Lecture notes for EECS 142:

Slide 1: Introduction

Prof. Ali Niknejad will be the principal lecturer for this course, but is unavailable today.

Dr. Joel Dunsmore will be presenting this lecture, and several lectures during the semester related to RF measurements incorporated as part of this course

Contact information: Dr_Joel@agilent.com (until his Cal email gets set up).

Brief Biography:

Since graduating from Oregon State University with a BSEE (1982) and an MSEE (1983), Joel Dunsmore has worked for Agilent Technologies (formerly Hewlett-Packard) at the Sonoma County Site. He received his Ph.D. from Leeds University in 2004. He is a senior design engineer working for the Component Test Division. He was a principle contributor to the HP 8753 and PNA family of network analyzers, responsible for RF and Microwave circuit designs in these products. Recently, he has worked in the area of non-linear test including differential devices, and mixer measurements. He has received 21 patents related to this work, has published numerous articles on measurement technology, as well as consulting on measurement applications. He has taught electrical circuit fundamentals at the local university, and presented several short courses and seminars through ARFTG, MTT, EMC, and Agilent.

There will be some changes in the course material, and especially in the lab work, to focus on measurement-based simulation using brand-new modern tools including spectrum analyzers, network analyzers, high frequency oscilloscopes, and signal generators. There will be extensive use of simulation for linear and non-linear behavior.

Slide 2: Block Diagram of Communication System

The focus of this course is on circuit design for Communications Systems, and in most cases this will include frequency converting, in some way, the information.

POTS and Telegraph

In the oldest days, the Plain-Old-Telephone-System (POTS), provided direct connection between the source and the receiver, with essentially direct creation of the analog signal onto the line using a microphone which produced amplitude-modulated (AM) DC on the line. The receiver converted the AM signal back to audio (sound), through the speaker. POTS and the even older telegraph were some of the only communications system that did not use modulation and demodulation (up- and down-conversion) of signals.

Modern Systems

For our purposes, all modern systems of communication are comprised of the five components we have in this slide. The exception is the case of direct digital connections, where the information source and sink are essentially connected together. However, all wireless devices, and many wire connections, make use a communications channel.

The Channel

Starting at the right-hand side of the slide, we see that many things can be considered a channel. In wireless system, such as a cell phone, the channel is free-space (air). But, consider your walkie-talkie or cell-phone, used in a city. You may be on one side of a building, and your friends are on the other. Does the signal skip around the building, or go right through. If the signal bounces around the building, does it take one or more paths to the receiver? Are more paths better or worse? It turns out that the answer is "it depends!" What it depends upon is the channel properties, and how the information is coded and decoded. For wireless communications, the channel's main attributes are loss and delay. The loss of a channel can change dramatically over frequency, with some frequencies (60 GHz and 100 GHz) being very lossy due to oxygen or water absorption. Delay in the signal can cause multipath response to arrive at different times, with different phases, effectively cancelling out the signal at the receiver. For wireless devices, the channel consists of the portion that the communication system that the designer has effectively no control over. So, the goal of the wireless design is to create a circuit that uses the channel most effectively.

As the cost of radio components continues to shrink, one approach to maximizing the channel is to use lots of radios to split the signal up and send it over many channels, using multiple-input-multiple-output (MIMO) designs. These techniques have only been put into practice with the last 10 years.

Wired Channels

Wired channels are those for which a direct connection is made between the transmitter and the receiver. For RF signals, these have traditionally been made using **coaxial cables** which exhibit broad bandwidth and relatively low loss, **waveguide** which is limited in their bandwidth but provide much lower loss and can handle much higher power, and **fiber optic cables**, where the modulator/transmitter essentially converts the baseband signal to an optical signal, and the receiver detects and modulated light.

That's about all we're going to talk about channels for now. If we make good progress in the Lab, you'll have a chance to investigate some of the properties of channel propagation for yourself.

