

# Piezoelectric MEMS Speaker Design

Jason Bissias  
12 May 2025

## I. INTRODUCTION

Modern consumer headphones are seeing an increasing trend of convenience, primarily by decreasing their size. One solution seen in modern production headphones is to use micro-electromechanical system (MEMS) speakers integrated onto printed circuit boards. By nature of what MEMS speakers entail, these systems are significantly smaller and thinner than traditional membrane speakers. Furthermore, and more importantly for our purposes, they are cheaper as they can be constructed using modern lithographic means, allowing for batch processing, whereas traditional speakers require large coils and permanent magnets that are comparatively difficult to assemble. [1]

There are many different types of MEMS speakers that have been developed, including transducer mechanisms based on piezoelectric, electrodynamic, electrostatic, and thermoacoustic principles, as seen in Fig. 1. The piezoelectric transducer mechanism is the most widely used and developed among these systems, as it sees usage in other popular consumer systems such as ink-jet printer heads [1]. Thus, this paper will focus on these systems.

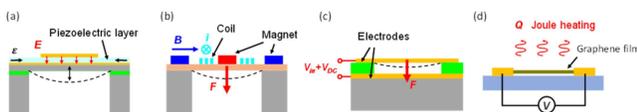


Figure 1: Schematics of various transducer mechanisms. (a) Piezoelectric (b) Electrodynamic (c) Electrostatic (d) Thermoacoustic [1]

Through the miniaturization and bulk manufacturing capabilities of MEMS speakers, entirely integrated machines can be swiftly made. Typical and high-end consumer products may not opt for this method, as they often desire some amount of uniqueness in sound signature and style, depending on the specific user, so the more modular system currently used by audio device manufacturers may prefer their old ways. However, for commercial settings, these devices may see significant cost savings. For example, in any profession that requires a headset (notably in customer service), it may be beneficial to cheaply mass produce such devices.

This paper will describe a fully integrated wireless audio device for economical mass-production, focused on the audio output portions of the device. It is reasonable to expect such product to include a microphone as input, however for the purposes of brevity this consideration will be avoided.

## II. SYSTEM ARCHITECTURE

The architecture of the entire (considered) system comprises of the speakers themselves, a shift register, a Bluetooth receiver circuit, and an antenna. Conventional audio output systems use a digital-to-analog converter (DAC) along with an amplifier to properly render digitally stored and transmitted audio. However, the presented speaker array system actually can directly work with digital signals, this process being called digital sound reconstruction (DSR) [2]. This is accomplished by setting each microspeaker in a

group to handle one of the bits in each sample, at which point the group will create a “click” to render the bit. For example, for 4-bit audio, one would need 15 microspeakers for the 16 possible values, as diagrammed in Fig. 2. This is typically scaled to 16-bits with 255 microspeakers [2, 3].

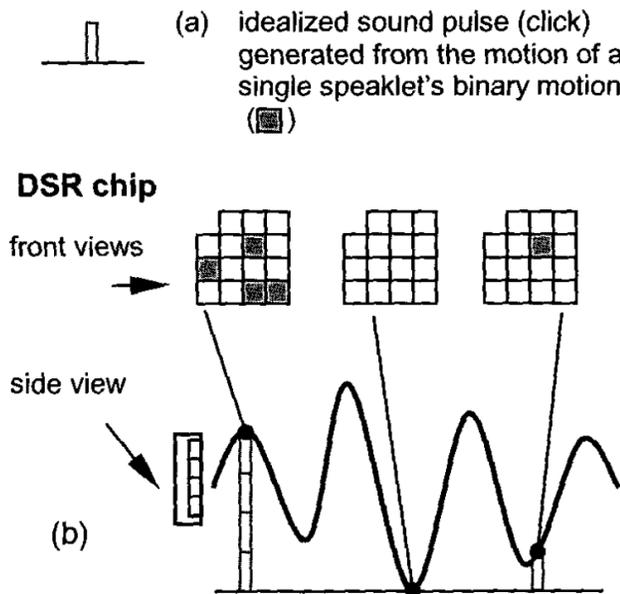


Figure 2: DSR with a 15 microspeakers (4-bit) chip. (a) conceptually shows a a single click. (b) shows how a culmination of microspeakers come together at different times to reconstruct a sound wave. [2]

