Manuscript March 1, 2018

On Cuba, Diplomats, Ultrasound, and Intermodulation Distortion

University of Michigan Tech Report CSE-TR-001-18

Chen Yan[1], Kevin Fu[2], and Wenyuan Xu[1]

[1]Department of Systems Science and Engineering, Zhejiang University
[2]Computer Science & Engineering, University of Michigan

March 1, 2018

Abstract

This technical report analyzes how ultrasound could have led to the AP news recordings of metallic sounds reportedly heard by diplomats in Cuba. Beginning with screen shots of the acoustic spectral plots from the AP news, we reverse engineered ultrasonic signals that could lead to those outcomes as a result of intermodulation distortion and non-linearities of the acoustic transmission medium. We created a proof of concept eavesdropping device to exfiltrate information by AM modulation over an inaudible ultrasonic carrier. When a second inaudible ultrasonic source interfered with the primary inaudible ultrasonic source, intermodulation distortion created audible byproducts that share spectral characteristics with audio from the AP news. Our conclusion is that if ultrasound played a role in harming diplomats in Cuba, then a plausible cause is intermodulation distortion between ultrasonic signals that unintentionally synthesize audible tones. In other words, acoustic interference without malicious intent to cause harm could have led to the audible sensations in Cuba.

1 Introduction

This technical report analyzes how intermodulation distortion of multiple inaudible ultrasonic signals could have unintentionally produced audible side effects and harm to diplomats.

In early 2017, diplomats in Cuba suffered hearing loss and brain damage after hearing strange metallic sounds. The news media published reports ranging from scientific analysis of sound recordings [16, 24, 27] to the diplomatic implications [11, 10, 12, 13]. The mystery deepened after physicians published two dueling JAMA papers on neurological damage to diplomats [25, 18]. The news media remained flummoxed on what may have caused the neurological damage [15, 17, 8]. Several news reports suggested that an ultrasonic weapon could have caused the harm. Other experts suggested toxins or viruses. The cause remains a mystery. The substantiated facts include:

- Ultrasonic tones are inaudible to humans.
- Diplomats in Cuba heard audible sounds.

Therefore, any sounds perceived by diplomats are not likely the ultrasound itself. We were left wondering:

1. How could ultrasound create audible sensations?

2. Why would someone be using ultrasound for in the first place?

Assumptions and Limitations. We assume that the sound came from ultrasound, then work backwards to determine the characteristics of the ultrasonic source that would cause the observed audible sensations.

There could be added distortion in the AP audio, so we cannot assume the recordings reflect what humans actually perceived. In one video, the AP news is seen playing a sound file from one iPhone to a second iPhone, essentially making a recording of a recording. Each traversal through a speaker or microphone will add distortion and filtering.

Our experiments focus on spectral properties rather than the effect of amplitude or distance. It might be worthwhile to replicate our experiments with a high powered array of ultrasonic transducers. We do not explore non-ultrasonic hypotheses such as toxins, RF, or LRADs. We also do not consider direct mechanical coupling such as sitting on an ultrasonic vibrator.

Why Ultrasound? It is well known that audible sounds typically propagate omnidirectionally and are difficult to confine to parts of a room. In contrast, ultrasound tends to propagate within a narrower beam than audible sound and can focus a beam towards a more specific area. News reports cited diplomats discussing sounds that were narrowly confined to a room or parts of a room. This type of observation is strongly correlated with ultrasound. We believe that the high-pitched audio signals confined to a room or parts of a room are likely created by ultrasonic intermodulation distortion.

How to Produce Audible Sound from Ultrasound? Humans cannot hear airborne sounds at frequencies higher than 20 kHz, i.e., ultrasound. Yet the AP news reported that "It sounds sort of like a mass of crickets. A high-pitched whine, but from what? It seems to undulate, even writhe." The AP's spectrum plot shows a strong audible frequency at 7 kHz. We believe that this 7 kHz sound is caused by intermodulation distortion, which can down-convert the frequency of ultrasound into the audible range—resulting in high-pitched noises. Nonlinearity typically causes Intermodulation distortion. The engineering question boils down to: assuming an ultrasonic source, how can the audible byproducts consist of a mixture of several tones around 7 kHz separated by 180 Hz?

Sources of Ultrasound. There are many potential sources of ultrasound in office, home, and hotel environments. Energy efficient buildings often use ultrasonic room occupancy sensors in every room (Figure 1). Ultrasonic emitters can repel rodents and other pests. HVAC systems and other utilities with pumps or compressors can vibrate entire buildings. Certain burglar alarm sensors, security cameras, and automated doors use ultrasound for detection of movement.

Researchers from Illinois recently proposed using specially crafted ultrasound to jam microphones [20]. If sounds from an ultrasonic jammer (Figure 2) were to collide with an eavesdropping device attempting to covertly exfiltrate a signal over an ultrasonic carrier, these two signals could combine to produce audible byproducts in both air and microphones. However, if there were an ultrasonic jammer present, it would have likely jammed all nearby microphones, including the microphone used to record the metallic sounds. This would make jammers an unlikely cause.

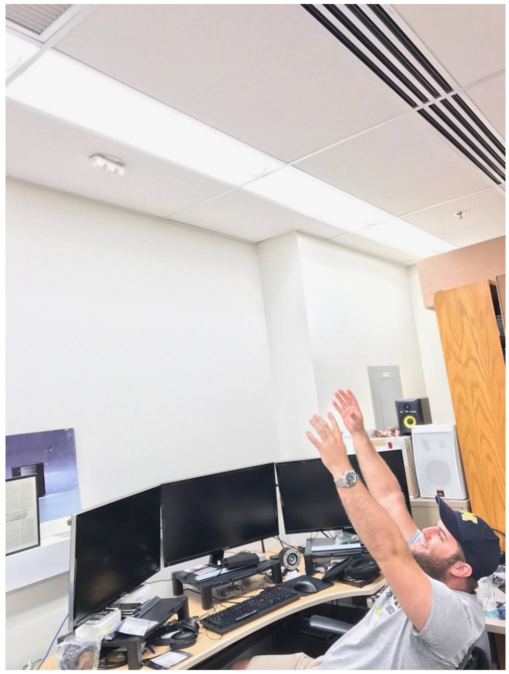

Figure 1: *Michigan Computer Science & Engineering Ph.D. student Connor Bolton notices that an ultrasonic room occupancy sensor in the ceiling had been bathing his experiments with unwanted 25 kHz sounds.*

When introducing additional ultrasonic interferce to an ultrasonic jammer, the signals might render the jammer ineffective while causing unwanted audible byproducts to humans and nearby microphones.

There are also hailing devices such as the Long Range Acoustic Device (LRAD) that many people claim use ultrasound. There may be LRADs that use ultrasound, but modern LRADs tend to use parametric audible sound below 3 kHz. Using an array of several dozen piezo speakers that emit sound in a synchronized fashion to improve directionality, an LRAD can generate sound waves where the wavelength is much smaller than the size of the speaker. Under such conditions (which also tend to be true for ultrasonic emissions), the sound will propagate in a tight, directional beam—allowing long distance delivery of sound.

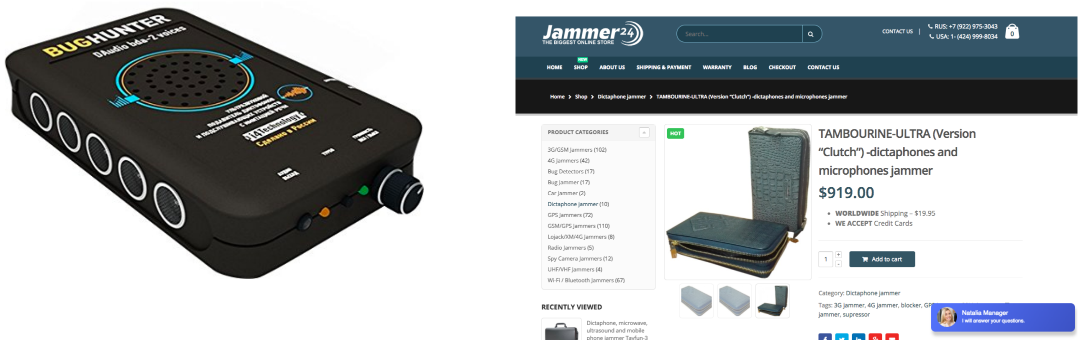

Figure 2: *Commercial products with several ultrasonic transducers can jam nearby microphones. One manufacturer sells a clutch, presumably for fashionable people to jam microphones at cocktail parties.*

2 Spectral Analysis of AP News Audio

We initiate our study with two observations from the AP news: (1) the original audio recordings and (2) description on the high-pitched sounds heard by those in Cuba. Our goal is to construct ultrasonic signals that can lead to similar spectral and audible characteristics.

Audio clips. The AP News [16] published several recordings from Cuba described as high-pitched or metallic cricket sounds[1]. As a common method to analyze signals, frequency spectrum is obtained by Fourier transforms of the original sounds. The AP news performed the spectrum analysis on a smartphone (Figure 3) and shown a spectrum plot centered at 7 kHz (Figure 4 and 5). The spectrum plot demonstrates that there are roughly more than 20 different frequencies embedded in the audio recording.

We acquired the audio from the AP News[2], which claims include a recording of what some U.S. embassy workers heard in Havana. The recording extracted from the video is 5 seconds long, and sampled at 44.1 kHz with 32-bit floats. We analyzed the sound in time (Figure 6), frequency (Figure 7), and time-frequency domains (Figure 8).