Baseband – Information sources and sinks

The information sources are and sinks are really another way of describing how the communications systems interface with the human operator. A cell phone has at least 3 interfaces: Audio (does anyone just talk on a cell phone?), video (www and photos), and textual (texting or SMS). The information source can be processed in many ways to reduce the time or energy it takes to transmit the information, or increase the range and reliability of the signal. Voice signals are coded using various techniques (CODECs) to reduce the information bandwidth required to transmit them. Uncoded, the POTS system used about 3 kHz of bandwidth to transmit a voice signal. But voice signals have a lot of redundancy in them, so that it is possible to compress the signals dramatically. Compare the space require for a CD (which is essentially uncoded) to that for an MP3 or IPOD which are compressed but to a very high fidelity. For voice communications, much less fidelity is required, which means more compression can occur. For digital communications (texting) only enough bits need to be sent to encode the text. Of course, the earliest texting was using morse code to transmit messages. There were only 3 symbols, dot, dash and space, and the encoding of the alphabet was incorporated in these 3 symbols.

The baseband compressing and coding does not always lower the bit count, in fact for some systems, such as CDMA, each data bit is encoded (broken up) into many more bits (sometimes called chips), which form the coded signal. The chip rate can be much higher than the bit rate, and thus the coded signal requires a much higher bandwidth than the uncoded signal. This is done to provide a certain redundancy that allows the coded signal to be smaller than the noise, or interfering signals, yet still be decoded.

The receiver side of the system provides the information sink, where the signal is decoded and converted into a usable form, typically for human consumption.

Now that we've discussed the parts of the system that we aren't going to study, let's discuss the parts of the system that we will be studying, and designing.

Slide 3: A Cell Phone

A cellphone provides a very good example of a communication system. Here we show a greatly simplified block diagram of a cell phone. In practice, the cell phone will have 3 or four radio channels, each with multiplexing of the transmitter and receiver depending upon the bands used.

The cell phone bands use carrier frequencies in the range of 800-950 MHz, 1700-1950 MHz, and around 2100 MHz, in various bands for various services. We will concern ourselves with the radio portion, and maybe just a bit more. The antenna is often in the realm of the radio designer, and may be incorporated in the design of the power amplifier or receiver. The output of the radio is typically an intermediate frequency that is digitized by the DSP, or may be a baseband signal. Similarly, the DSP will produce a digitally generated analog signal that is used by the transmitter. The rest of the radio is very interesting, but not the RF designer.

Slide 4: Non-Linear and Time Varying Circuits

Here we have the pre-requisites for this course.

This may well be your first course where the non-linear behavior of the devices is both not desired, and cannot be avoided. Digital circuits do their best to encapsulate a model of the underlying electronics in such a way that a gate can be described only by its logic family and its logical function. The interaction between gates is assumed ideal, and one of the only non-ideal issues that you deal with is timing (race-conditions) between gates. Digital electronics is a logical based discipline, whereas RF design tends to have a bit of art to it. Sometimes it seems to be a Dark Art (as they say in Harry Potter), but an art none-the-less. Art is another way of saying that the interactions are so complex, it is difficult to model accurately and with confidence. We'll learn in the lab that the non-ideal parts of the design tend to be the cause of 80% of the design difficulties.

Time-varying circuits are even more fun than non-linear circuits. With non-linear circuits, such as amplifiers and oscillators, you study the signal and its harmonics. With time-varying circuits, such as mixers which often have multiple input signals, you get to study the interactions of these signals. And, as we'll learn, you can use time varying properties to make non-linear circuits, and vice-versa.

Slide 5: Zoom in on Receiver

This is a block diagram of a generic receiver circuit. Each of these blocks provide a key contribution to the system, and are required to avoid difficulties associated with the channel.

The **Antenna** is the first element of the receiver, and is what converts electromagnetic radiation into a detectable electric voltage or current. Antenna designs can be trivially simple or staggeringly complicated. The simplest antennas are quarter wave above a ground plane, or half-wave dipoles. The antenna patterns of these types are well studied, and the impedance of the half-wave dipole is approximately 73 ohms, which can be well matched to a nominal 75 ohm cable. Matching impedance is important as maximum power is coupled into the receiver when the receiver provides a conjugate match to the antenna. While matching elements may be used, their finite Q (loss factor) will cause additional loss and are best avoided when possible. The wavelength of an RF signal is found by dividing the velocity of the wave by the frequency. In air, the wavelength of 1 GHz is approximately 30 cm.