The first realized issue with this system, however, is that though in theory, compared to a single microspeakers, a higher sound pressure level (SPL) and more linear (less distorted) sound reconstruction is possible, there practically is less capability in sub-1 kHz audio reconstruction [4], unless the microspeakers are ideally designed for this use. One remedy may be in a ceramic PZT, as opposed to the typical sol-gel PZT [5]. This is because a ceramic PZT has a significantly higher piezoelectric constant compared to thin-film PZTs, thus allowing for greater SPL with the same driving voltage and architecture. Recent advances in material processing has allowed for this previously hard-to-integrate material feasible.

Combining these two concepts, a capable yet cheap speaker system is very possible. The MEMS speakers are able to be made directly on silicon and directly integrated to the digital signal source, in this case a shift register connected to the Bluetooth receiver circuit with an integrated antenna as its input. A typical design for Bluetooth antennas is the Printed Inverted F Antenna (PIFA) [6]. These antennas can be printed onto silicon substrates with a benzocyclobutane (BCB) dielectric layer between the silicon and the antenna to reduce losses caused by the low-resistivity of silicon [7]. Taking into account the surrounding structure of the antenna (since the ground plane is likely comparable to the size of the antenna itself), it can be tuned to the 2.45 GHz center frequency of Bluetooth operation [6, 7]. Fig. 3 shows the basic layout of this antenna. The antenna connects with the rest of the system via the CPW feed.

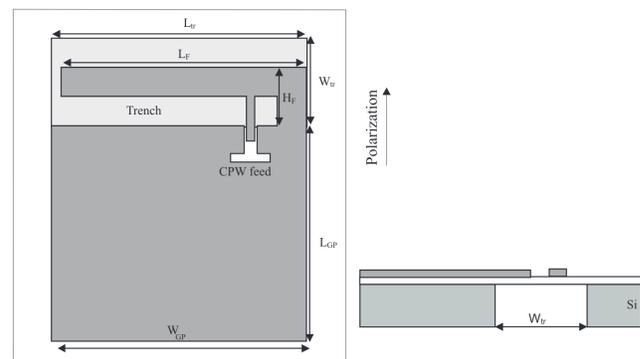


Figure 3: PIFA on Si Substrate [8]

The shift-register is used to parallelize the serial Bluetooth signal for each microspeakers cluster. The integration of the shift-register, even at 16-bit resolution, is relatively trivial and uses conventional & established MOSFET architectures. Similarly, the Bluetooth receiver is also trivial, and can use well-established fully integrated technology [8]. Even in its most basic form, 1 Mb/s of data reception is easily feasible, which is more than plenty for this use [9]. Fig. 4 shows a block diagram of this integrated receiver,

with the shift-register being the “CODEC Interface.” The “Host Interface” could simply be any external buttons the product may have, and the clock would be built on the chip.

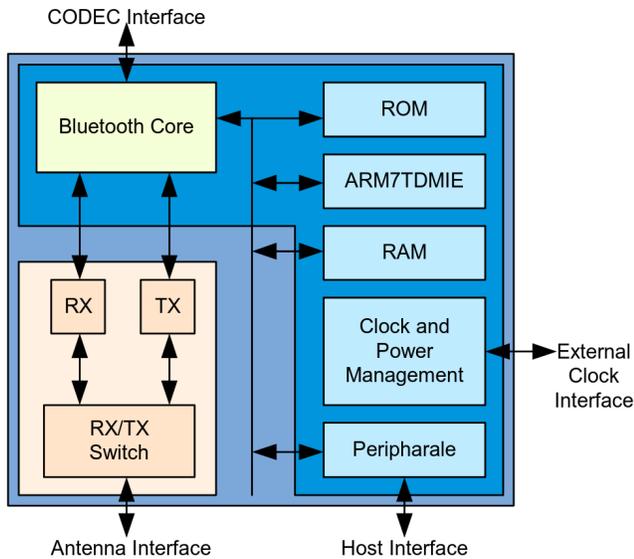


Figure 4: Block diagram of Bluetooth receiver circuit [9].

