Review

Review of Recent Development of MEMS Speakers

Haoran Wang [1], Yifei Ma [2], Qincheng Zheng [2], Ke Cao [2], Yao Lu [2] and Huikai Xie [2,3,*]

1   Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611, USA; wanghaoran@ufl.edu
2   School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China; 3120200697@bit.edu.cn (Y.M.); 1120171582@bit.edu.cn (Q.Z.); 3220210564@bit.edu.cn (K.C.); y.lu@bit.edu.cn (Y.L.)
3   BIT Chongqing Center for Microelectronics and Microsystems, Chongqing 400030, China
*   Correspondence: hk.xie@ieee.org

Abstract: Facilitated by microelectromechanical systems (MEMS) technology, MEMS speakers or microspeakers have been rapidly developed during the past decade to meet the requirements of the flourishing audio market. With advantages of a small footprint, low cost, and easy assembly, MEMS speakers are drawing extensive attention for potential applications in hearing instruments, portable electronics, and the Internet of Things (IoT). MEMS speakers based on different transduction mechanisms, including piezoelectric, electrodynamic, electrostatic, and thermoacoustic actuation, have been developed and significant progresses have been made in commercialization in the last few years. In this article, the principle and modeling of each MEMS speaker type is briefly introduced first. Then, the development of MEMS speakers is reviewed with key specifications of state-of-the-art MEMS speakers summarized. The advantages and challenges of all four types of MEMS speakers are compared and discussed. New approaches to improve sound pressure levels (SPLs) of MEMS speakers are also proposed. Finally, the remaining challenges and outlook of MEMS speakers are given.

Keywords: microelectromechanical systems; MEMS; loudspeakers; microspeakers

1. Introduction

With the rapid advancement of consumer electronics, the worldwide audio market has been seeing a growing trend towards smaller devices with lower power consumption and better performance in the last decade. Speakers, as one of the core components in mobile electronic devices such as laptops, smartphones, wireless earbuds, and human-machine interfaces, are highly demanded to be smaller, lighter, and more power efficient. Currently, speakers in those mobile electronic devices are dominated by conventional speakers with bulky moving coils, which are still challenging to be batch fabricated since voice coils and permanent magnets must be assembled [1]. The miniaturization of these conventional speakers also has a negative impact on the sound quality and reaches some limits due to the employed materials and the fabrication approaches [2]. For example, the plastic or polymer diaphragms of conventional speakers are too soft to be used as high-quality radiator surfaces [3]. The simplification of the mechanical suspensions and electromagnetic parts in the miniaturization would lead to reduced bandwidths and increased nonlinearities, thus deteriorating the sound quality [2,4]. It is also difficult for conventional manufacturing technologies to achieve high dimensional precision and good reproducibility in the miniaturization of speakers.

By contrast, microelectromechanical systems (MEMS) speakers, or microspeakers, have been drawing more and more attention due to their inherent advantages, e.g., small form factors, low power consumption, batch fabrication, and potential on-chip integration with electronic circuits. Many researchers have developed MEMS speakers based

on various transduction mechanisms and achieved promising results, including electro-dynamic MEMS speakers [5–7], electrostatic MEMS speakers [8,9], piezoelectric MEMS speakers [10–12], and thermoacoustic MEMS speakers [13,14]. Various materials and fabrication approaches have also been explored for developing MEMS speakers [15–17]. The performances of MEMS speakers have been evaluated and compared with conventional speakers in terms of several key specifications, such as device footprint, output sound pressure level (SPL), power consumption, bandwidth, and total harmonic distortion (THD) [2,9,10,17]. Among them, the SPL and bandwidth are two widely used parameters to evaluate the acoustic performance of MEMS speakers. THD, defined as the sum of all power radiated in frequencies other than the fundamental frequency relative to the total emitted sound power, is an important parameter to evaluate the sound quality of MEMS speakers [9].

To date, MEMS speakers have been developed mainly for in-ear applications (e.g., hearing aids) and headphones [15,18]. It is challenging for MEMS speakers and conventional electrodynamic microspeakers as well to achieve both high SPL output and flat audio frequency response due to the vibration mode complexity of the diaphragm and the limited space for actuation. Thus, the actuation method, structure, and electrode pattern design of the diaphragm as well as the enclosure design are crucial to the overall response and performance of a MEMS speaker. Both finite element analysis (FEA) and lumped element modelling (LEM) are typically employed to study the effects of various design parameters and to optimize the overall performance of MEMS speakers [19,20]. Several approaches in terms of material selection [12], special structural design [21,22], and electrode configuration [23,24] have also been demonstrated to achieve the better acoustic performance of MEMS speakers. Extensive research efforts have been devoted to developing better MEMS speakers with promising results demonstrated, which is evidenced by a large amount of literature produced.