The **Input Filter** provides the first selectivity in the receiver. It is used to "channelize" the receiver, or make sure that it is only sensitive to signals in its desired band of operation. This filter can also provide matching between the antenna and the next stage.

A **Low Noise Amplifier** is often the first active stage of a receiver, although it is sometimes integrated with the first converting or demodulating stage. The LNA must amplifier the signal to a usable level, without also amplifying the noise. The quality of goodness (or figure of merit) or an LNA is the noise figure, and is defined as the ten times the log of the ratio of the signal plus noise the output to the signal plus noise at the input, in dB. The less used noise factor is the simply the linear version of noise figure: signal plus noise at the output over signal plus noise at the input. The noise figure of the

LNA essentially determines the smallest signal that it can detect. We'll talk about other considerations of the LNA later.

Image Filters can be found at the input and output of the next stage, the mixer or down-converter stage. The image filters are used to keep unwanted products of the mixer from propagating out to other circuits. For example, the local oscillator (LO) in the mixer can leak through an re-radiate out of the antenna. If you want to know why all electronic devices must be turned off during take-off and landing, even if you are not using them, this is one of the reasons: sometimes the filter designs aren't so good.

The **Mixer** or down-converter, provides the frequency translation to bring the RF signal down to a frequency that can be directly manipulated to remove the modulation. This circuit can be created in many ways. Often, the job of the LNA is to boost the signal such that the noise created in the mixer is not the limitation of the receiver.

The **LO** or Local Oscillator provides the means to allow the mixer to down convert the signal. Often the fundamental of the LO is much lower than the carrier frequency, and multiplication is used to create the drive signal to the LO port of the mixer.

The output of the mixer might then go to the **IF Amplifier** which is often provides for variable gain to condition the signal for best performance in the final conversion stage. Depending upon the radio design, the final conversion to **Baseband** is done either with a second converter, or a high speed DSP. In either case, the final conversion depends upon the modulation method used. For complex modulation, which uses both amplitude and phase (or frequency), two output signals are obtained. In complex modulation, the outputs represent the real and imaginary parts of the RF modulation, often called the In-phase (**I**) or Quadrature-phase (**Q**) signals. Sometimes this type of receiver is called a complex detection or synchronous detection.

One thing we haven't talked about so far is synchronization of the receiver with the source. The frequency of the internal LO in the wireless device can easily drift with battery voltage or temperature changes, even if it is locked to an internal reference. Because of this, many systems make use of **carrier lock** or **clock-recovery** methods to try to reconstruct the original carrier frequency. These techniques typically consist of sending known patterns of modulation, from which any frequency offset in the wireless device can be determined by perceived errors in the received data. However, poor analog design in the receiver can make these errors appear worse due to the response of the receiver having a non-ideal time response. An example might be the response of an amplifier to an input signal may depend upon the size of an earlier signal. This is sometimes called **memory effect** and can cause inter-symbol interference and failure of pre- and post-distortion compensation. It is an area of intense research, and currently the state-of-the-art is to simply minimize the effect, rather than fine ways to compensate for it.

Slide 6: The Spectrum Seen by a Receiver

The LNA at the input of the receiver will see a waveform comprised of the **Signal, Noise,** and **Interference.** The signal can be either very large (standing right next to the radio tower) or very small, that is, it can by quite dynamic, and we call the ratio of the biggest signal that can be received to the smallest signal that can be detected the **Dynamic Range** of the receiver. We've already talked about the LNA, which can introduce its own noise into the system. The noise introduced by the LNA is low end limit of the dynamic range. Additionally, the LNA might not be able to operate with very high signals, its gain will **Saturate,** and it will no longer provide amplification of the desired input signal. If this were the only signal present, then this would not really be a problem. In case of amplitude modulated signals, however, the distortion caused by large signals can create significant errors, including inter-modulation distortion. These distortion products become much larger as the saturation of the amplifier increases, as shown in our demonstration. Even in an FM system, where there is only one signal present, saturation of the amplifier can cause out-of-band harmonics that may need to be filtered.