After zooming in and looking through the time signal, we found nothing remarkable. No modulation appears in the waveform (at least not ASK), and the waveform does not resemble FSK or

[1] `https://www.apnews.com/88bb914f8b284088bce48e54f6736d84`
[2] https://www.youtube.com/watch?v=Nw5MLAu-kKs

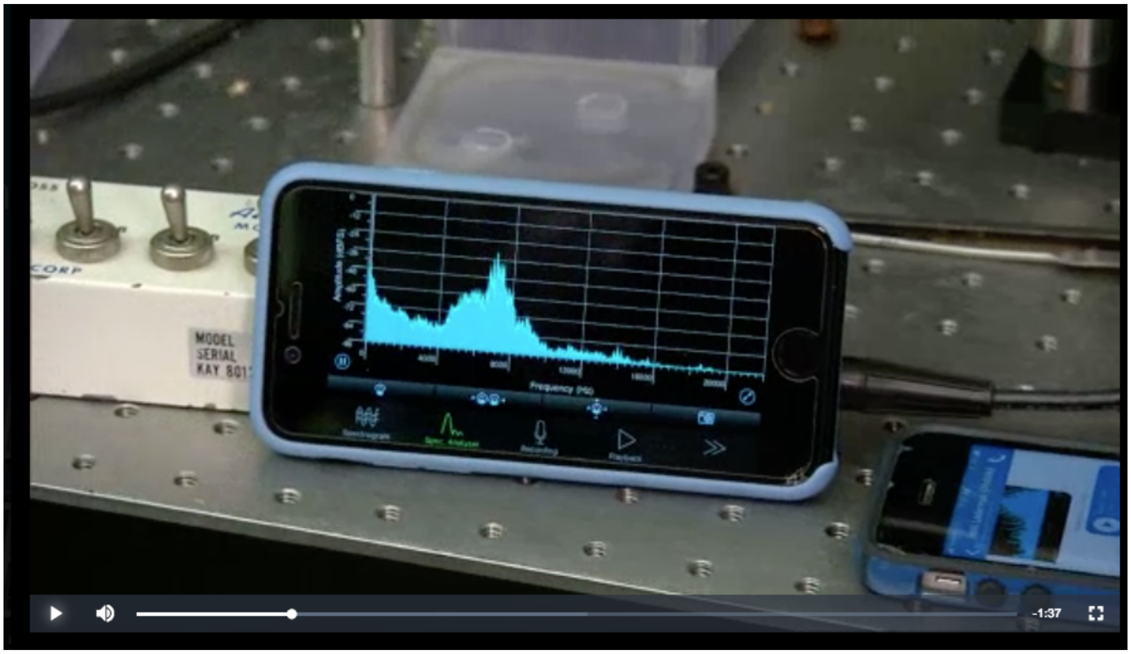

Figure 3: *Screen shot of the AP news itself showing a screen shot of a recording of yet another recording from Cuba. Note that the recording device appears to have removed the spectrum above 14 kHz.*

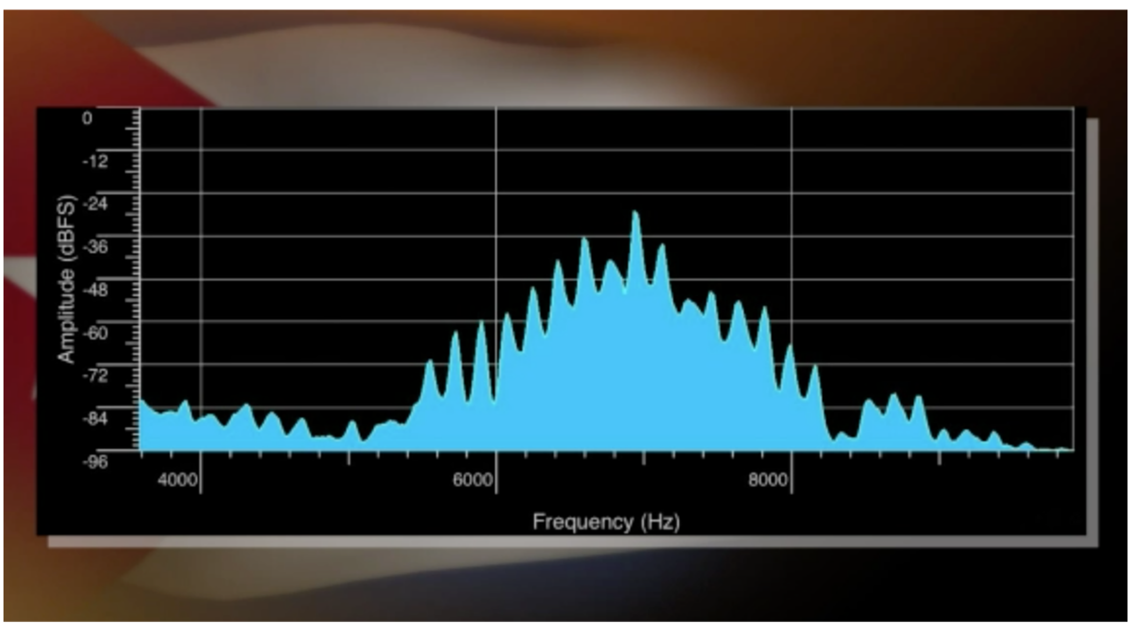

Figure 4: *Screen shot of the AP news showing the Fourier transform of what appears to be a recording of sounds heard in Cuba.*

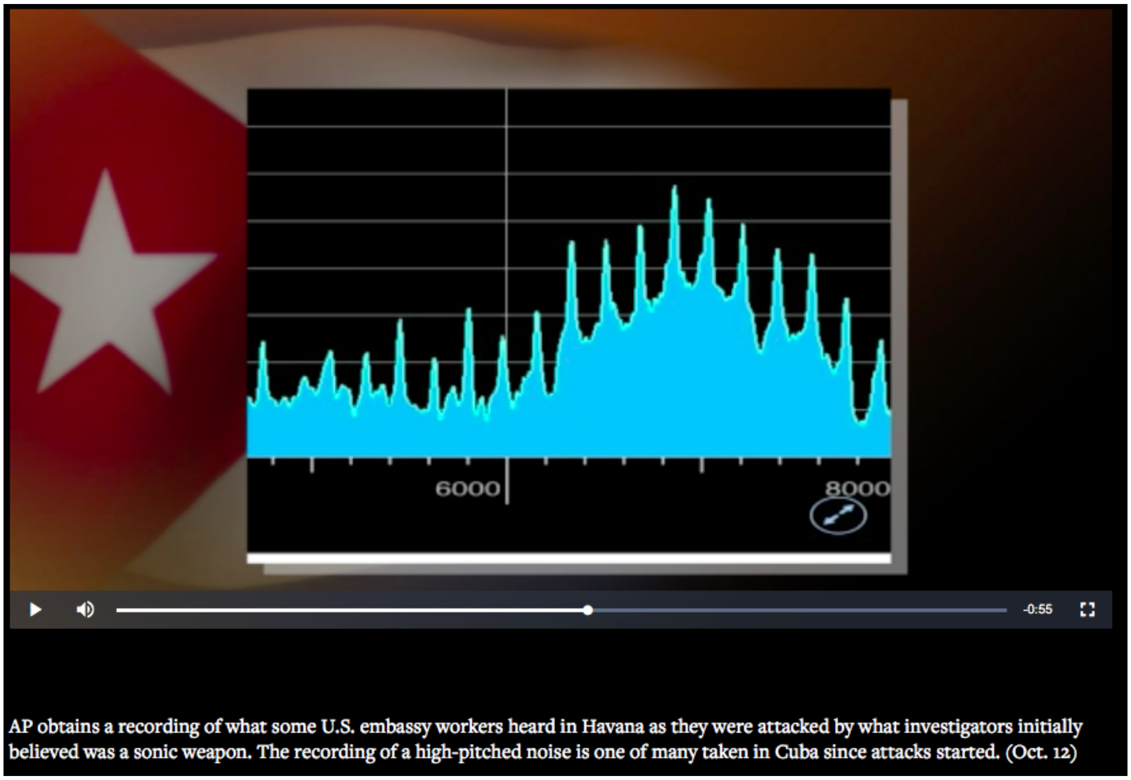

Figure 5: *Screen shot of the AP news analyzing a different recording showing emphasis on spectrum near 7 kHz.*

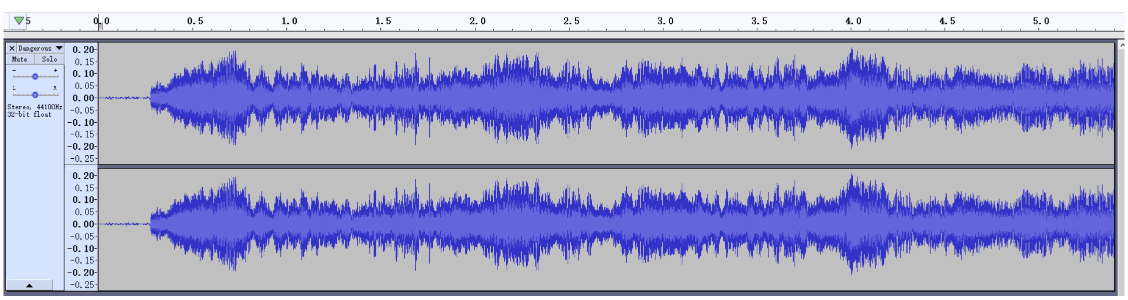

Figure 6: *The time domain signal of metallic sounds extracted from the AP News video.*

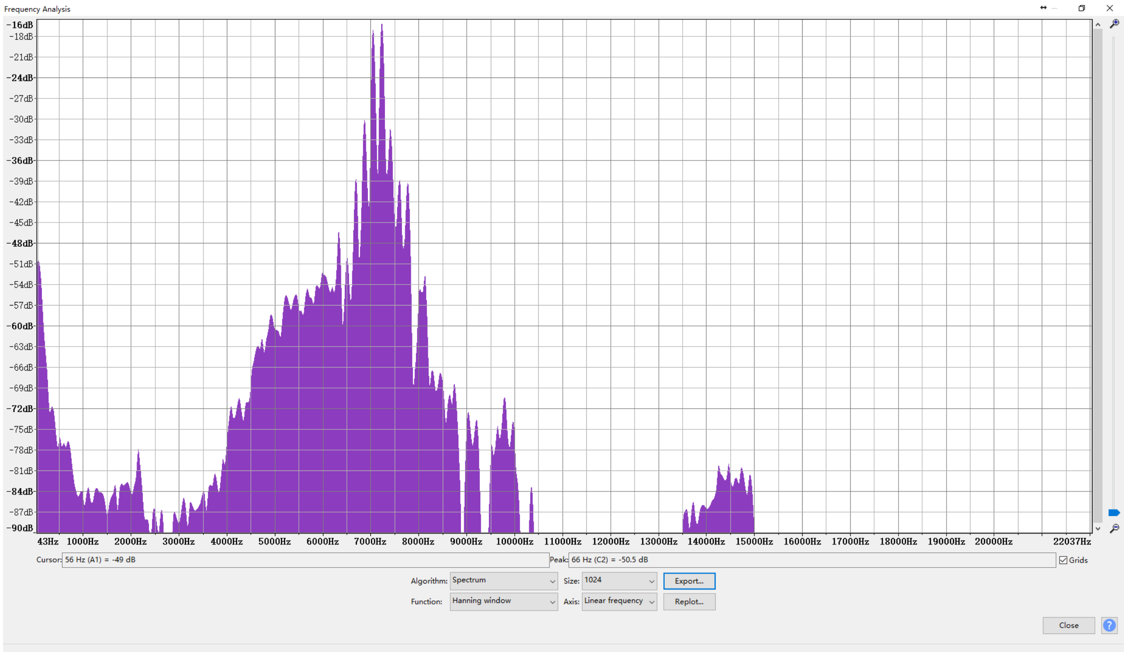

Figure 7: *The spectrum of metallic sounds extracted from the AP News video. The spectrum of the AP news audio ends abruptly at 15 kHz. We suspect this is an artifact of either the AP audio filtering, YouTube audio filtering that is known to roll off beginning at 16 kHz, or iPhone audio filtering that begins to roll off at 21 kHz on our equipment.*

PSK, among other common modulation schemes. We wondered for a moment if someone might be playing a joke on us with fake audio if after demodulation, a message were to tell us it's all a joke. So for giggles, we tried AM demodulation. The resulting signal sounds like a F1 engine and is not likely meant as a message.