### III. PACKAGING APPROACH

The smallest reported microspeakers have a radius of 400  $\mu\text{m}$ , which produces an 255-speaker array with a footprint of about  $7.7 \times 7.7 \text{ mm}$  [2]. Meanwhile, an antenna tuned to 2.45 GHz is going to be about 20-25 mm long [6], and the Bluetooth receiver circuit may be about 90  $\text{mm}^2$  [9]. The only necessary physical connection the device needs is to the battery, which for this application the typical battery size is comparable to the footprint of the antenna structure [10]. The chip therefore may be connected directly to the battery using wirebonding from the necessary contact pads. This method is chosen as wirebonding is a very mature process, and likely the most common of the interconnection methods [11]. By extension, it is one of the cheapest methods. Lastly, since these interconnections do not involve high speed or high power, wirebonding will more than suffice.

With these specifications in mind, the headphones can therefore be simple on-ear systems, which may measure just a few centimeters in every direction. They will be able to be slim and light, partially thanks to the low profile of MEMS speakers and partially thanks to the low power consumption (therefore small battery) the system as a whole consumes. The single largest component of the product will be in the battery, which does not need to be too large since power consumption is almost certainly below a Watt [1, 8] for the entire system. A rudimentary to-scale block diagram of this product is shown in Fig. 5. In total, all of the components will simply sit within the plastic cup of the headphone, with the bottom of the wafer facing the grille, allowing for the speakers to output sound clearly and for the antenna to be facing away from the user’s head.

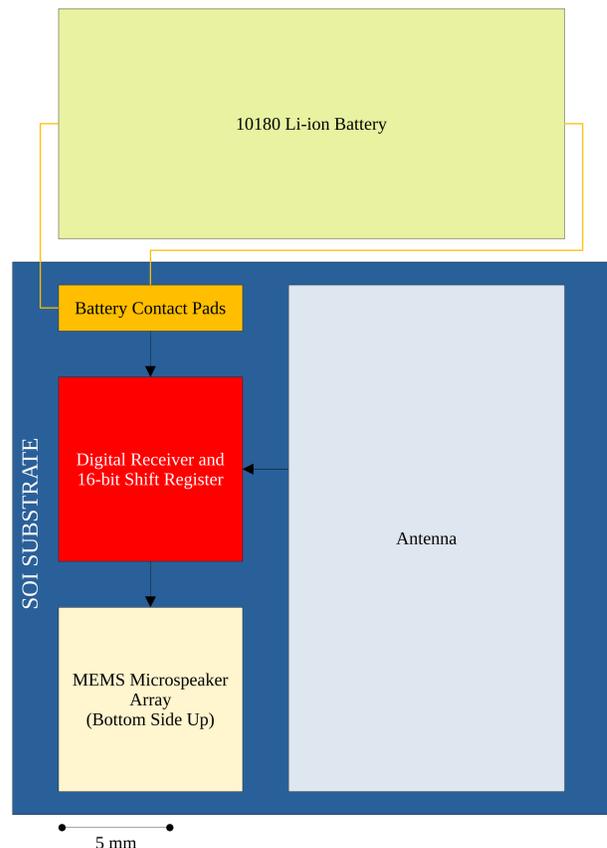


Figure 5: To-scale block diagram of proposed MEMS speaker system.

#### IV. MECHANICAL/ENVIRONMENTAL CONSIDERATIONS

There may be some concern over the reliability of a relatively unprotected MEMS device, since by principle the only barrier between the microspeakers and the outside world is a mesh grille, and the antenna needs to be as unobstructed as possible. The digital microelectronics can be reasonably protected via a passivation layer at the surface, such as SiN, however the microspeakers and antenna still need to be exposed. At this stage, there is no serious concern, as the ceramic PZT membranes are expected to be relatively durable and the antenna is decently protected from the fact that it is printed into the substrate.

