

# Gyrophone: Recognizing Speech From Gyroscope Signals

Yan Michalevsky Dan Boneh

*Computer Science Department  
Stanford University*

Gabi Nakibly

*National Research & Simulation Center  
Rafael Ltd.*

## Abstract

We show that the MEMS gyroscopes found on modern smart phones are sufficiently sensitive to measure acoustic signals in the vicinity of the phone. The resulting signals contain only very low-frequency information (<200Hz). Nevertheless we show, using signal processing and machine learning, that this information is sufficient to identify speaker information and even parse speech. Since iOS and Android require no special permissions to access the gyro, our results show that apps and active web content that cannot access the microphone can nevertheless eavesdrop on speech in the vicinity of the phone.

## 1 Introduction

Modern smartphones and mobile devices have many sensors that enable rich user experience. Being generally put to good use, they can sometimes unintentionally expose information the user does not want to share. While the privacy risks associated with some sensors like a microphone (eavesdropping), camera or GPS (tracking) are obvious and well understood, some of the risks remained under the radar for users and application developers. In particular, access to motion sensors such as gyroscope and accelerometer is unmitigated by mobile operating systems. Namely, every application installed on a phone and every web page browsed over it can measure and record these sensors without the user being aware of it.

Recently, a few research works pointed out unintended information leaks using motion sensors. In Ref. [34] the authors suggest a method for user identification from gait patterns obtained from a mobile device's accelerometers. The feasibility of keystroke inference from nearby keyboards using accelerometers has been shown in [35]. In [21], the authors demonstrate the possibility of keystroke inference on a mobile device using accelerometers and mention the potential of using gyroscope measurements as well, while another study [19] points to the benefits of exploiting the gyroscope.

All of the above work focused on exploitation of motion events obtained from the sensors, utilizing the expected kinetic response of accelerometers and gyroscopes. In this paper we reveal a new way to extract information from gyroscope measurements. We show that

gyroscopes are sufficiently sensitive to measure acoustic vibrations. This leads to the possibility of recovering speech from gyroscope readings, namely using the gyroscope as a crude microphone. We show that the sampling rate of the gyroscope is up to 200 Hz which covers some of the audible range. This raises the possibility of eavesdropping on speech in the vicinity of a phone without access to the real microphone.

As the sampling rate of the gyroscope is limited, one cannot fully reconstruct a comprehensible speech from measurements of a single gyroscope. Therefore, we resort to automatic speech recognition. We extract features from the gyroscope measurements using various signal processing methods and train machine learning algorithms for recognition. We achieve about 50% success rate for speaker identification from a set of 10 speakers. We also show that while limiting ourselves to a small vocabulary consisting solely of digit pronunciations ("one", "two", "three", ...) and achieve speech recognition success rate of 65% for the speaker dependent case and up to 26% recognition rate for the speaker independent case. This capability allows an attacker to substantially leak information about numbers spoken over or next to a phone (i.e. credit card numbers, social security numbers and the like).

We also consider the setting of a conference room where two or more people are carrying smartphones or tablets. This setting allows an attacker to gain simultaneous measurements of speech from several gyroscopes. We show that by combining the signals from two or more phones we can increase the effective sampling rate of the acoustic signal while achieving better speech recognition rates. In our experiments we achieved 77% successful recognition rate in the speaker dependent case based on the digits vocabulary.

The paper structure is as follows: in Section 2 we provide a brief description of how a MEMS gyroscope works and present initial investigation of its properties as a microphone. In Section 3 we discuss speech analysis and describe our algorithms for speaker and speech recognition. In Section 4 we suggest a method for audio signal recovery using samples from multiple devices. In Section 5 we discuss more directions for exploitation of gyroscopes' acoustic sensitivity. Finally, in Section 6 we discuss mitigation measures of this unexpected threat. In

particular, we argue that restricting the sampling rate is an effective and backwards compatible solution.

## 2 Gyroscope as a microphone

In this section we explain how MEMS gyroscopes operate and present an initial investigation of their susceptibility to acoustic signals.