The spectral plot (Figure 7) shows major frequency components around 7 kHz. The peaks (6,704 Hz, 6,883 Hz, 7,070 Hz, 7,242 Hz, 7,420 Hz) are separated by approximately 180 Hz.

However, in the waterfall plot (Figure 8), the major frequencies (in yellow) do not change over time. This lack of change again suggests that there is no frequency-related modulation, such as FM or FSK. So wherever the sound comes from, it produces a mixture of several tones around 7 kHz separated by 180 Hz.

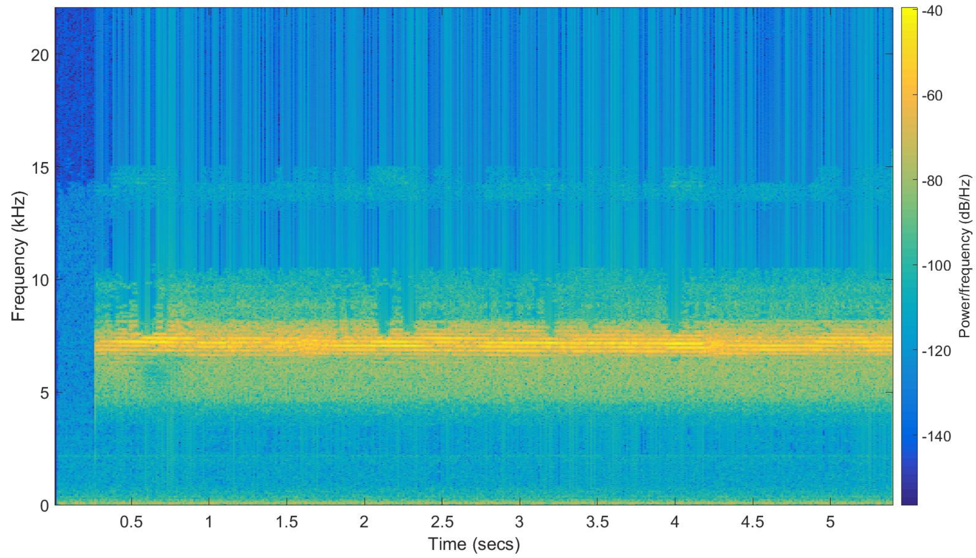

Figure 8: *The spectrogram-time plot (waterfall) of metallic sounds extracted from the AP News video.*

3 Simulation: Intermodulation Distortion of Ultrasound

Intermodulation distortion (IMD) is the result of multiple signals propagating through nonlinear systems. Without loss of generality, a nonlinear system can be modeled as the following polynomial equation:

$$s_{out} = a_1 s_{in} + a_2 s_{in}^2 + a_3 s_{in}^3 + \cdots + a_n s_{in}^n$$

where s_{in} is the system input and s_{out} is the system output. The $a_n s_{in}^n$ for $n > 1$ is called the nth order IMD. When s_{in} contains multiple frequency tones, the IMDs introduce new frequency components.

3.1 Simulation of 20 kHz and 21 kHz IMD

To illustrate the principle of intermodulation distortion independent of what may have happened in Cuba, let $s_{in} = s_1 + s_2$, where $s_1 = \sin(2\pi f_1 t)$ and $s_2 = \sin(2\pi f_2 t)$. When $f_1 = 20$ kHz and $f_2 = 21$ kHz, the spectrum of s_{in} will have two spikes with one at 20 kHz and another at 21 kHz

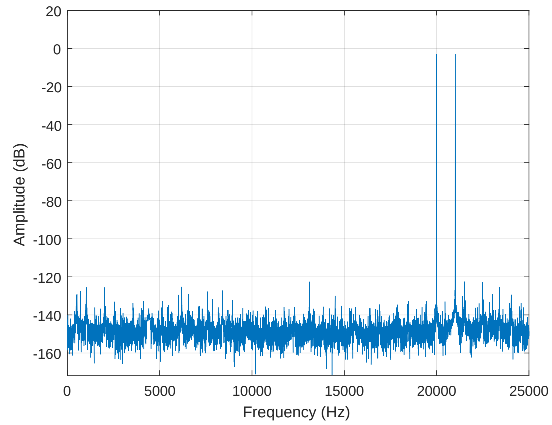

Figure 9: *Spectrum of our input signal with pure tones at 20 kHz and 21 kHz to illustrate effects of IMD in a nonlinear medium.*

(Figure 9).

After the signals pass through the nonlinear system, s_{out} will contain new frequency components that are determined by the order of IMD. Figures 10–11 show the spectrum of the 2nd, 3rd, 4th, and 5th IMDs. For example, the 2nd order IMD introduces new frequencies at $f_2 - f_1$ (1 kHz), $f_2 + f_1$ (41 kHz), $2f_1$ (40 kHz), and $2f_2$ (42 kHz). Notice that $f_2 - f_1$ is below 20 kHz and audible. The 4th order IMD introduces both $f_2 - f_1$ (1 kHz) and $2f_2 - 2f_1$ (2 kHz).

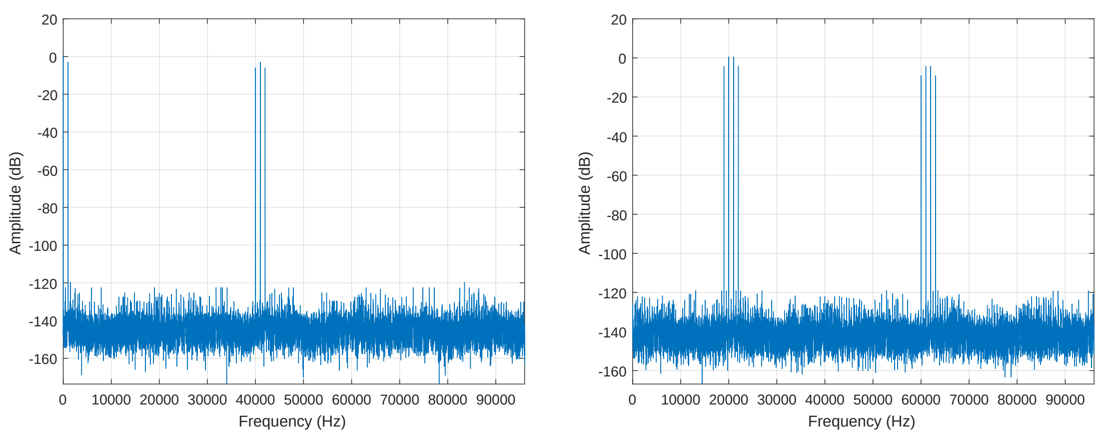

Figure 10: *Spectrum of the (L) 2nd and (R) 3rd order IMD for* $(f_1, f_2) =$ *(20 kHz, 21 kHz).*

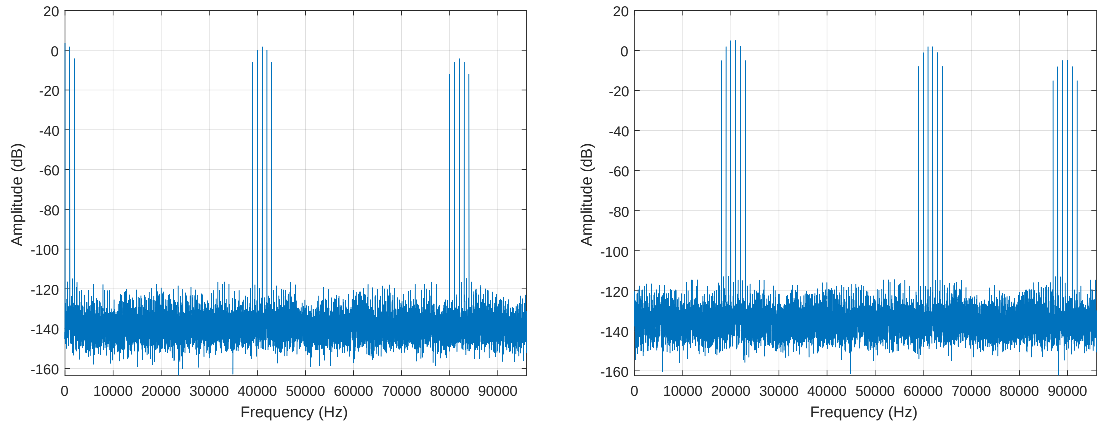

Figure 11: *Spectrum of the (L) 4th and (R) 5th order IMD for* $(f_1, f_2) =$ *(20 kHz, 21 kHz).*

3.2 Simulation of IMD of Three Ultrasonic Tones

In practice, most signals contain multiple tones. To illustrate the effects of IMD on three ultrasonic tones, let us explore the case of three signals at 25 kHz, 32 kHz, and 32.18 kHz. That is, $s_{in} = s_1 + s_2 + s_3$, where $s_1 = \sin(2\pi f_1 t)$, $s_2 = \sin(2\pi f_2 t)$, and $s_3 = \sin(2\pi f_3 t)$, $f_1 = 25$ kHz, $f_2 = 32$ kHz, and $f_3 = 32.18$ kHz. We selected 32.18 kHz to mimic the observation of a 180 Hz separation in the AP news spectrum.