With so many research efforts paid to the development of MEMS speakers, significant progress has been made in their commercialization. For example, piezoelectric MEMS speakers developed by Usound have reached the market. With a chip size of 6.7 mm × 4.7 mm × 1.58 mm, the developed piezoelectric MEMS speaker can generate a high SPL of around 116 dB in an acoustic coupler, under a driving voltage of 15 V [25]. The TDK Corporation has developed a series of piezoelectric speakers called PiezoListen. With a thickness of as small as 0.49 mm and footprints ranging from 20 mm × 10 mm to 66 mm × 30 mm, the developed speakers can be installed on almost any kinds of displays or surfaces to generate sound over a wide frequency range from 400 Hz to 20 kHz [26]. In addition, Audio Pixels has successfully implemented a digital sound reconstruction (DSR) technique in a commercially feasible manner and developed MEMS speaker arrays to generate high quality sounds [27]. Furthermore, by using moving beams with electrostatic actuation to generate sound inside silicon chips, Arioso Systems has developed MEMS speakers with high-fidelity sound and CMOS-compatible process for in-ear applications [28].

In order to better leverage the existing achievements, it is necessary to sort out the recent development of MEMS speakers, understand the barriers, compare different types of MEMS speakers, and point out the future perspectives with respect to these challenges. Thus, the main purpose of this article is to provide a state-of-the-art review of MEMS speakers and a future outlook as well.

This review article is organized as follows. In Section 2, we introduce the theories and modeling of MEMS speakers, including device concepts, LEMs, and several transduction mechanisms. In Section 3, we review different types of MEMS speakers, including their fabrication technologies, characterization results, and approaches to improve the SPLs of MEMS speakers with regard to structures, materials, and actuation methods. The focus is on the piezoelectric MEMS speaker. In Section 4, we compare and discuss the performances of different MEMS speakers. In Section 5, we summarize the review and discuss future perspectives of MEMS speakers.

2. Theory and Modeling of MEMS Speakers

2.1. Basic Structure

In general, the main structure of a MEMS speaker consists of an acoustic diaphragm, an actuation mechanism, and an air chamber. When an AC voltage is applied to drive the MEMS speaker, a bending moment will be generated by the actuation mechanism, forcing the diaphragm to vibrate and thus generating a sound pressure output. Considering a circular vibrating diaphragm, as shown in Figure 1, the pressure amplitude can be calculated based on the Helmholtz equation and the Rayleigh integral and is readily given by [29]:

$$P(z) = \rho(2\pi f)^2 \int_0^a \frac{w(r)}{\sqrt{z^2 + r^2}} r dr \tag{1}$$

where ρ is the air density, f is the vibration frequency, a is the radius of the diaphragm, $w(r)$ is the vibration amplitude at the radial distance of r, and z is the distance from the diaphragm to the listener.
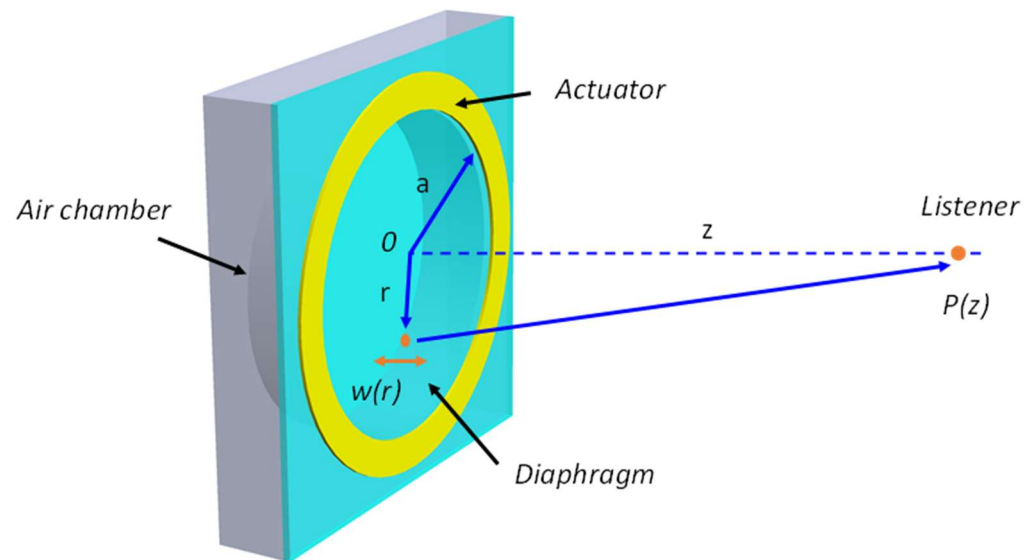