Another concern may be that of heat generated. However, since the device is rather low-power, not much heat is likely to be generated. Additionally, the lack of solid barrier with the outside world may be a benefit in this case, allowing for heat to be easily dissipated.

#### V. MANUFACTURING FLOW

Combining the steps for each individual system from [2], [3], [5], and [7], Fig. 6 shows an abbreviated process flow for the entire integrated chip. The first steps are to create the MEMS speakers, as this process involved attaching a ceramic PZT substrate onto the Si substrate. The majority of the ceramic PZT will be etched, leaving behind only the speaker diaphragm regions. Then, it is the turn of the antenna structure, and lastly the planar MOSFETs for the 16-bit shift register. All of the systems will be interconnected appropriately, the contacts for power will be implanted near the MOSFETs as they are the most sensitive components to power instability, and anything that may be covered will be coated by the aforementioned SiN passivation layer.

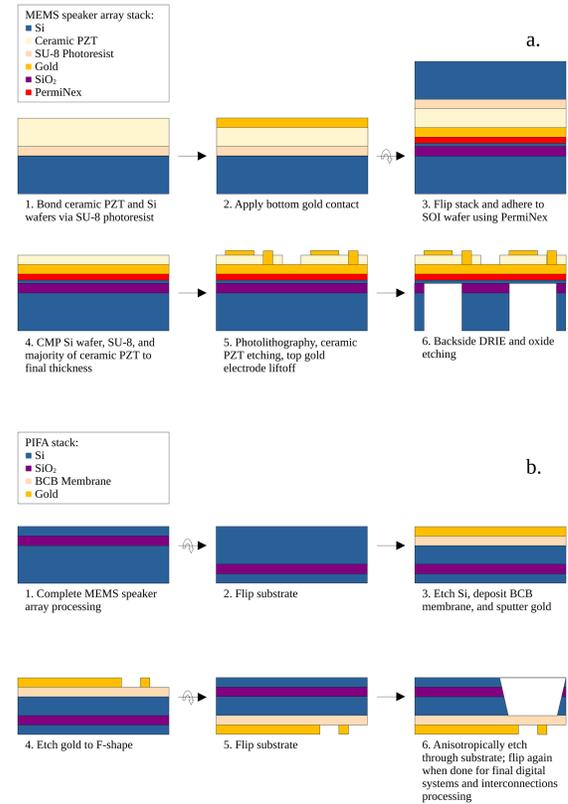


Figure 6: Process flow for a. MEMS speaker array, b. PIFA.

For proper digital speaker design, the resonant frequency should be at 44.1 kHz [2]. Using values for a similar material [5] used in their ceramic PZT from [12], and formulas provided by [13], to maintain the 400  $\mu\text{m}$  diameter of a speaker, the thickness of the ceramic PZT needs to be 1.8  $\mu\text{m}$ . To maintain the approximate thickness [5] made their ceramic PZT layer into, the diameter of each microspeaker needs to be 650  $\mu\text{m}$ , so the array would have a side length total of at least 10.4 mm. Studying [5, 13], there does not seem to be any reason the smaller speakers are not possible, so for the present the design will continue with those specifications in mind.

$F_0 = \frac{1.654t}{a^2} \sqrt{\frac{E}{\rho(1-\nu^2)}} [13]$		
$F_0$ (kHz) [3]	44.1	
$\rho$ (kg/m <sup>3</sup> ) [12]	7.9E3	
$\nu$ [13]	~ 0.3	
$E$ (GPa) [12]	40	
$t$ ( $\mu\text{m}$ )	0.002	5.500
$a$ ( $\mu\text{m}$ )	400	700

*Table 1: Equation and values for determining suitable ceramic PZT membrane dimensions*