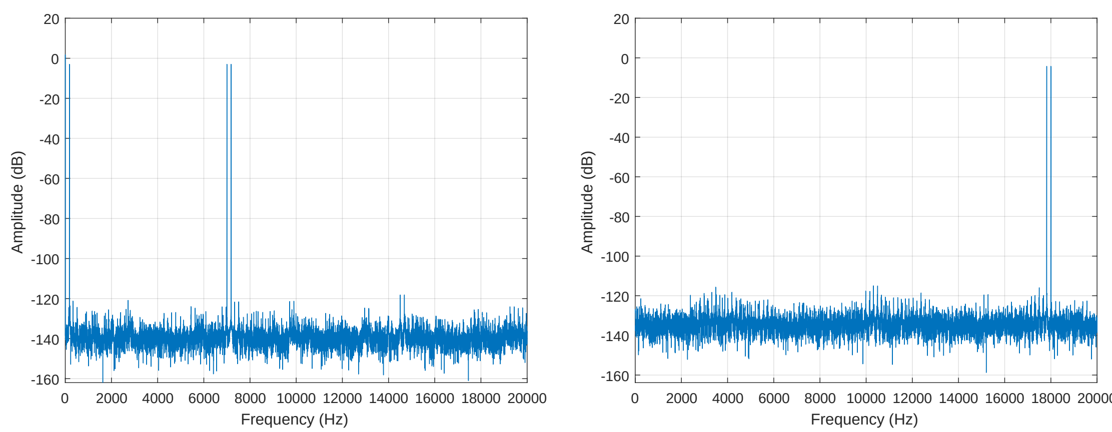

Figure 12: *Audible spectrum of the (L) 2nd and (R) 3rd order IMD for 25 kHz, 32 kHz, and 32.18 kHz tones.*

When there are more than two signals, intermodulation happens between each pair of the signals. In our case, the 2nd order IMD introduces new frequencies (below 20 kHz) at $f_2 - f_1$

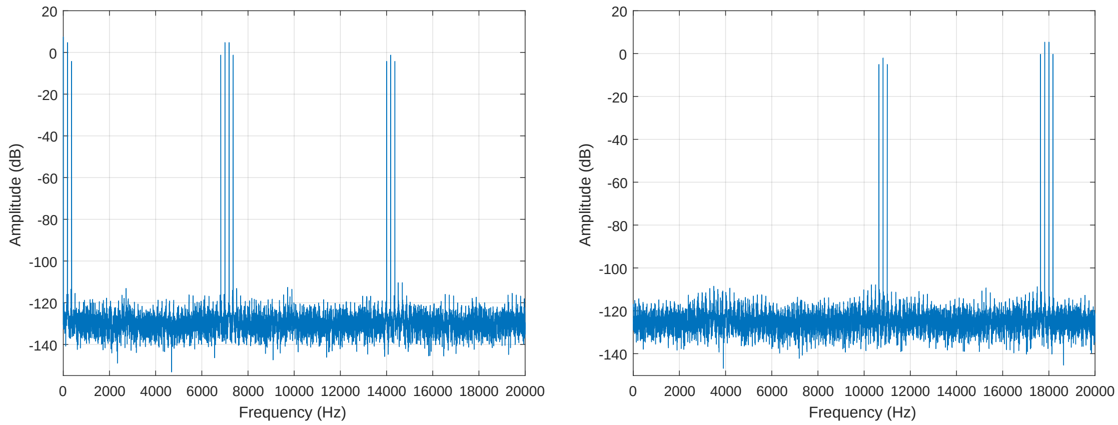

Figure 13: *Audible spectrum of the (L) 4th and (R) 5th order IMD for 25 kHz, 32 kHz, and 32.18 kHz tones.*

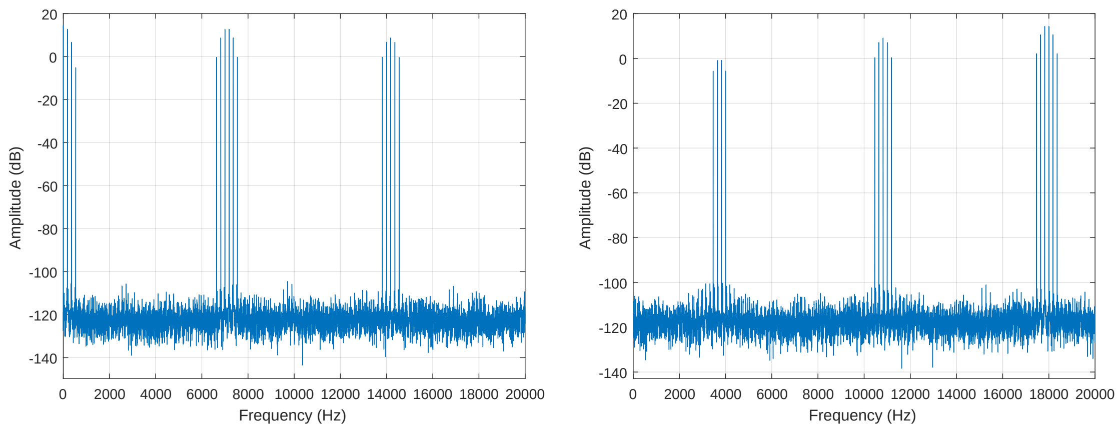

Figure 14: *Audible spectrum of the (L) 6th and (R) 7th order IMD for 25 kHz, 32 kHz, and 32.18 kHz tones.*

(7 kHz), $f_3 - f_2$ (180 Hz), and $f_3 - f_1$ (7.18 kHz). If there are more signals (e.g., another $f_4 = 31.82$ kHz), more IMD products are generated — $f_4 - f_1$ (6.82 kHz), and $f_3 - f_4$ (360 Hz), and existing IMD frequencies are enhanced ($f_2 - f_1$ (180 Hz)). The higher order IMD products (4th, 6th, 8th, etc.) will generate more frequencies around the existing ones (7 kHz and 180 Hz) with a separation of 180 Hz, and create new frequencies. For example, the 4th order IMD introduces new frequencies (below 20 kHz) at $f_3 - f_2$ (180 Hz), $f_3 - f_2$ (360 Hz), $2f_2 - f_3 - f_1$ (6.82 kHz), $f_2 - f_1$ (7 kHz), $f_3 - f_1$ (7.18 kHz), $2f_3 - f_2 - f_1$ (7.36 kHz), $2f_2 - 2f_1$ (14 kHz), $f_2 + f_3 - 2f_1$ (14.18 kHz), and $2f_3 - 2f_1$ (14.36 kHz). With the increase of IMD orders, there will be more frequency peaks

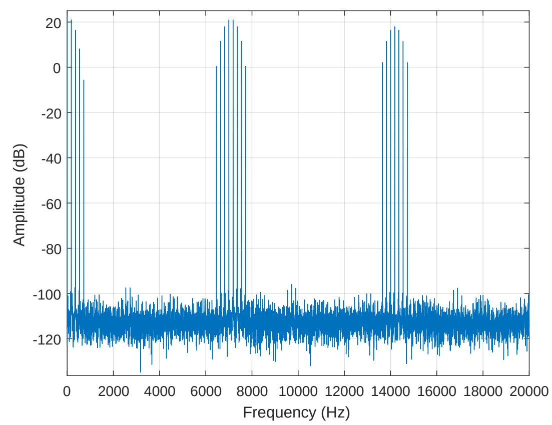

Figure 15: *Audible spectrum of the 8th order IMD for 25 kHz, 32 kHz, and 32.18 kHz tones.*

rippling around 180 Hz, 7 kHz, and 14 kHz. Each ripple will be separated by 180 Hz.

Now consider the audible frequencies produced by all the IMDs up to and including the 8th order summed together in Figure 16. The peaks near 7 kHz are beginning to resemble the AP news spectrum.

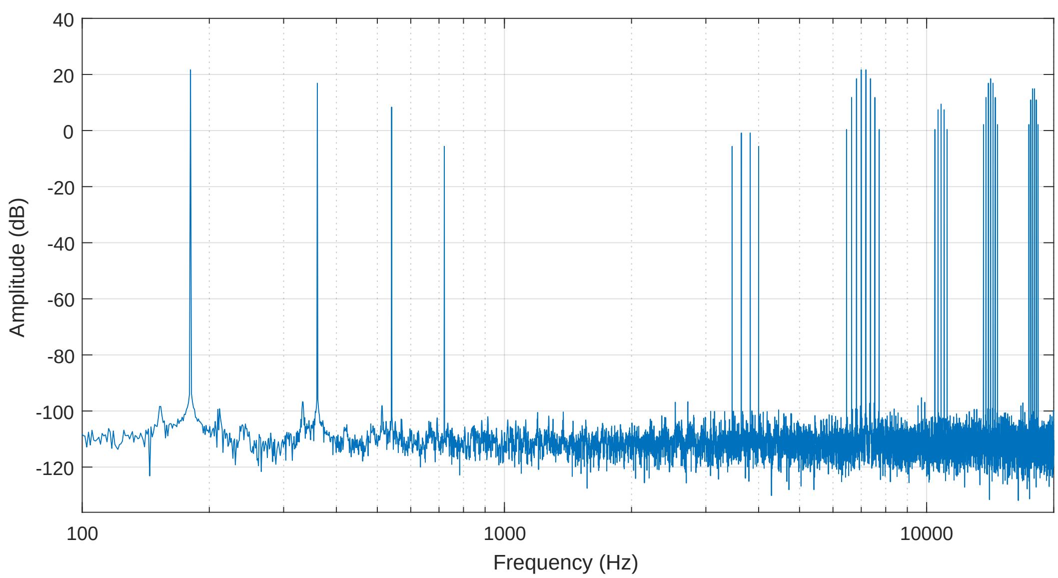

Figure 16: *Log-scale cumulative audible spectrum of 2nd though 8th order IMD for 25 kHz, 32 kHz, and 32.18 kHz.*

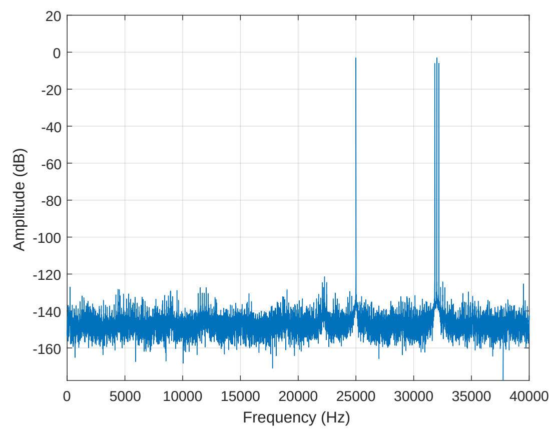

Figure 17: *Calculated spectrum of 25 kHz tone and 180 Hz AM modulated over a 32 kHz carrier.*

3.3 Simulation of IMD of Ultrasonic Modulation

To generate similar intermodulation of three ultrasonic tones, it is feasible to explore the IMD for two signals where one is modulated on an ultrasonic carrier. In particular, to generate signals similar to the recording, i.e., signals centered at 7 kHz with a serial of multiples of 180 Hz signals nearby, we can utilize two signals and their intermodulation. Let $s_{in} = s_1 + s_2$. One of the signals can be a single tone, $s_1 = \sin(2\pi f_1 t)$, and the other will be a signal that is modulated with a single tone of 180 Hz. In particular, we utilize amplitude modulation (AM) that produces double-sideband and transmitted carrier. For example, when the baseband signal is a single tone at $f_m = 180$ Hz, and the carrier signal is at $f_c = 32$ kHz, AM with transmitter carrier will produce an output of $s_2 = \sin(2\pi f_c t) + \sin(2\pi f_c t)\sin(2\pi f_m t)$, which can be seen as the combination of three signals at f_c (32 kHz), $f_c + f_m$ (32.18 kHz), and $f_c - f_m$ (31.82 kHz), as shown in Figure 17. When IMD happens between such an AM signal and a $f_1 = 25$ kHz single tone, the result will be exactly the same as the previous example — signals around 7 kHz, 180 Hz, 14 kHz, and more.