### 2.1 How does a MEMS gyroscope work?

Standard-size (non-MEMS) gyroscopes are usually composed of a spinning wheel on an axle that is free to assume any orientation. Based on the principles of angular momentum the wheel resists to changes in orientation, thereby allowing to measure those changes. Nonetheless, all MEMS gyros take advantage of a different physical phenomenon  $\pm$  the Coriolis force. It is a fictitious force (d'Alembert force) that appears to act on an object while viewing it from a rotating reference frame (much like the centrifugal force). The Coriolis force acts in a direction perpendicular to the rotation axis of the reference frame and to the velocity of the viewed object. The Coriolis force is calculated by  $F = 2m\mathbf{v} \times \boldsymbol{\omega}$  where  $m$  and  $\mathbf{v}$  denote the object's mass and velocity, respectively, and  $\boldsymbol{\omega}$  denotes the angular rate of the reference frame.

Generally speaking, MEMS gyros measure their angular rate ( $\boldsymbol{\omega}$ ) by sensing the magnitude of the Coriolis force acting on a moving proof mass within the gyro. Usually the moving proof mass constantly vibrates within the gyro. Its vibration frequency is also called the resonance frequency of the gyro. The Coriolis force is sensed by measuring its resulting vibration, which is orthogonal to the primary vibration movement. Some gyroscope designs use a single mass to measure the angular rate of different axes, while others use multiple masses. Such a general design is commonly called *vibrating structure gyroscope*.

There are two primary vendors of MEMS gyroscopes for mobile devices: STMicroelectronics [15] and InvenSense [7]. According to a recent survey [18] STMicroelectronics dominates with 80% market share. Tear-down analyses show that this vendor's gyros can be found in Apple's iPhones and iPads [17, 8] and also in the latest generations of Samsung's Galaxy-line phones [5, 6]. The second vendor, InvenSense, has the remaining 20% market share [18]. InvenSense gyros can be found in Google's latest generations of Nexus-line phones and tablets [14, 13] as well as in Galaxy-line tablets [4, 3]. These two vendors' gyroscopes have different mechanical designs, but are both noticeably influenced by acoustic noise.

#### 2.1.1 STMicroelectronics

The design of STMicroelectronics 3-axis gyros is based on a single driving (vibrating) mass (shown in Figure 1). The driving mass consists of 4 parts  $M_1$ ,  $M_2$ ,  $M_3$  and  $M_4$  (Figure 1(b)). They move inward and outward simultaneously at a certain frequency<sup>1</sup> in the horizontal plane. As shown in Figure 1(b), when an angular rate is applied on the Z-axis, due to the Coriolis effect,  $M_2$  and  $M_4$  will move in the same horizontal plane in opposite directions as shown by the red and yellow arrows. When an angular rate is applied on the X-axis, then  $M_1$  and  $M_3$  will move in opposite directions up and down out of the plane due to the Coriolis effect. When an angular rate is applied to the Y-axis, then  $M_2$  and  $M_4$  will move in opposite directions up and down out of the plane. The movement of the driving mass causes a capacitance change relative to stationary plates surrounding it. This change is sensed and translated into the measurement signal.

#### 2.1.2 InvenSense

InvenSense's gyro design is based on the three separate driving (vibrating) masses<sup>2</sup>; each senses angular rate at a different axis (shown in Figure 2(a)). Each mass is a coupled dual-mass that move in opposite directions. The masses that sense the X and Y axes are driven out-of-plane (see Figure 2(b)), while the Z-axis mass is driven in-plane. As in the STMicroelectronics design the movement due to the Coriolis force is measured by capacitance changes.

## 2.2 Acoustic Effects

It is a well known fact in the MEMS community that MEMS gyros are susceptible to acoustic noise which degrades their accuracy [22, 24, 25]. An acoustic signal affects the gyroscope measurement by making the driving mass vibrate in the sensing axis (the axis which senses the Coriolis force). The acoustic signal can be transferred to the driving mass in one of two ways. First, it may induce mechanical vibrations to the gyros package. Additionally, the acoustic signal can travel through the gyroscope packaging and directly affect the driving mass in case it is suspended in air. The acoustic noise has the most substantial effect when it is near the resonance frequency of the vibrating mass. Such effects in some cases can render the gyro's measurements useless or even saturated. Therefore to reduce the noise effects vendors manufacture gyros with a high resonance frequency (above

<sup>1</sup>It is indicated in [1] that STMicroelectronics uses a driving frequency of over 20 KHz.

<sup>2</sup>According to [43] the driving frequency of the masses is between 25 KHz and 30 KHz.