From here, per [5], the first step in producing the MEMS speakers is to adhere the ceramic PZT onto a Si wafer (not the final substrate) using SU-8 photoresist, followed by grinding it down to 100  $\mu\text{m}$ . The bottom electrode, made of gold, is then sputtered onto the top of the ceramic PZT. Both the final SOI substrate and the gold side of ceramic PZT are spin-coated with PermiNex, an adhesive material, and the entire stack is now flipped upside-down. From here, the stack receives ample CMP, clearing away the previous Si wafer & SU-8 and bringing the ceramic PZT to its final thickness. Now, the ceramic PZT is patterned and wet-etched, the top electrode is sputtered onto the top of the ceramic PZT, and to finish Si and SiO<sub>2</sub> on the bottom side of the speaker is etched via deep reactive-ion etching and oxide etching, leaving behind only the Si that is directly touching the ceramic PZT. The bottom is etched in this manner to provide the necessary acoustic cavity and, of course, to allow the membrane to be flexible.

The next phase in this process is to create the antenna. [8] specifies that a relatively large trench needs to be micromachined below where the F-shaped antenna will lie, as seen in Fig. 3. [13] points out that isotropic etching produces very rounded and subsequently imprecise trenches, therefore anisotropic etching should be used instead. However, it is further specified that anisotropic etching involves etching at 54.74° off-axis on (100)-oriented Si, as in on the [111] plane.

Therefore, the trench is more of a pyramid or a trapezoid (depending on when the etch is terminated), rather than an upright rectangle, hence the disparity in the trench shapes between Fig. 3 and Fig. 6. Before this micromachining can take place, however, the BCB membrane first needs to be deposited onto the substrate. Due to the orientation of the MEMS speakers, that portion of the processing necessitated a SOI wafer where the oxide (as in the bottom surface) is actually near the top. The substrate is flipped, and the BCB polymer may be applied like a conventional photoresist [7], by preparing, applying, and baking the BCB [13]. At this point, the gold antenna and ground plane is sputtered on. [8] describes a ground plane of about 2.6  $\times$  2.2 mm, assuming the antenna length is 2.5 mm. It, however, is tuned to 24 GHz, 10 times higher than the desired center frequency. This antenna should be more like 20-25 mm long [6], so it can be expected that the ground plane should follow suit and be somewhere around 22.5 mm long on each side. Then comes the anisotropic etch described before, wherein the substrate is flipped again. Per [14], as long as the volume of the underside trench is equivalent to the calculated rectangular prism, its character and center frequency should be equivalent. With that complete, the digital circuitry may be applied between the antenna and the microspeaker array.

The digital circuitry uses standard CMOS processing that, for brevity, will not be detailed here. The important specification to note, per [8], is that for these purposes a fairly rudimentary 130 nm planar CMOS process is utilized. The interconnections are then created via copper interconnects. The 16-bit shift register may also be designed using this same process. An introductory previous report [15] indicates that the necessary performance and size constraints would be easily realizable with this process node for this application, as the presently discussed

process node is about 10 times smaller, more than accounting for the quadrupled footprint the shift register would otherwise take up, allowing for a total footprint of about  $600 \mu\text{m}^2$ .

## VI. CHALLENGES AND MITIGATION

The main challenge in the processing of this product is in the attachment of the ceramic PZT and subsequent etching. The grain size of the ceramic PZT is up to  $3 \mu\text{m}$  in size [5]. The layers (Si, Au, PermiNex) it is sandwiched between may keep it intact, and it seems that etching will not cause rough/otherwise problematic interfaces at the edges. Still, some semblance of delicacy and accuracy is needed for this portion of the processing flow so as to ensure long-term reliability. Since the microspeaker array is the first portion of the device that is manufactured, on-wafer testing can be utilized to reduce wasted materials and time, should a particular batch face issues, before the rest of the components are manufactured onto the wafer.

The other possible trouble may be fragility of each die, since a significant portion of the substrate has been etched out. For this reason, the microspeaker array and antenna's trench have been brought as centrally to each die as possible, so that it does not crumble at the edges.