The spectrum of the simulated IMD through the 7th order products with input of 25 kHz and 180 Hz AM modulated on a 32 kHz carrier is depicted in Figure 18 and Figure 19 (log-scale).

If the baseband signal is not a 180 Hz tone, but music or something else with many tones, it will only change the separation (f_m) of the smaller peaks. The recovered signals always remain at around 7 kHz, 14 kHz, and 18 kHz. If we only consider the strongest 2nd order product, there will

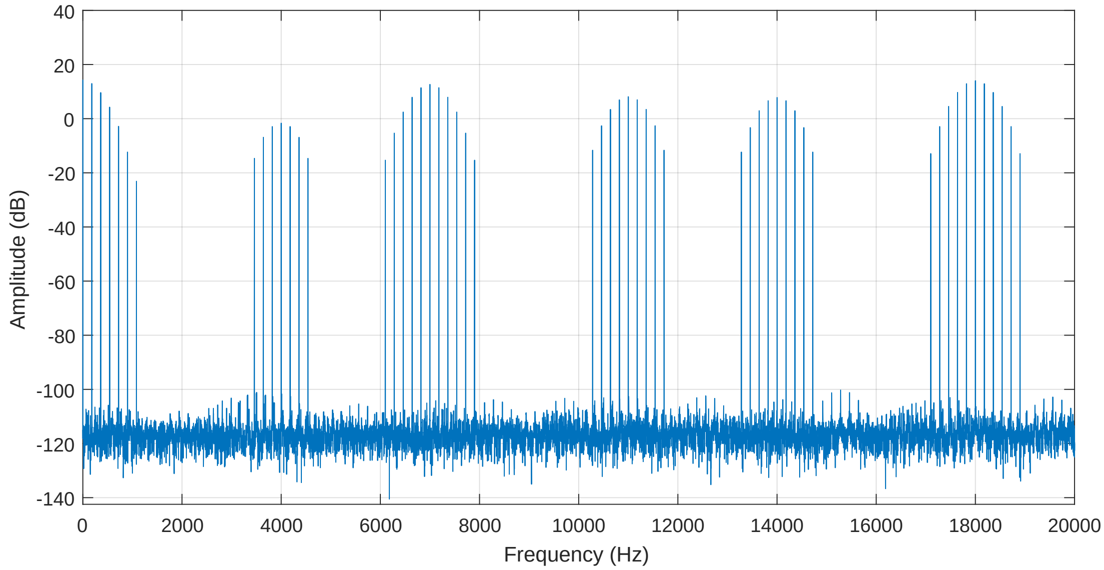

Figure 18: *Cumulative audible spectrum of 2nd though 7th order IMD for 25 kHz tone and 180 Hz AM modulated over a 32 kHz carrier.*

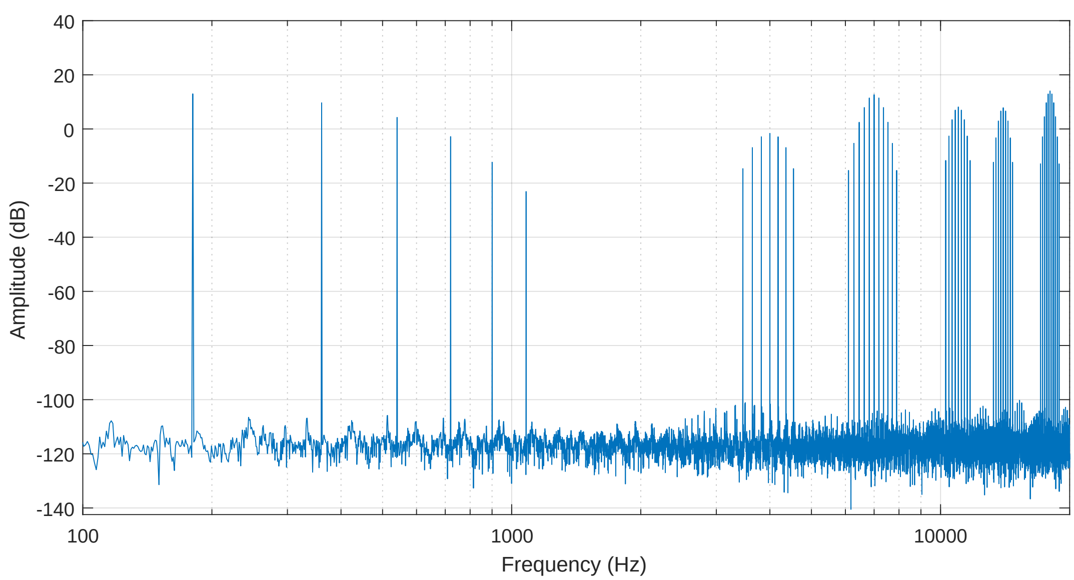

Figure 19: *Log-scale cumulative audible spectrum of 2nd though 7th order IMD for 25 kHz tone and 180 Hz AM modulated over a 32 kHz carrier.*

be signals at only around 7 kHz.

3.4 Discussion of Simulation Results

Different systems (e.g., recording devices) have different nonlinear properties that determine the strength of each order of IMD products. In the simulations, we use a_i coefficients of unity weight for the strengths. If we were to obtain the recording devices and emitters from Cuba, we could deduce the coefficients. We surmise that the reason that there are no obvious frequencies at 4 kHz, 11 kHz, and 18 kHz in the original AP news recording is because the intermodulation products at the odd orders are weak relative to the 2nd and 4th order IMDs on whatever devices recorded sounds in Cuba.

The IMD can also happen multiple times. IMD may occur during air-borne transmission. The IMD can happen again inside the circuitry of a microphone as well as in the human inner ear itself. Thus, the perceived sounds will differ depending upon where one is listening and what are the characteristics of the microphone.

3.5 Summary of IMD Simulation

Our simulations confirmed the feasibility of reproducing the acoustic spectrum of the AP news recording with the intermodulation distortion of multiple ultrasonic signals. Notice that in the spectrum of the AP news recording, there were also frequency components at 180 Hz (not obvious), 360 Hz, 540 Hz, and around 14 kHz.

4 Experiments

With the theories validated by the Matlab simulations, our next step was to generate real ultrasonic signals that caused audible sensations that mimic the sounds heard in Cuba.

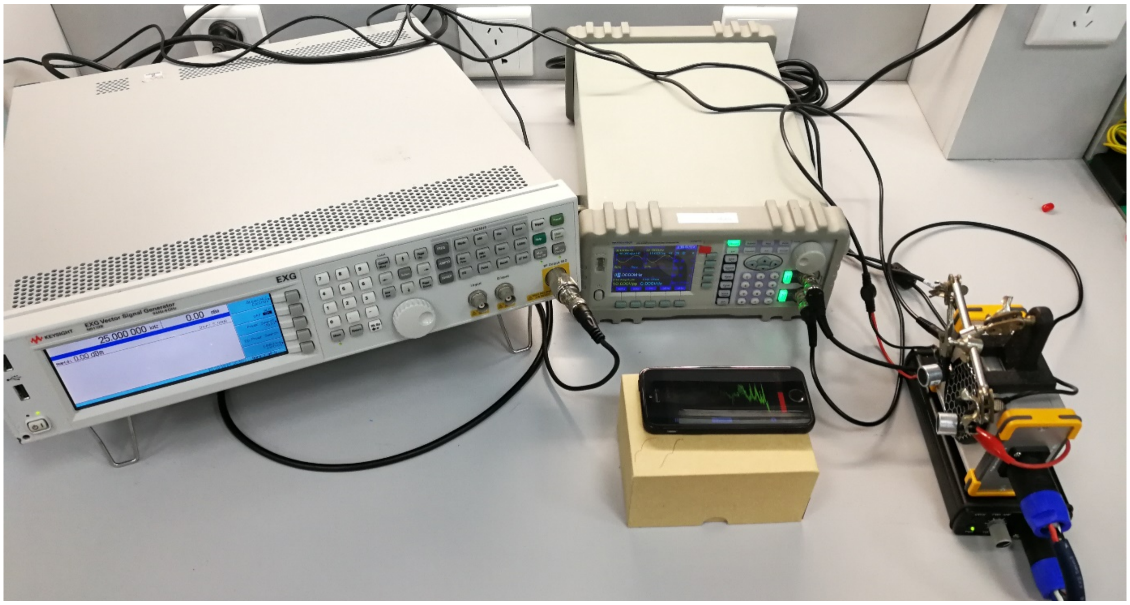

Figure 20: *Our benchtop equipment to carry about the proof of concept reproduction of tones heard in Cuba. Note, we would expect emitters to be smaller than a paperback book in practice, if not smaller. We use large equipment because of our general-purpose laboratory.*

4.1 Experimental Setup

Our experiments tested several different emitters and frequencies. We primarily use one wide-band ultrasonic speaker in combination with a multitude of fixed ultrasonic transducers to artificially create IMD. Our fixed transducers are centered at 25 kHz or 32 kHz depending on the experiment. Each fixed transducer has enough tolerance to support 180 Hz sidebands from AM modulation.

The wide-band ultrasonic speaker is a Vifa Speaker[3]. The 25 kHz and 32 kHz transducers are driven at 7 Vpp. At least two ultrasonic signals are necessary to reproduce our experiment. We use a basic function waveform generator for the fixed 25 kHz ultrasonic transducer, and a modulation-capable signal generator for the dynamic ultrasonic source. We used a Keysight N5172B EXG X-Series RF Vector Signal Generator for the AM modulation, but many function generators also have modulation capabilities. We validated the sound waves generated by our experiment with a measurement microphone with a frequency response of 4 Hz–100 kHz[4].