Figure 1. Schematic of a MEMS speaker with a piston-move diaphragm and the geometries for sound pressure calculation.

When the vibration of the acoustic diaphragm is simplified as a piston on an infinite baffle, the effective sound pressure output $P_e(z)$ and the sound pressure level (SPL) in decibels (dB) can be further simplified as:

$$P_e(z) = \frac{P(z)}{\sqrt{2}} = \frac{\sqrt{2}\pi\rho S w f^2}{z} \tag{2}$$

$$SPL = 20 \lg(\frac{P_e(z)}{P_{ref}}) \tag{3}$$

where S and w are the surface area and vibration amplitude of the diaphragm, respectively. The reference effective sound pressure value P_{ref} is 20 μPa [30]. Typically, the SPLs of MEMS speakers are measured by microphones placed at 1 cm away from the MEMS speakers in open air. For MEMS speakers specifically developed for in-ear or hearing-aid applications, their SPLs are measured in a 2cc coupler (a coupler with a volume of 2 cm^3 that conforms to the ANSI S3.7 and IEC 60318-5 standards) [31].

The acoustic diaphragm is important in MEMS speaker designs. According to Equation (2), the sound pressure output generated by the acoustic diaphragm is proportional to its surface area and vibration amplitude, and the square of the working frequency. Thus, generating high sound pressure output at lower frequencies is more challenging, which re-

quires larger deflections under the same diaphragm size constraint, as indicated in Figure 2. Figure 2 plots the required deflection amplitudes for circular diaphragms with different frequencies and different diameters to achieve a 90 dB SPL at 1 cm. This plot shows the decreasing trend of the required deflection amplitudes with the increasing frequencies and the diaphragm sizes and gives a general indication of the design values. As can be seen, for a circular diaphragm with a diameter of 4 mm, achieving a 90 dB SPL at 1 cm requires a diaphragm deflection of 5.9 μm, 94.4 μm, and 1.05 mm at frequencies of 4 kHz, 1 kHz, and 300 Hz, respectively.
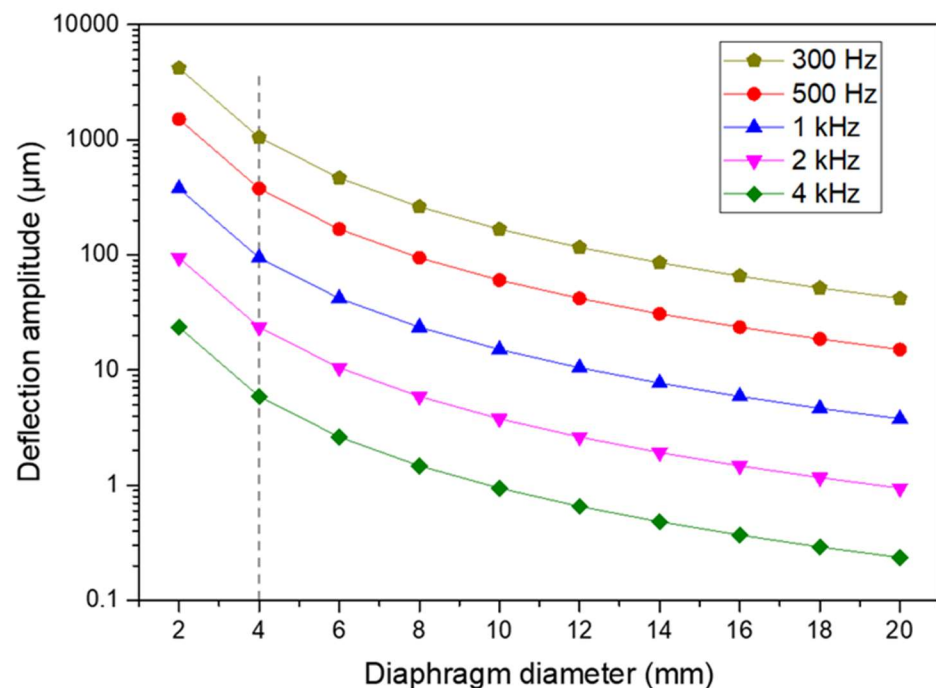

Figure 2. Required deflection amplitudes for different diaphragm diameters to achieve a 90 dB SPL at 1 cm at frequencies of 300 Hz, 500 Hz, 1 kHz, 2 kHz, and 4 kHz.