## VII. CONCLUSION AND FUTURE WORK

Described in this paper is a MEMS-based microspeaker system designed with cost-effective mass-production in mind. Nearly the entirety of the electrical components can be easily processed onto a single standard SOI wafer using long-established and mature process techniques that the industry is well accustomed to using. The performance of the device is expected to be reasonable, especially for the low cost, and mass-production to the scale possible fits the idealized

use-case of widespread customer-service adoption.

## VIII. REFERENCES

- [1] Wang, H. et al. [2021] "Review of Recent Development of MEMS Speakers." *Micromachines* **2021**, *12*, 1257. <http://doi.org/10.3390/mi12101257>.
- [2] Diamond, B. et al. [2003] "Digital Sound Reconstruction Using Arrays of CMOS-MEMS Microspeakers." *TRANSDUCERS '03. 12th International Conference on Solid-State Sensors, Actuators and Microsystems. Digest of Technical Papers, vol.1*, <http://doi.org/10.1109/SENSOR.2003.1215297>.
- [3] Casset, F. et al. [2015] "A 256 MEMS membrane digital loudspeaker array based on PZT actuators." *Procedia Engineering, 120. Proc. Eurosensors XXIX*. <http://doi.org/10.1016/j.proeng.2015.08.563>.
- [4] Mayrhofer, D. et al. [2020] "A New Method of Digital Sound Reconstruction." *DAGA 2020 Hannover*. [https://pub.dega-akustik.de/DAGA\\_2020/data/articles/000093.pdf](https://pub.dega-akustik.de/DAGA_2020/data/articles/000093.pdf)
- [5] Wang, H. et al. [2019] "A Piezoelectric MEMS Loud Speaker Based on Ceramic PZT." *2019 20th International Conference on Solid-State Sensors, Actuators and Microsystems & Eurosensors XXXIII (TRANSDUCERS & EUROSENSORS XXXIII)*. <http://doi.org/10.1109/TRANSDUCERS.2019.8808528>
- [6] Texas Instruments Incorporated [2013] "AN-1811 Bluetooth Antenna Design" <https://www.ti.com/lit/an/snoa519b/snoa519b.pdf>.

- [7] Deen, J. [2012] "Chapter Six - Substrate-Integrated Antennas on Silicon." *Advances in Imaging and Electron Physics*, vol. 174. <https://doi.org/10.1016/B978-0-12-394298-2.00006-5>.
- [8] Texas Instruments Incorporated [2002] "Product Brief: Bluetooth™ Transceiver Complete in CMOS." [https://www.ti.com/pdfs/wtbu/bluetoothcmos\\_120402.pdf](https://www.ti.com/pdfs/wtbu/bluetoothcmos_120402.pdf).
- [9] Xia, B. et al. [2003] "A GFSK demodulator for low-IF Bluetooth receiver." *IEEE Journal of Solid-State Circuits*, vol. 38, no. 8. <http://doi.org/10.1109/JSSC.2003.814424>.
- [10] Ufine Battery [2024] "Exploring Lithium Battery Dimensions: Different Battery Sizes." <https://www.ufinebattery.com/blog/different-battery-sizes/>.
- [11] Harman, G. [2010] *Wire Bonding in Microelectronics*. ISBN: 978-0-07-164265-1
- [12] Ling, Y. [2022] "PZT-5H and PZT-5A." *HE-SHUAI BLOG*. <https://www.he-shuai.com/pzt-5h-and-pzt-5a/>.
- [13] Campbell, A. [2013] *Fabrication Engineering at the Micro- and Nanoscale, Fourth Edition*. ISBN: 978-0-19-986122-4.
- [14] Pozar, M. [2012] *Microwave Engineering, Fourth Edition*. ISBN: 978-0-470-63155-3.
- [15] Bissias, J. [2023] "Analysis and Designing a 4-Bit Shift Register Using VLSI Techniques." <https://jbissias.xyz/pdf/jasonBissiasVLSIFinalReport.pdf>.