But the maximum signal present in the system is often an **Interfering** signal, which will cause the amplifier to limit, such that the smaller desired signal will not be amplified and passed to the next stage.

Further complicating this problem is the fact that the signal that is received may have been sent out and reflected off of many surfaces before it arrives at the receive antenna. Each of these portions of the power can combine at the antenna, but because they are delayed differently, the phases may add or cancel. As you walk along with your phone, you may find "dead spots" where all the signal amplitudes vary dramatically in the matter of a few feet. This can be caused by either shadowing effects or multipath cancellation.

.

Slide 7: Interference

In a full **Duplex** system, where the transmitter and receiver operate at the same time, on different frequencies, it is necessary to filter out the transmit signal so as to not saturate the receiver. For all systems, it is necessary to not be susceptible to in-band signals for other radios operating in the system. Because many radios share the same frequency band, the band select filter must be wide enough to support all the possible transmit frequencies. Finally, if the design of the input band select filter is not good, or the radio is not well shielded, out-of-band radio signals can cause the LNA to saturate, and the input signal will be blocked.

We will show a response of an amplifier with a two tones at the input. One tone represents the desired signal, and the other represents the blocking signal. When the blocking signal gets large, the desired signal will start to compress.

Slide 8: Power Levels.

For cellular networks, the minimum signal level is typically less than -100 dBm, and the max signal at the receiver might be as high as -20 dBm before the receiver saturates. This is 8 orders of magnitude!

The slide shows the computations for power in dBm, which is the standard definition of power in dB. When we convert this to voltage, we see that we are working with really small voltages. We want to convert this voltages in the **Analog-to-Digital Converter (ADC)** and if we determine the minimum detectible voltage by assuming a reference voltage (typically 1.5 to 3 volts) and a number of bits (8 to 14 typically). For a 12 bid ADC operating from a 3 volt supply, the minimum detectible step is about 75 mV. How much gain do we need then, to detect the 100 uV signal that the antenna supplies? And really we should amplify this up to nearly full scale if we want to detect all the small changes in the signal that a complex modulation would entail.

Slide 9: Filtering

Filtering needs to remove the out-of-band signals that we don't want, and in some systems, we can use narrowband IF filtering to remove even signals that are in-band for our system. Consider a broadcast television receiver (or your XM satellite receiver). It has to have a very broad band input to allow the entire band (500 MHz or so) of signals to be received. But each signal is nearly as large as the desired signal, and their all too small for the ADC to detect. The radio architecture allows us to first amplifier the whole band so that the signal is well above the noise of the first converter. The first converter mixers the whole band of signals down to an IF band, but then if we follow the mixer with a narrow-band filter, one that can receive only one of the desired signals, we can eliminate the rest of the signals before our final VGA stage. This is why there is often a very narrow band filter at the output of a mixer as well. Many technologies are used to generate these narrow band filters, and in the lab you will get a chance to try to design a few yourselves. In the classroom demonstration, a few examples of filters will be shown.

Slide 10: Frequency Translation

The baseband signal must be converted to the appropriate frequency for the transmission channel. Frequency translation takes place in devices commonly called mixers, but sometimes called modulators or demodulators. In all of these, some non-linear operation is required to create the new frequencies (linear time-invariant circuits cannot create new frequencies). The carrier frequency is, in some way, the average frequency that the modulated signal produces. In AM modulation, the baseband is imposed on the amplitude of the signal, so the carrier frequency is essentially constant (a bit of math will show that there will be sidebands associated with the AM signal's frequency content); this is sometimes call the envelope of the carrier. Your local radio station uses AM modulation; it is very simple but very wasteful of bandwidth and energy. If an I/Q modulator is used, AM modulation will result in identical information on the I and Q channels. For simple AM, it is possible to directly modulate the carrier using such elements as variable gain amplifiers, or variable gain attenuators, with the baseband signal driving the gain control.

Other modulations include FM, or phase modulation, and cause the carrier frequency to change during modulation. With FM or phase modulation, the magnitude of $I^2 + Q^2$ will be constant. FM or phase modulation is sometimes achieved by directly driving the voltage of frequency tuning line of a **Voltage Controlled Oscillator (VCO)**.