[3] `https://www.avisoft.com/usg/vifa.htm`

[4] National Instruments Inc., G.R.A.S. 46BE 1/4" CCP Free-field Standard Microphone Set, `http://www.ni.com/pdf/manuals/G.R.A.S._46BE.pdf`

We used SpectrumView and Ultrasonic Analyzer to produce the Fourier transforms of the sounds we measured with the microphone in our lab. Note, microphones can also add distortion and may differ from what a human would have heard in the room. In our Fourier transforms of the IMD recorded by our measurement microphone, we noticed very clear a 7 kHz tone and a few peaks that are 180 Hz, 360 Hz, 540 Hz, 720 Hz away from 7 kHz.

4.2 Experiment with Three Ultrasonic Tones

As shown in Figure 21, we generate ultrasound at three different frequencies (25 kHz, 32 kHz, 32.18 kHz) with three devices—two 32 kHz ultrasonic transducers (for 32 kHz and 32.18 kHz) and a wide-band ultrasonic speaker (for 25 kHz). A smartphone with recording and spectrum analysis applications listen to the ultrasonic sources, which are driven by two signal generators. The spectrum, the magnified spectrum around 7 kHz, and the waterfall plot appear in Figures 21–23. The experimental findings are consistent with results of simulation, except for the 3.5 kHz and 11 kHz signals, which might be caused by imperfections of the ultrasonic speakers. Notice that the logarithmic scale spectrum resembles what we observed in the simulations, which supports the nonlinearity model.

4.3 Experiment with Modulation

Our experiments tested a couple modulation schemes, including AM and FM. The FM (Figure 25) does not appear to match well with the AP News recording, but the AM modulation does (Figure 24).

4.4 Experiments with Video Demonstrations

The following videos show our experiments in action. The white appliance is the Keysight N5172B EXG X-Series RF Vector Signal Generator for the AM modulation, and it is connected to the silver ultrasonic speaker with orange rims on the right (the ultrasonic Vifa Speaker); the grey appliance is the signal generator that drives the fixed ultrasonic transducers. Note, in the picture above, we

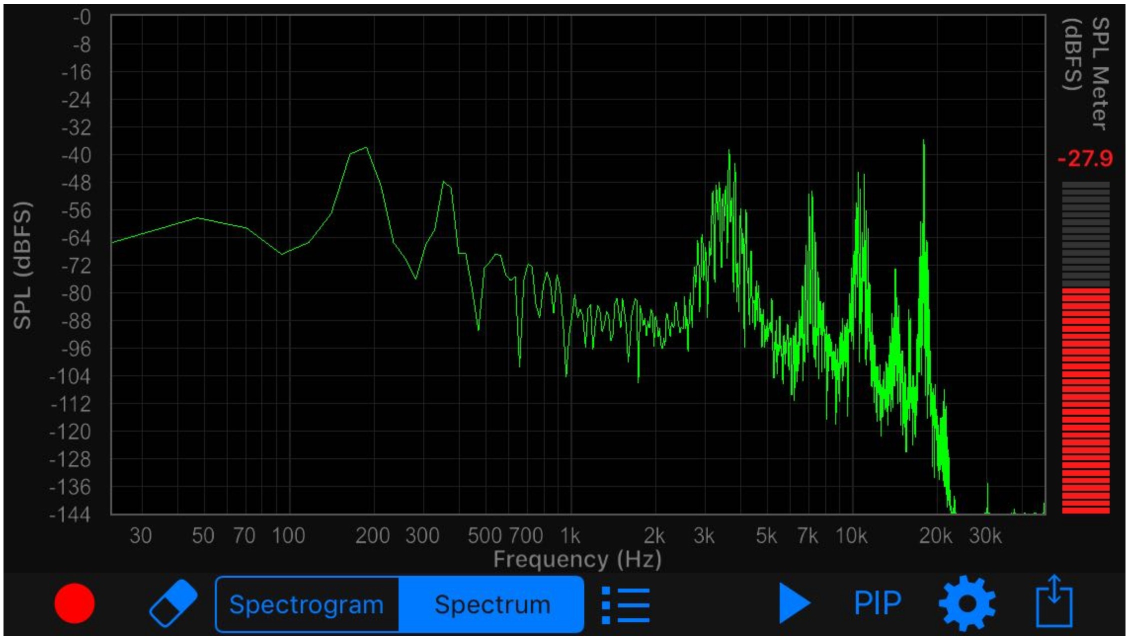

Figure 21: *Spectrum recorded during an IMD experiment playing three ultrasonic tones (25 kHz, 32 kHz, 32.18 kHz).*

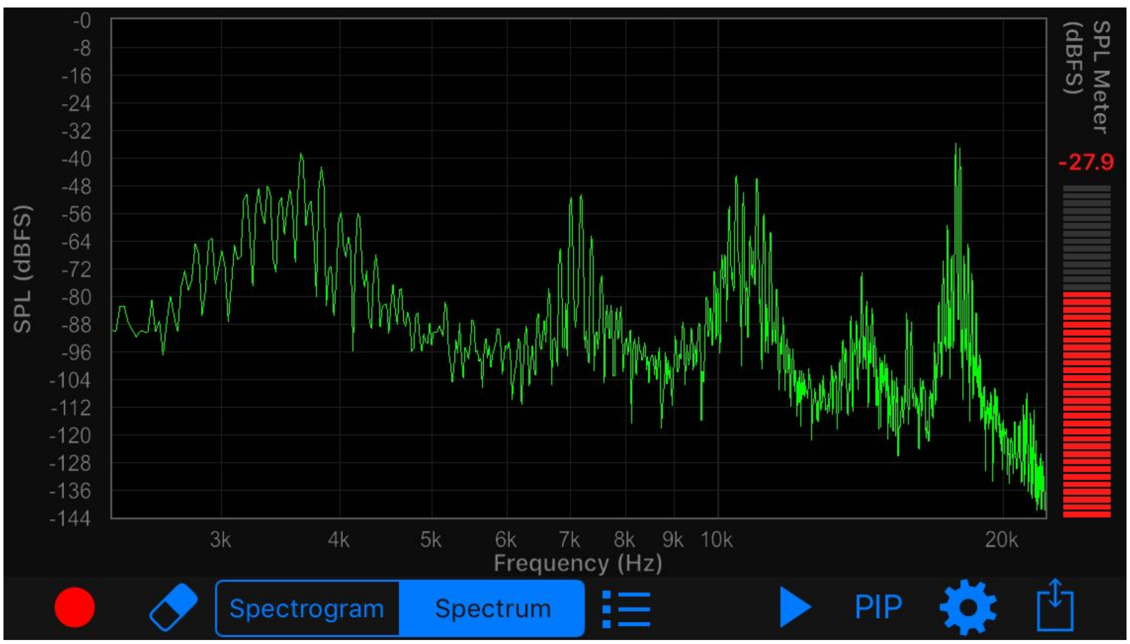

Figure 22: *Magnified spectrum of the signals near 7 kHz during an IMD experiment playing three ultrasonic tones (25 kHz, 32 kHz, 32.18 kHz)*

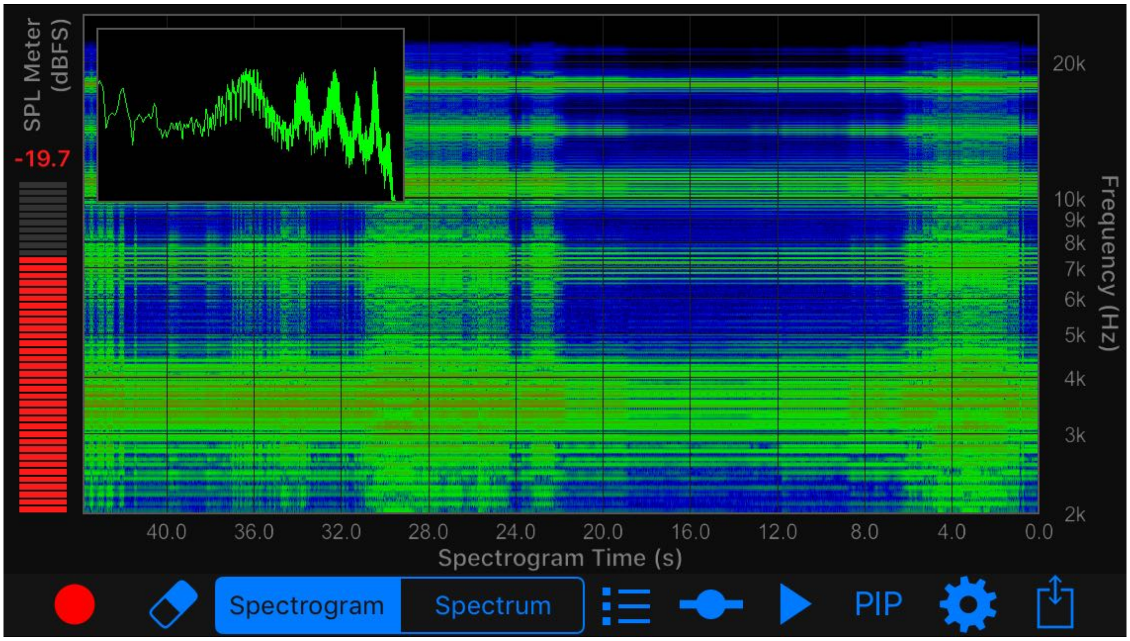

Figure 23: *Waterfall plot during an IMD experiment playing three ultrasonic tones (25 kHz, 32 kHz, 32.18 kHz).*

have two fixed ultrasonic transducers instead of one. The black smartphone in the middle serves as a spectrum analyzer.

