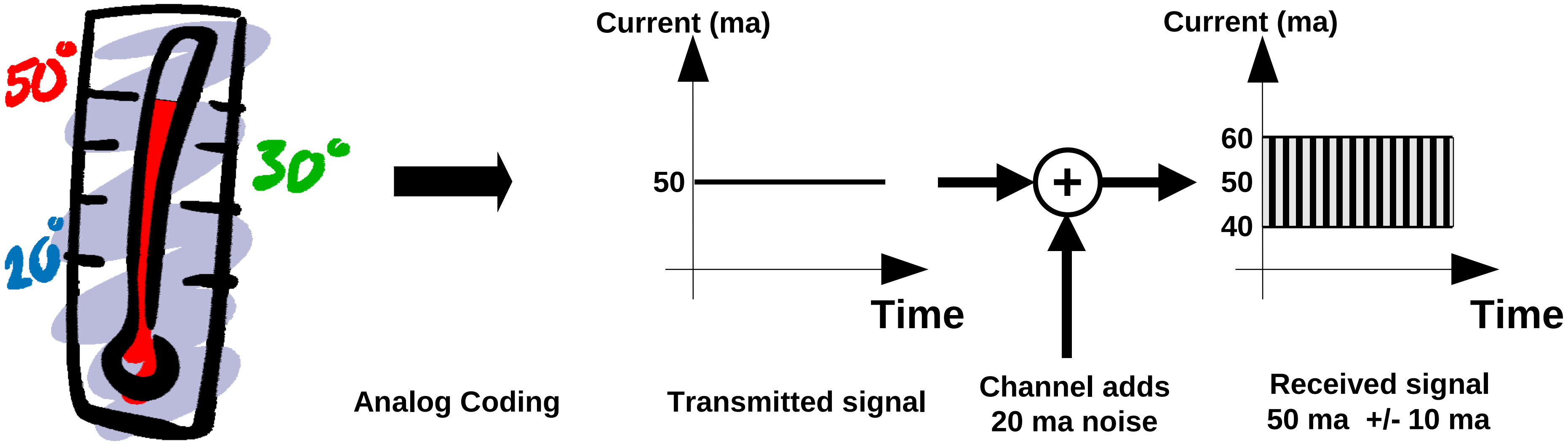

Figure 1 – Analog Transmission

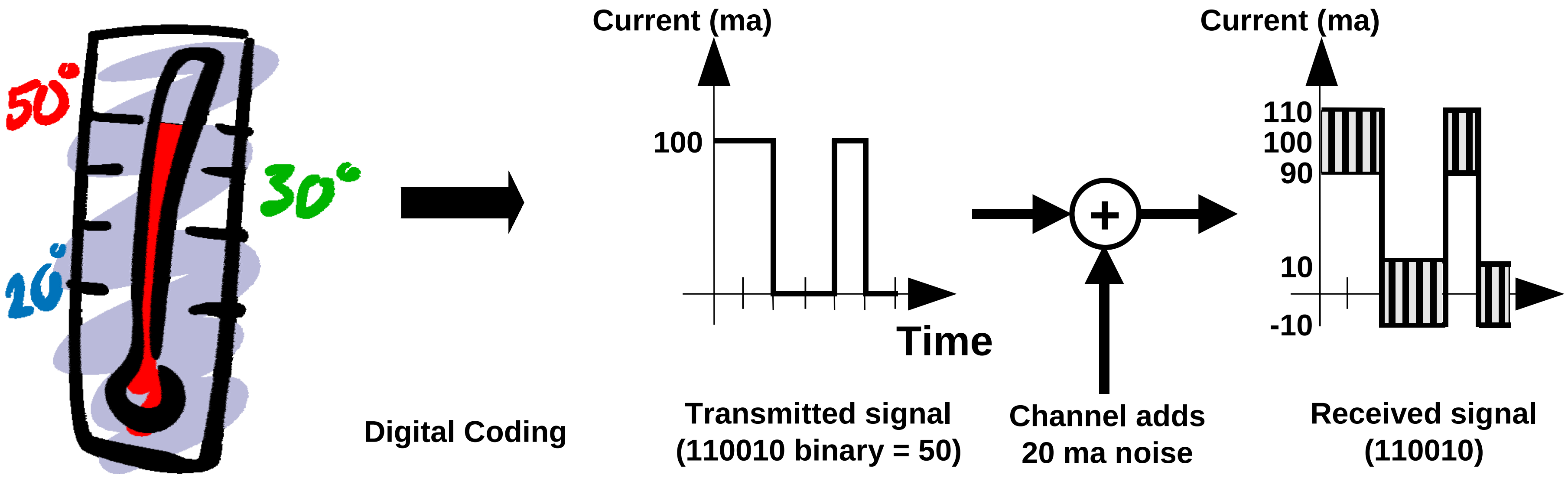