With more complex schemes, both magnitude and phase of the I and Q signal will be different, and both carry information. These complex modulations are often created in the baseband, and for cost and simplicity reasons, are output from the baseband by at a low frequency, typically a few MHz. This is done because it can be quite difficult to synchronize the FM control with the AM control, and the RF carrier oscillators, and RF amplitude modulators have non-ideal characteristics that yield errors in the modulation; but there are relatively simple circuits for creating an I/Q signal imposed on a low frequency carrier. So, if a baseband signal is created at some lower **Intermediate Frequency (IF)**, then it is necessary to up-convert the signal to the final RF carrier frequency. We use mixers for this up conversion.

The basic up-converting mixer operates by multiplying the IF input by the LO frequency, to create an RF output, down converters work by multiplying the RF input by the LO frequency to create and IF output. The multiplication can be a true multiplication (as in a Gilbert cell mixer) or achieved through a sampling or chopping process, which is more common in diode or FET-ring mixers, or Gilbert cell mixers driven very hard on the current sources.

In the conversion process, both the sum and difference of the input and LO frequencies are created at the output. This can cause some serious system issues, if the IF frequency is low, as the desired image at the output is quite close to the undesired image. Consider a baseband signal with an IF of 1 MHz. If it is up-converted directly to 1.801 GHz, using a 1.8 GHz LO, there will be an image at 1.799 GHz, as well as leakage of the LO out of the mixer at 1.8 GHz. This will require a very good filter to reject 1.799 GHz and pass 1.801 GHz (too good). Filtering with a fractional bandwidth of less than 3% is quite hard, so it is common to do multiple up-conversions to make the filtering easier. In this example, the 1 MHz base band might be up converted to 301 MHz, and using special filter types (called **Surface Acoustic Wave (SAW)** filters, which can have very narrow bandwidths but don’t work so well above 1 GHz) and a 300 MHz LO; the 299 image is removed by the SAW filter. Then the 301 MHz signal is up-converted again using a 1.5 GHz LO, resulting in an 1.801 GHz desired signal, and a 1.199 GHz image, which can be easily filtered.

In modern systems, if the base band IF signal can be made high enough (10-100 MHz) it may be practical to do a single conversion and eliminate one stage of mixers.

Slide 11: Carrier Frequency

The carrier frequency is most often determined by the channel used for communications, and is strictly regulated. When the BW of the modulation is small with respect to the carrier frequency, non-ideal effects in the mixer such as frequency roll-off or phase-shift with frequency are less important. That's why some of the high data-rate communications proposed want to use very high frequencies for their carriers. 60 GHz is now proposed for use with 1 GBit wireless LAN schemes.

If the radio transmitter or receiver portion does not have flat frequency response over the channel bandwidth, then errors in the response will appear to be errors in the modulation, and can degrade the performance of the channel. For some systems, such as communication satellites, the hardware of the radio is specified very tightly, as they don't like to give up any performance due to non-ideal hardware. In fact, many of the satellite systems have very wide bandwidths (500 MHz at 12 GHz), which are often reconfigured to transmit voice, data, video or paging signals, as the demand (and money) changes.

Slide 12: Frequency Synthesis

The frequency synthesis in a wireless system is typically provided by a dividing an RF VCO by some value, comparing it to a portion of a frequency reference, which may also be divided, in a phase/frequency detector. The output of the detector is integrated to drive the tune line of the VCO such that the phase/frequency error becomes zero. The area of frequency synthesis is huge, with many difficulties. For example, in this scheme how is a frequency generated that is not a rational (N/M) product of the frequency reference? If the frequency reference was 10 MHz, how would you get an output frequency of Pi GHz. One technique that is used is called **Fractional-N** synthesis, where the Fout divider has more than one modulus (N can vary, in the classic example between N and N+1). Utilizing precise timing, the VCO operates for several cycles of the reference frequency with a divide number of N, then occasionally is divided by N+1. This allows the average output of the VCO to be any arbitrary number. But, to avoid having the VCO snap between two frequencies, careful loop filtering, and analog phase interpolation is used to keep Fout constant at the desired frequency.