Science of Synthesizing Audible Sounds from Ultrasonic Intermodulation Distortion. How can inaudible ultrasonic signals lead to audible byproducts? When multiple ultrasonic tones pass through a nonlinear medium such as air or a microphone, the result is intermodulation distortion[5]. In our experiment, we have two signals. One is a 180 Hz sine wave AM modulated over a 32 kHz ultrasonic carrier. The second is a simple 25 kHz ultrasonic sine wave. The smartphone displays the Fourier transform of repeated intermodulation distortion in the air and smartphone microphone circuitry. The 2nd order intermodulation distortion includes the difference between the two signals, which appears centered at 7 kHz and peppered with sidebands from the modulated 180 Hz. The higher order intermodulation distortion products create additional ripples in the spectrum at 7 kHz as well as several other frequencies. Matlab simulations predict the strong 7 kHz intermodulation distortion product, and we suspect the 4 kHz tones are the result of secondary intermodulation

[5] `https://youtu.be/wA2MZshrafk`

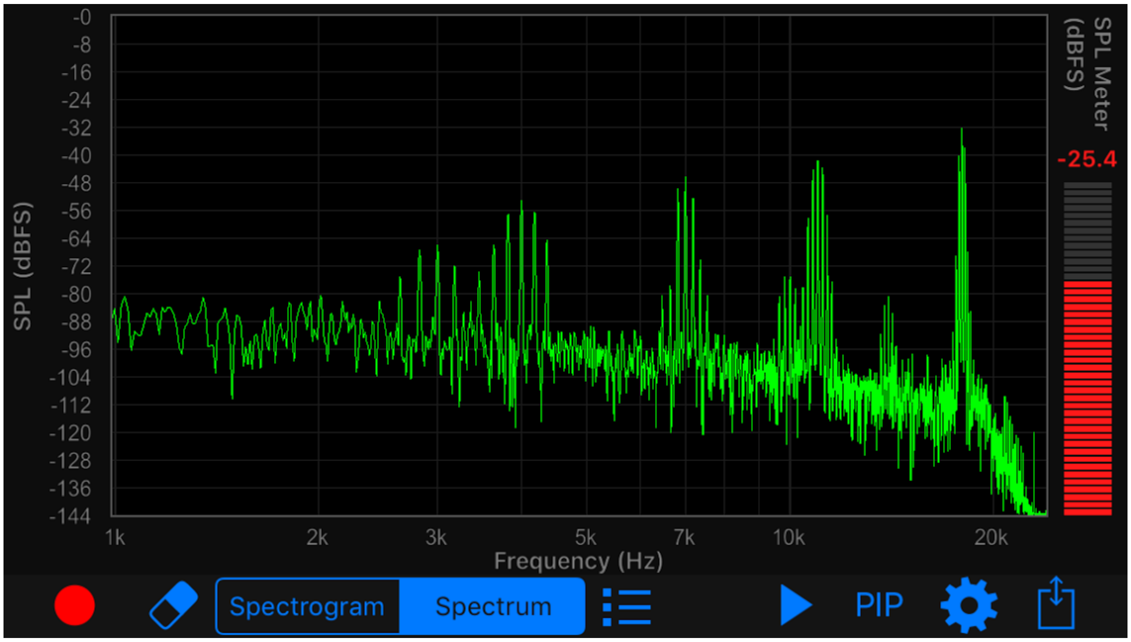

Figure 24: *Spectrum of sounds heard by a smart phone when playing 25 kHz and 180 Hz AM modulated on a 32 kHz carrer. The IMD spectrum resembles the ripples near 7 kHz in the AP news spectrum.*

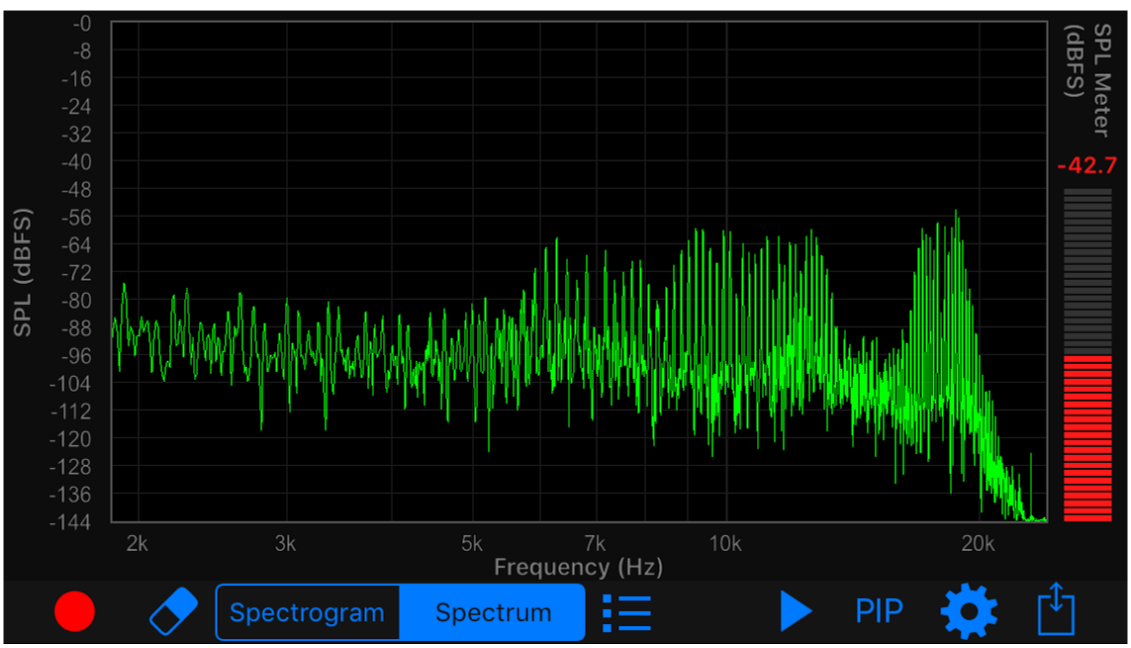

Figure 25: *Spectrum of sounds heard by a smart phone when playing 25 kHz and 180 Hz FM modulated on a 32 kHz carrer.*

distortion in the microphone.

At the beginning of the video, only the AM modulated signal (32 kHz carrier & 180 Hz sinusoidal baseband) is played through the ultrasonic Vifa Speaker, and the modulated ultrasounds cannot be heard or seen on the spectrum, which is out of the range of the spectrum plots. Once the signal generator starts to drive the fixed ultrasonic transducer to transmit a 25 kHz tone, we observe the IMD, as the spectrum analyzer shows, and can hear the high-pitched sounds.

Localized Audible Sounds Synthesized from Ultrasonic Intermodulation Distortion. Using two signal generators of low-intensity ultrasonic tones, we demonstrate synthesis of audible byproducts below 20 kHz[6]. Note, there are likely two cascading instances of intermodulation distortion: Once in the air that nearby humans can perceive, and a second time in the microphone of this smartphone. Thus, recordings of sound in Cuba are unlikely to match perfectly what humans perceived. In this experiment, our smartphone sensed a 4 kHz tone, but the student conducting the experiment could not hear a 4 kHz tone. Also note that the smartphone microphone has a frequency response that tapers off quickly after 20 kHz.

Absence of Audible Intermodulation Distortion from Single Ultrasonic Tone. Using two signal generators of ultrasonic tones, we demonstrate that the audible byproducts disappear when we disable one of the ultrasonic sources[7]. This is because at least two signals are necessary to elicit intermodulation distortion from a nonlinear medium such as air or microphone amplification circuitry.

Covertly Exfiltrating a Song with an Ultrasonic Carrier. This proof of concept shows two things: (1) how ultrasound can be used to covertly exfiltrate data (in this toy example, the audio from a memetastic song serves as a stand-in for eavedropping a conversation) and (2) how the covert channel becomes audibly overt when a second ultrasonic tone interferes. In this video[8], there are three microphones, two ultrasonic transmitters, and one audible speaker. One GRAS

[6] `https://youtu.be/ZTLjs4dbnEA`
[7] `https://youtu.be/o9jqwk83PSM`
[8] `https://youtu.be/w7_J1E5g8YQ`

ultrasonic microphone, one audible microphone on the iPhone plotting the FFT, and one audible microphone on the video recording device. The Vifa dynamic ultrasonic speaker emits the music modulated on an ultrasonic carrier. A small ultrasonic emitter sends out a single 32 kHz tone. A computer processing the ultrasonic signals from the G.R.A.S. microphone demodulates the signal and plays the resulting data, which is the song except when IMD causes corruption of the demodulation.

4.5 Discussion of Experiments

Our ultrasonic experiments create small, focused areas where one can perceive the audible sounds. Only where the ultrasonic beams cross do the sounds become apparent. Moving even a few inches from the beam can change the pitch, intensity, and sensation.

Our experiments were carried out in a lab at extremely low amplitudes to ensure the safety of the researchers.

The IMD products generated in our lab differ from the AP news recording in that we notice a set of tones at 4 kHz. IMD can happen between two signals and among more than two signals. To illustrate, we carried out experiments with multiple ultrasonic signals. While the student carrying out the experiment did hear the 7 kHz tone with his own ears, he could not hear the 4 kHz tone. We suspect that non linearities in our measurement microphone created this additional 4 kHz IMD. This observation is consistent with IMD we have found in other microphones from our previous research on ultrasonic cybersecurity [28].

5 Safety and Neurological Implications

There are two important questions that affect humans. What types of ultrasound can lead to hazardous situations or harm, and what are the neurological effects on humans?

Safety: Hazards, Hazardous Situations, Harm. We find little consensus on the risks of human exposure to air-borne ultrasound [21, 1]. Airborne ultrasonic waves on their own are not necessarily harmful, but it may become harmful at large intensity or when in direct contact exposure to a vibrating source. In direct contact exposure rather than by air, ultrasound can cause thermal injuries [1]. OSHA warns of potential harm from subharmonics of ultrasound [3], and appears to set a safety threshold in an abundance of caution. Health Canada [1] sets stricter safety requirements for the intensity of airborne ultrasound based on "plausible" risks of heating and cavitation as well as auditory and subjective effects. Canada sets a conservative 110 dB safety limit on emissions of airborne ultrasound.

According to the news [16], "The AP reported last month that some people experienced attacks or heard sounds that were narrowly confined to a room or parts of a room." Such a sensation is typical for ultrasound because ultrasound is more directional than audible sound and infrasound. Ultrasound can be focused on a certain area. Therefore, ultrasound would match the symtoms of discomfort.

Neurological Effects of Ultrasound. Researchers analyzed the effects of intense sounds on humans, but we find that the outcomes include large safety margins to make up for lack of consensus [4]. The *Handbook Human Vibration* [9] and an ISO standard [2] explore the physiological effects of low frequency vibrations and sounds. We have found little in the way of reproducible control trials for ultrasonic vibrations aside from folklore. Neurologists who examined the injured diplomats published their findings in JAMA [25], and suggest that the neurological damage is real. However, there are limitations to the retroactive study [18]—namely, causality is difficult to establish without a control trial or elimination of other null hypotheses. Our report does not itself contribute any new findings on neurological harm.

6 Alternative Explanations

While our results do not rule out other potential causes, the results do rule in the notion that ultrasound without harmful intent could have led to accidental harm to diplomats in Cuba.

We originally suspected subharmonics of ultrasound as the cause, but this hypothesis would not align well with the spectral analysis by the AP news. Rather than evenly spaced ripples in the frequency domain, we would expect to see frequencies at $1/n$ submultiples of the fundamental frequency for integers n if subharmonics were to blame.

180 Hz happens to be the high end of the fundamental frequencies of average male conversational voices. It may be coincidence that the tones are 180 Hz apart, but it could also indicate some kind of voice eavesdropping modulated over ultrasound and gone awry.