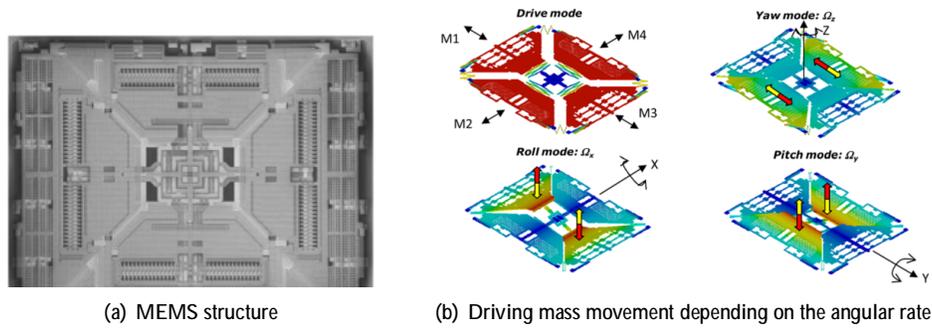


Figure 1: STMicroelectronics 3-axis gyro design (Taken from [16]. Figure copyright of STMicroelectronics. Used with permission.)

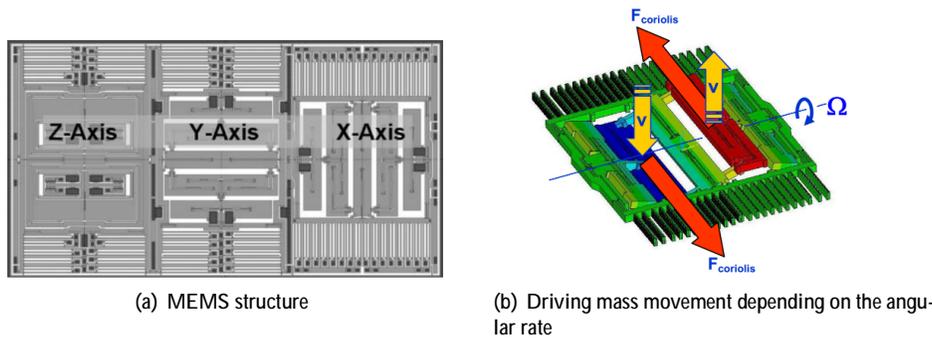


Figure 2: InvenSense 3-axis gyro design (Taken from [43]. Figure copyright of InvenSense. Used with permission.)

20 KHz) where acoustic signals are minimal. Nonetheless, in our experiments we found that acoustic signals at frequencies much lower than the resonance frequency still have a measurable effect on a gyro’s measurements, allowing one to reconstruct the acoustic signal.

### 2.3 Characteristics of a gyro as a microphone

Due to the gyro’s acoustic susceptibility one can treat gyroscope readings as if they were audio samples coming from a microphone. Note that the frequency of an audible signal is higher than 20 Hz, while in common cases the frequency of change of mobile device’s angular velocity is lower than 20 cycles per second. Therefore, one can high-pass-filter the gyroscope readings in order to retain only the effects of an audio signal even if the mobile device is moving about. Nonetheless, it should be noted that this filtering may result in some loss of acoustic information since some aliased frequencies may be filtered out (see Section 2.3.2). In the following we explore the gyroscope characteristics from a standpoint of an acoustic sensor, i.e. a microphone. In this section we exemplify these characteristics by experimenting with Galaxy S III which has an STMicroelectronics gyro [6].

#### 2.3.1 Sampling

Sampling resolution is measured by the number of bits per sample. More bits allow us to sample the signal more accurately at any given time. All the latest generations of gyroscopes have a sample resolution of 16 bits [9, 12]. This is comparable to a microphone’s sampling resolution used in most audio applications.

Sampling frequency is the rate at which a signal is sampled. According to the Nyquist sampling theorem a sampling frequency  $f$  enables us to reconstruct signals at frequencies of up to  $f/2$ . Hence, a higher sampling frequency allows us to more accurately reconstruct the audio signal. In most mobile devices and operating systems an application is able to sample the output of a microphone at up to 44.1 KHz. A telephone system (POTS) samples an audio signal at 8000 Hz. However, STMicroelectronics’ gyroscope hardware supports sampling frequencies of up to 800 Hz [9], while InvenSense gyros’ hardware support sampling frequency up to 8000 Hz [12]. Moreover, all mobile operating systems bound the sampling frequency even further  $\pm$  up to 200 Hz  $\pm$  to limit power consumption. On top of that, it appears that some browser toolkits limit the sampling frequency even further. Table 1 summarizes the results of our experi-

		Sampling Freq. [Hz]
Android 4.4	application	200
	Chrome	25
	Firefox	200
	Opera	20
iOS 7	application	100 [2]
	Safari	20
	Chrome	20

Table 1: Maximum sampling frequencies on different platforms

ments measuring the maximum sampling frequencies allowed in the latest versions of Android and iOS both for application and for web application running on common browsers. The code we used to sample the gyro via a web page can be found in Appendix B. The results indicate that a Gecko based browser does not limit the sampling frequency beyond the limit imposed by the operating system, while WebKit and Blink based browsers does impose stricter limits on it.

