EECE 5698 Fall 2025
Prelab 3: Ultrasound & Nonlinear Intermodulation

Instructor: Prof. Kevin Fu
TAs: Hui Zhuang, Nuntipat Narkthong
Last updated: 10/06/2025, 09 AM

Submission Deadline: Oct 16, 2025 by 11:45 AM.
Submit your report as an individual on Canvas.

Part 1: Ultrasound

As mentioned in lab02, sound is a type of mechanical waves. The frequency of the waves decides how well different systems can sense the waves. For example, the human hearing systems can sense sounds approximately in the range of 20-20k Hz. Older people will often hear a smaller range of sounds as their hearing systems age.

Above 20k Hz, sounds become inaudible/undetectable by most humans. This is why we call sounds with higher frequencies ultrasounds. Note that the concepts of audible and inaudible should always be defined with respect to a certain system. Common commercial microphones can easily pick up sounds with frequencies higher than 20k Hz, so these sounds are still "audible" to electronic microphones. It is worth noting that this gap between the audible ranges of humans and microphones create a space of adversarial exploitations because an adversary can emit some malicious sound signals that the owner of microphone devices cannot hear but can be picked up by microphones. For example, there exists previous research of using ultrasound beacons emitted by localization devices and received by user smartphones to track users and compromise privacy without users' awareness.

In lab03, we study how adversaries may use inaudible ultrasounds to compromise microphone-based speech recognition systems. This requires more efforts than simply emitting an ultrasound beacon, as now we need to come up with a way to make the microphone systems not only pick up ultrasound signals, but also interpret them as human spoken audible signals because these systems are designed to mimic human hearing and speech systems. Simply put, the question is how to convert ultrasound signals into audible speech signals. And the answer is through modulation and demodulation.

Part 2: Signal Modulation

The core of the ultrasound *dolphinattack* we will explore in lab03 is an amplitude modulation (AM) and demodulation process. Essentially, adversaries modulate the baseband signals that they want to stealthily inject into the microphones onto inaudible ultrasound with AM, and exploit the nonlineary characteristics of the microphone circuits as a natural demodulator to recover the baseband signals.

AM

Amplitude modulation piggybacks low-frequency baseband information on high-frequency carriers. If you are not familiar with AM, we refer you to the Wiki page of AM and this online demo for detailed explanations. You should pay close attention to the “Analysis”, “Spectrum”, “Modulation index (depth)”, and “Demodulation methods” sections of the wiki. In summary, the modulated signals that will be generated by the adversary and inputted into the targeted microphones will take the form of

$$s_{in}(t) = D \cdot m(t) cos(2\pi f_c t) + cos(2\pi f_c t) = [1 + D \cdot m(t))] cos(2\pi f_c t) \quad (1)$$

where m(t) is the baseband signal, f_c is the carrier frequency, and D is the modulation depth.

To understand the math behind AM, you need to know the Trigonometric Product-to-sum identities. Assume the simplest case where the baseband and carrier signals are single frequencies of f_m and f_c respectively, you should understand why the AM signals (after modulation) will contain frequency components of f_c - f_m and f_c + f_m. Furthermore, note that the AM scheme also adds the carrier again, resulting in a f_c component in the modulated signals. For example, with a 50 Hz baseband signals and a 500 Hz carrier, you should get the following time and frequency domain figures when the modulation depth are 100% and 50% respectively.

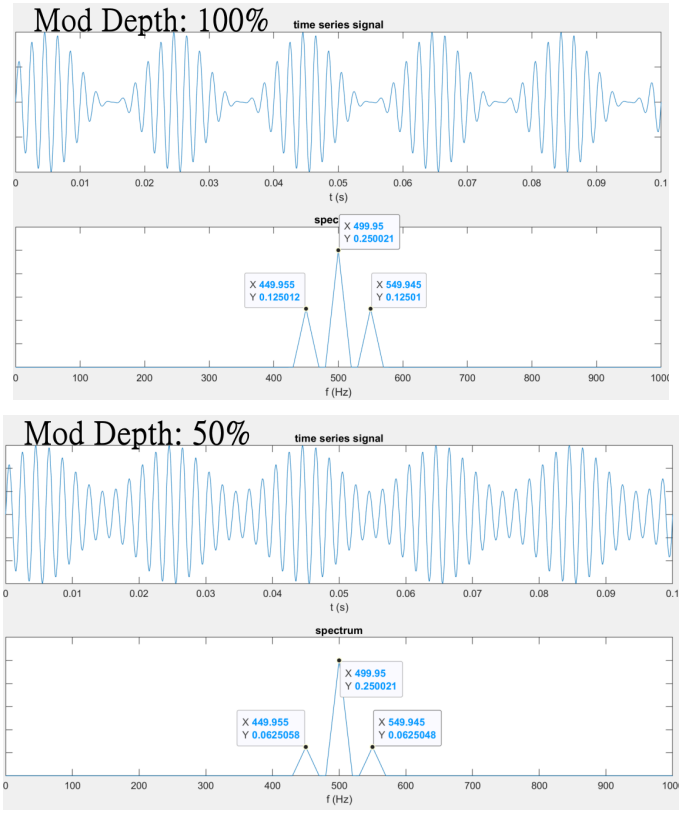

Fig. 0

Question 1:

- Submit a figure that replicates Fig. 0. Specifically, generate by yourself 0.1s signals of the 50 Hz baseband signal and the 500 Hz carrier with the same amplitudes and a sample rate of 100k Hz in either Matlab/Python or Audacity, and perform AM in Matlab/Python according to Equation (1). Show the time and frequency domain modulated signals when the modulation depth is 100% and 50% respectively.
- When plotting the figures, clearly label the frequency axis and only show 0-1000 Hz for your spectrums. The Y axis does not matter.