7 Related Work

The notion of using audible and inaudible sound to cause auditory and sensory illusions is not new. Our results build upon the following research.

Research from the music community used AM modulation on ultrasound to generate focused audible sound [19]. This research evolved into a company called Holosonics[9] with a product called Audio Spotlight for music, personalized sound, and museum exhibits, among other artistic applications. Companies such as the LRAD Corporation[10] produce products that deliver higher intensity sounds with military application to crowd control and long-distance hailing at sea. However, modern LRADs use audible parametric sound rather than ultrasound. Projects such as Soundlazer[11] allow the hobbyist engineer to play with ultrasonic generation of audible tones. Musicians have also used intermodulation distortion of *audible* tones to synthesize additional audible tones from nonlinearities of the inner ear [14]. Campbell even describes his realization of hearing synthesized combination tones (also known as intermodulation distortion) while listening to a movement in Sibelius's Symphony #1 [6].

Several researchers use ultrasound to fool sensors such as microphones. The BackDoor paper from Illinois [20] uses ultrasound and intermodulation distortion to jam eavesdropping microphones and watermark music played at concerts. A team from Korea uses both audible and ultrasonic tones to cause malfunctions in flight stability control of drones by acoustic interference at the resonant frequency of MEMS gyroscopes [22].

[9] https://www.holosonics.com/
[10] https://www.lradx.com/
[11] http://www.soundlazer.com/

In our past research, we use audible and ultrasonic tones to test the cybersecurity of computer systems. The DolphinAttack paper [28] uses ultrasound and intermodulation distortion to inject inaudible, fake voice commands into speech recognition systems including Siri, Google Now, Samsung S Voice, Huawei HiVoice, Cortana, Alexa, and the navigation system of an Audi automobile. Researchers from Princeton [23] investigate inaudible voice commands from ultrasound on Android phones and Amazon Echo. The Walnut paper [26] exploits nonlinear amplifiers, permissive analog filters, and signal aliasing to adulterate the output of MEMS accelerometers with sound waves at the resonant frequency of the sensor found in applications such as Fitbits, airbags, and smartphones. The sensors effectively serve as unintentional demodulators of the sound. Our upcoming Blue Note [5] paper analyzes the physics of why hard drives and operating systems get corrupted or spontaneously reboot when subjected to certain ultrasonic tones or by clicking on a link to a web site that plays maliciously crafted sound through the victim computer's mechanically coupled speakers.

We have urged more attention to the physics of cybersecurity [7], and the events in Cuba provide more evidence of the need to understand the causal relationships between physics and cybersecurity.

8 Unresolved Questions

Our report only rules in ultrasound and intermodulation distortion as a cause. It does not eliminate other hypotheses. In particular, several mysteries remain:

- How could ultrasound penetrate walls into homes and offices? Could an emitter be outside the premises or planted inside? Was it primarily air-borne, or did it originate as contact vibration?
- At what level of intensity could IMD products cause harm to humans? We know of no nontrivial lower bounds. Based on our reading of various safety documents, we believe most countries set conservative thresholds for airborne ultrasound from an abundance of caution and to compensate for uncertainty. While there are anecdotes and folklore for harm from

airborne ultrasound, we have found no primary sources that confirm this aside from stories about extremely intense sounds above 155 dB.

- What about standoff distance? Our report does not investigate distance. We do not have a facility to safely test high intensity ultrasound, but might look into it in the future if can borrow an airport runway.
- Could audible tones be a symptom or cause? Without a control study, it would be difficult to distinguish a cause from a symptom. It's possible that the audible sensations are byproducts from contact vibration or some other ultrasonic source.

9 Conclusion

Two inaudible ultrasonic signals mixing in a nonlinear medium could easily lead to an audible intermodulation distortion product. Although little is known about how audible sound waves can cause neurological damage rather than merely be correlated with neurological damage, the safety community has studied how certain audible sounds can cause pain and hearing damage. Thus, ultrasonic intermodulation distortion could produce harmful, audible byproducts. The safety warnings on audible frequencies and intensities would apply to these byproducts.

While our experiments do not eliminate the possibility of malicious intent to harm diplomats, our experiments do show that whoever caused the sensations may have had no intent for harm. The emitter source remains an open question, but could range from covert ultrasonic exfiltration of modulated data to ultrasonic jammers of eavesdropping devices or perhaps just ultrasonic pest repellents. It's also possible that someone was trying to covertly deliver data into a localized space using ultrasound to say, activate a sensor or other hidden device. Our experiments show that tones modulated on an ultrasonic carrier by one or more parties could have collided invisibly to produce audible byproducts. These audible byproducts can exist at frequencies known to cause annoyance and pain. Other theories include solid vibration (e.g., unwittingly standing on a covert transmitter) at ultrasonic frequencies for prolonged periods—leading to bodily harm. In such a case, audible intermodulation distortion could represent a harmless side effect rather than the cause of harm. Although our tests focus on frequencies rather than amplitudes or distances, we believe that high amplitude ultrasonic signals could easily produce high amplitude audible signals as unintentional byproducts capable of harm to hearing.

Acknowledgments

This research is supported by NSF CNS-1330142. The views and conclusions contained in this paper are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of NSF. We thank the U.S. Department of State and the members of the U.S. Foreign Service for their dedication to representing America abroad.

Change Log, Errata

1. March 1, 2018: Release 1.0. Technical feedback is welcome, and we will periodically update this report as new facts come to light.

References

[1] Guidelines for the Safe Use of Ultrasound: Part II Industrial and Commercial Applications. pages 1–43, 1991. Health Canada.

[2] Mechanical Vibration and Shock: Evaluation of human exposure to whole-body vibration. pages 1–7, 1997. ISO Standard 2631-1.

[3] OSHA Technical Manual: Noise: Appendix C Ultrasound. pages 1–3, 2013.

[4] C H Allen, I Rudnick, and H Frings. Some Biological Effects of Intense High Frequency Airborne Sound. *The Journal of the Acoustical Society of America*, 20(2):221–221, March 1948.

[5] Connor Bolton, Sara Rampazzi, Chaohao Li, Andrew Kwong, Wenyuan Xu, and Kevin Fu. Blue Note: How Intentional Acoustic Interference Damages Availability and Integrity in Hard Disk Drives and Operating Systems. In *Proceedings of the 39th Annual IEEE Symposium on Security and Privacy*, May 2018.

[6] Murray Campbell and Clive Greated. *The Musician's Guide to Acoustics*. OUP Oxford, April 1994.

[7] Kevin Fu and Wenyuan Xu. Inside Risks: Risks of Trusting the Physics of Sensors. *Communications of the ACM*, 61(2):20–23, February 2018.

[8] Tim Golden and Sebastian Rotella. The Sound and the Fury: Inside the Mystery of the Havana Embassy. *Pro Publica*, February 2018.

[9] M J Griffin. *Handbook of Human Vibration*. Academic Press, December 2012.

[10] Gardiner Harris. Tillerson Suggests Cuba Could Have Stopped 'Targeted Attacks' on U.S. Diplomats. *The New York Times*, December 2017.

[11] Gardiner Harris. U.S. to Open Formal Inquiry on Americans Sickened in Cuba. *The New York Times*, January 2018.

[12] Gardiner Harris, Julie Hirschfeld Davis, and Ernesto Londoño. U.S. Expels 15 Cuban Diplomats, in Latest Sign Détente May Be Ending. *The New York Times*, October 2017.

[13] Gardiner Harris and Adam Goldman. Illnesses at U.S. Embassy in Havana Prompt Evacuation of More Diplomats. *The New York Times*, September 2017.

[14] Gary S Kendall, Christopher Haworth, and Rodrigo F Cádiz. Sound Synthesis with Auditory Distortion Products. *Computer Music Journal*, 38(4):5–23, 2014.

[15] Gina Kolata. Diplomats in Cuba Suffered Brain Injuries. Experts Still Don't Know Why. *The New York Times*, February 2018.

[16] Josh Lederman and Michael Weissenstein. Dangerous sound? What Americans heard in Cuba attacks. *AP News*, October 2017.

[17] Alexis C Madrigal. The Case of the Sick Americans in Cuba Gets Stranger. *The Atlantic*, February 2018.

[18] Christopher C Muth and Steven L Lewis. Neurological Symptoms Among US Diplomats in Cuba. *JAMA*, pages 1–3, February 2018.

[19] F Joseph Pompei. The Use of Airborne Ultrasonics for Generating Audible Sound Beams. *Journal of Audio Engineering Society*, 47(9), September 1999.

[20] Nirupam Roy, Haitham Hassanieh, and Romit Roy Choudhury. BackDoor: Making Microphones Hear Inaudible Sounds. In *ACM MobiSys '17*, pages 2–14. ACM Press, 2017.

[21] Bozena Smagowska. Effects of Ultrasonic Noise on the Human Body—A Bibliographic Review. *International Journal of Occupational Safety and Ergonomics JOSE*, 19(2), 2013.

[22] Yunmok Son, Hocheol Shin, Dongkwan Kim, Youngseok Park, Juhwan Noh, Kibum Choi, Jungwoo Choi, and Yongdae Kim. Rocking Drones with Intentional Sound Noise on Gyroscopic Sensors. In *24th USENIX Security Symposium (USENIX Security)*, pages 881–896, 2015.

[23] Liwei Song and Prateek Mittal. Inaudible voice commands. *CoRR*, abs/1708.07238, 2017.

[24] Richard Stone. Stressful conditions, not ‘sonic weapon,’ sickened U.S diplomats, Cuba panel asserts. *Science*, December 2017.

[25] Randel L Swanson II, Stephen Hampton, Judith Green-McKenzie, Ramon Diaz-Arrastia, M Sean Grady, Ragini Verma, Rosette Biester, Diana Duda, Ronald L Wolf, and Douglas H Smith. Neurological Manifestations Among US Government Personnel Reporting Directional Audible and Sensory Phenomena in Havana, Cuba. *JAMA*, pages 1–9, February 2018.

[26] Timothy Trippel, Ofir Weisse, Wenyuan Xu, Peter Honeyman, and Kevin Fu. WALNUT: Waging doubt on the integrity of MEMS accelerometers with acoustic injection attacks. *IEEE European Symposium on Security & Privacy*, pages 3–18, 2017.

[27] Emily Waltz. Was a Sonic Weapon Deployed in Cuba. *IEEE Spectrum*, December 2017.

[28] Guoming Zhang, Chen Yan, Xiaoyu Ji, Tianchen Zhang, Taimin Zhang, and Wenyuan Xu. DolphinAttack: Inaudible Voice Commands. *ACM CCS*, pages 103–117, 2017.