### 2.3.2 Aliasing

As noted above, the sampling frequency of a gyro is uniform and can be at most 200 Hz. This allows us to directly sense audio signals of up to 100 Hz. Aliasing is a phenomenon where for a sinusoid of frequency  $f$ , sampled with frequency  $f_s$ , the resulting samples are indistinguishable from those of another sinusoid of frequency  $jf - N \cdot f_s$ , for any integer  $N$ . The values corresponding to  $N \neq 0$  are called images or aliases of frequency  $f$ . An undesirable phenomenon in general, here aliasing allows us to sense audio signals having frequencies which are higher than 100 Hz, thereby extracting more information from the gyroscope readings. This is illustrated in Figure 3.

Using the gyro, we recorded a single 280 Hz tone. Figure 3(a) depicts the recorded signal in the frequency domain (x-axis) over time (y-axis). A lighter shade in the spectrogram indicates a stronger signal at the corresponding frequency and time values. It can be clearly seen that there is a strong signal sensed at frequency 80 Hz starting around 1.5 sec. This is an alias of the 280 Hz-tone. Note that the aliased tone is indistinguishable from an actual tone at the aliased frequency. Figure 3(b) depicts a recording of multiple short tones between 130 Hz and 200 Hz. Again, a strong signal can be seen at the aliased frequencies corresponding to 130 - 170 Hz<sup>3</sup>. We also observe some weaker aliases that do not correspond to the base frequencies of the recorded tones, and per-

<sup>3</sup>We do not see the aliases corresponding to 180 - 200 Hz, which might be masked by the noise at low frequencies, i.e., under 20 Hz.

haps correspond to their harmonics. Figure 3(c) depicts the recording of a chirp in the range of 420 - 480 Hz. The aliased chirp is detectable in the range of 20 - 80 Hz; however it is a rather weak signal.

### 2.3.3 Self noise

The self noise characteristic of a microphone indicates what is the most quiet sound, in decibels, a microphone can pick up, i.e. the sound that is just over its self noise. To measure the gyroscope's self noise we played 80 Hz tones for 10 seconds at different volumes while measuring it using a decibel meter. Each tone was recorded by the Galaxy S III gyroscope. While analyzing the gyro recordings we realized that the gyro readings have a noticeable increase in amplitude when playing tones with volume of 75 dB or higher which is comparable to the volume of a loud conversation. Moreover, a FFT plot of the gyroscope recordings gives a noticeable peak at the tone's frequency when playing tone with a volume as low as 57 dB which is below the sound level of a normal conversation. These findings indicate that a gyro can pick up audio signals which are lower than 100 Hz during most conversations made over or next to the phone. To test the self noise of the gyro for aliased tones we played 150 Hz and 250 Hz tones. The lowest level of sound the gyro picked up was 67 dB and 77 dB, respectively. These are much higher values that are comparable to a loud conversation.

### 2.3.4 Directionality

We now measure how the angle at which the audio signal hits the phone affects the gyro. For this experiment we played an 80 Hz tone at the same volume three times. The tone was recorded at each time by the Galaxy S III gyro while the phone rested at a different orientation allowing the signal to hit it parallel to one of its three axes (see Figure 4). The gyroscope senses in three axes, hence for each measurement the gyro actually outputs three readings  $\pm$  one per axis. As we show next this property benefits the gyro's ability to pick up audio signals from every direction. For each recording we calculated the FFT magnitude at 80 Hz. Table 2 summarizes the results.

It is obvious from the table that for each direction the audio hit the gyro, there is at least one axis whose readings are dominant by an order of magnitude compared to the rest. This can be explained by STMicroelectronics gyroscope design as depicted in Figure 1<sup>4</sup>. When the signal travels in parallel to the phone's  $x$  or  $y$  axes, the sound pressure vibrates mostly masses laid along the respective axis, i.e.  $M_2$  and  $M_4$  for  $x$  axis and  $M_1$  and  $M_3$

<sup>4</sup>This is the design of the gyro built into Galaxy S III.