There is a big trade-off between the quality of the output signal, in terms of stability or phase-noise, and the stepping speed of the synthesizer. If the loop filter is very narrow, the output of the oscillator will not have much phase noise due to noise on the tuning line, but it will be hard to change frequencies quickly. There continues to be ongoing research in the new methods to improve, or circumvent, these limitations.

Many systems use a form of **Frequency Domain Multiple Access (FDMA)** to separate users by frequency. In these systems, the frequency precision, and the stepping time between frequency changes are strictly controlled. In other systems, sometimes called **Spread Spectrum**, or **Ultra-Wide Band (UWB)** the timing of the frequency changes is critical as the source and receiver need to move between frequencies synchronously, or the connection is lost.

Slide 13: The Transmitter

As we described earlier, the transmitter is the dual of the receiver, up-converting where the receiver down converts, and this up-conversion process is often more difficult due to images of the mixer.

The transmitter typically must have much higher performance than the receiver, because it is more tightly regulated. A cell phone with a bad LO for the receiver won't connect to the system, but if it had a bad LO for the transmitter, it would affect other users as the transmit signal might lie on top of other users.

One way to improve the efficiency is to not transmit more power than is necessary. In modern systems, there is a software feedback that tells the radio to turn down its transmit power if the receiver has more than enough signal to provide good noise margin. The VGA portion of the transmitter is used for this function, and if the signal amplitude changes rapidly (think of a car driving fast through city blocks), the VGA must have very rapid response, as well.

The **Power Amplifier (PA)** is the heart of the transmitter, and by far the most costly part. Many companies have come and gone trying to strike gold with the ideal PA. As we will see, much of the quality of the transmit signal is due to the PA. The design of PA's has become much more complicated as more and better models of the PA performance are developed. One fascinating area is that of pre-distortion, where compression effects of the amplifier are compensated for by pre-emphasis of the signal in the baseband. But this only works if the model for the pre-emphasis properly models the non-linear behavior of the PA. Some of the most daunting tasks is accounting for PA memory effects, and many designs have been developed that try to minimize this problem so that the pre-distortion techniques are effective.

Finally the output filter of the transmitter must be very robust, as it must handle very large signals without distorting the signal, adding loss, or allowing spurious signals to pass out to the antenna. The output filters of base stations are largest component in the base station, and are designed using esoteric techniques and materials. Even tuning these filters to the proper frequency is a difficult task, and students have been awarded Ph.D.'s just for developing improvements in these areas (including your speaker today).

Slide 14: Transmitter Spectrum

Just as the transmitter doesn't want to have unwanted signals amplified out of the mixer, the amplifier itself cannot be allowed to produce unwanted, or out-of band signals. How might these signals occur? If a complex modulation is used, then multiple frequencies will be present in the signal at the same time. If the amplifier is operating close to saturation, limitations of the amplifier will create inter-modulation signals at its output. We can see this effect in the class room demonstration by driving the test amplifier hard with a two-tone signal.

The power of the transmitter is many orders of magnitude higher than the components in the receiver, and due to this, the efficiency of the transmitter is very important. In today’s cellular base stations, the cost of refrigeration (cooling the transmitter) is the highest single cost in the base station. Even a few percent improvement in efficiency can mean dollars to the bottom line.

The spectrum of a transmitted signal is very tightly controlled, and must have a precise form, fitting inside a spectral mask. Any deviations of the spectrum will lower the performance of the wireless device both for it’s own user, and maybe other users as well.

Slide 15-19: Topics

Here’s a list of the topics that we’ll cover in-depth during the class.

The LAB!

In the lab, we will be focusing on real-world examples using typical components. Your labs will typically be first simulation based; then you’ll have the opportunity to create your circuit on the bench. For a few of you, it’ll be the first time you get the opportunity to “solder your fingers together” as we like to say. During this portion, you’ll discover that the simulations may not have included effects that degrade or even dominate the performance of your circuits. Here’s where you’ll learn about the art of RF design. In fact, these lessons will apply to any state-of-the-art effort that you eventually go into, as even CPU designers are starting to face the limitations that RF characteristics place on increasing clock speed. Good luck and try to have a little bit of fun in the lab.