Non-linearity

We introduced aliasing as a type of ADC nonlinearity in lab02. Here, we introduce another nonlinear phenomenon that widely exists in semiconductors and microphone circuits. Essentially, a microphone can be roughly considered as a component with square-law non-linearity. Let the input signal be $s_{in}(t)$, the output signal $s_{out}(t)$ is [1]:

$$s_{out}(t) = As_{in}(t) + Bs_{in}^2(t) \quad (2)$$

where A is the gain for the input signal and B is the gain for the quadratic term s^2_{in}. A linear component takes a sinusoidal input signals of frequency f and outputs a sinusoidal signal with the same frequency f . In comparison, the nonlinearity of electric devices can produce harmonics and cross-products. It is worth mentioning that this type of non-linearity is an inherent

characteristic of semiconductor transistors's I-V curves. **No transistor has perfectly linear I-V curves even in the so-called "linear regions", but textbooks, designers, and users tend to approximate them as linear for simplity, which allowed for this attack.**

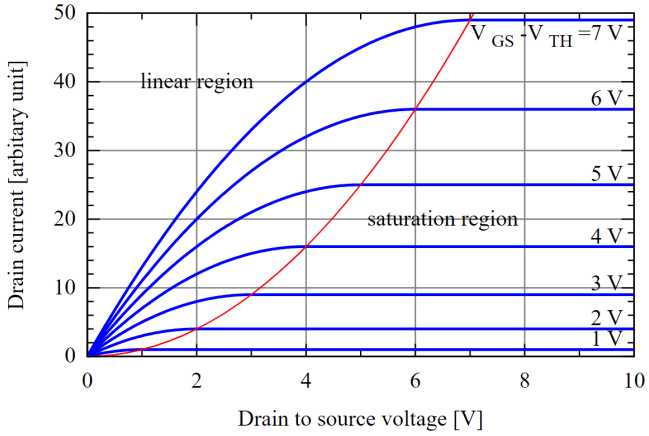

Fig. 1. There is no perfect linearity in even the so-called "linear regions".

Demodulation

Now think about what happens when the modulated signal passes through such nonlinear microphone systems. Plug Equation (2) into equation (1) and you should be able to recover the baseband signal m(t). Note that the conventional AM demodulation process often uses dedicated devices that perform envelope detection or product detection, which are different implementations/mechanisms for achieving demodulation. The nonlinear characteristics of microphone circuits just provide another unintentional and less efficient demodulation mechanism, which is exploited by an adversary who is aware of this. For example, Fig. 2 shows the $s_{out}(t)$ with the 50 Hz baseband and 500 Hz carrier in Question 1 when the modulation depth is 100% and A=1, B=0.2.

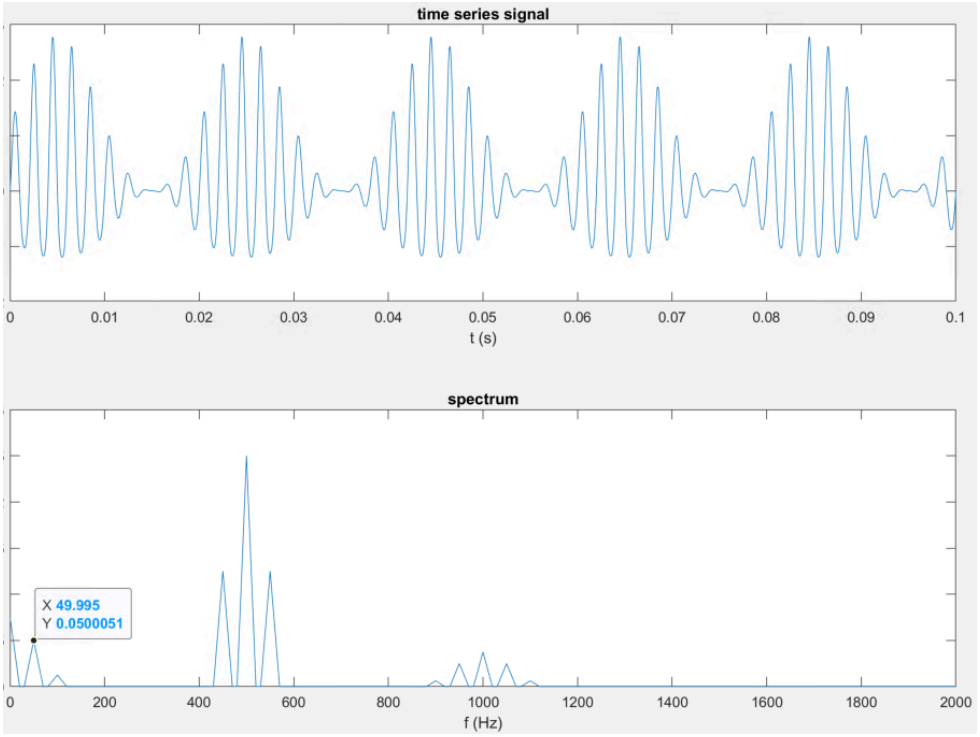

Fig. 2

Question 2:

- Show how a baseband signal m(t) = cos(2pi*f_m*t) is recovered by mathematical derivations using Equation (1) and (2). For simplicity, set A, B, and D as 1. Note that you don't need to do the full derivation. Just show us how the term cos(2pi*f_m*t) is generated in s_{out}.
- Submit a figure that replicates Fig. 2 when the modulation depth is 100% and A=1, B=0.2. You should have 50 Hz recovered. Show a frequency range of 0-2000 Hz for your spectrum.

[1] Zhang, Guoming, et al. "Dolphinattack: Inaudible voice commands." *Proceedings of the 2017 ACM SIGSAC conference on computer and communications security*. 2017.