Figure 2 – Digital Transmission

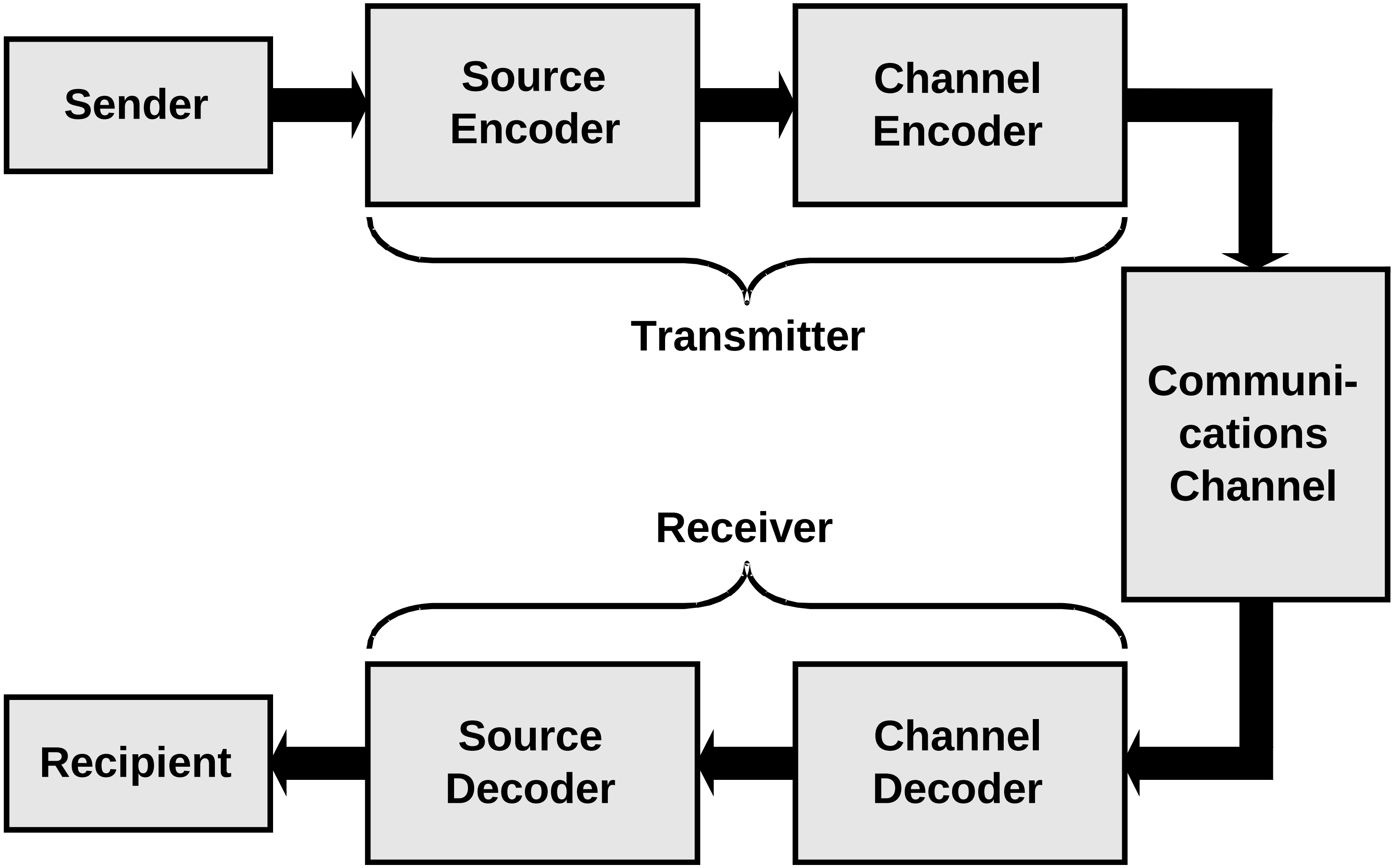

Figure 3 – Source and Channel Coding