MEMS speakers are usually designed to work in a frequency range from 20 Hz to 20 kHz, which is consistent with the hearing range of humans. Since the frequencies of audible sounds for humans in daily life typically varies from 100 Hz to 10 kHz, including speeches in low frequencies (300 Hz–3.4 kHz) and musical harmonics in high frequencies (>6 kHz) [32], MEMS speakers are normally evaluated in both of these low-frequency and high-frequency bands. In general, piezoelectric, electrodynamic, and electrostatic actuation are the three most commonly used approaches to excite acoustic diaphragms. Details of these transduction mechanisms will be introduced in Section 2.2.

In addition to the deflection, resonant frequency is another important design parameter of acoustic diaphragms. Most of MEMS speakers presented in literatures are developed based on deformable diaphragms with edges clamped on the substrate. Their fundamental resonant frequencies are dependent on the dimensions and material properties of the diaphragms. For a circular clamped vibrating diaphragm, the fundamental resonant frequency f_0 is given by [33]:

$$f_0 = 0.47 \frac{t}{a^2} \sqrt{\frac{E}{\rho_m (1 - v^2)}} \tag{4}$$

where t, a, E, ρ_m, and v are the thickness, radius, effective Young's modulus, mass density, and Poisson's ratio of the circular diaphragm, respectively.

The fundamental vibration mode of the clamped diaphragm is the so-called drum mode, whose deflection profile peaks at the center of the diaphragm and decreases from the center to the edge. When designing the fundamental resonant frequency, there are two considerations. On one hand, to achieve high SPL at low frequencies and thus improve the acoustic performance over a wide frequency range, the fundamental drum mode frequencies are typically designed at around 2 kHz to 3 kHz [16,34,35]. On the other hand, from the acoustic point of view, the drum mode vibration with the deformed emissive surface and higher harmonics stimulation due to nonlinearities will distort the acoustic wavefront, therefore causing sound distortions and deteriorating the sound quality [2]. Thus, some special diaphragm designs other than edge clamped diaphragms have been developed, such as rigid diaphragms with radial rib structures supported by suspension beams [2] and circular diaphragms supported by four flexible dual-curve actuators [21], in which piston mode vibrations at low frequencies can be employed to generate the sound while the drum mode vibrations can be shifted to high frequencies to avoid the sound distortion of MEMS speakers.

2.2. Transduction Mechanisms

MEMS speakers have been developed based on various transduction mechanisms, including the piezoelectric transduction [36], electrodynamic transduction [37], electrostatic transduction [38], and thermoacoustic transduction [13]. Figure 3 shows the schematics of MEMS speakers with different transduction mechanisms. Among them, MEMS speakers developed based on the first three types of transduction mechanisms rely on the mechanical vibration of the acoustic diaphragm to generate the sound. By contrast, thermoacoustic MEMS speakers produce the sound by the periodic contraction and expansion of the medium around the diaphragm due to the heat exchange between the diaphragm and the surrounding medium.
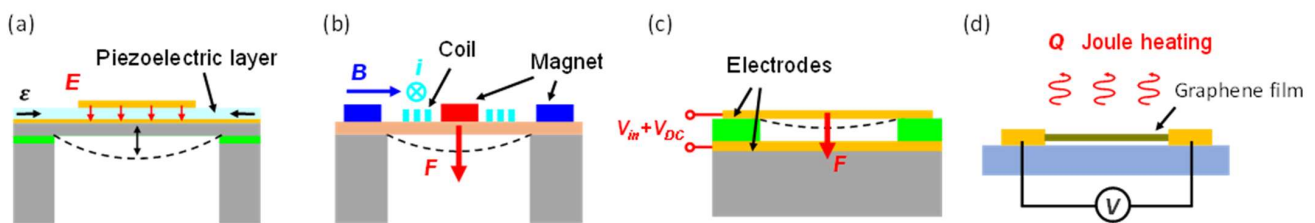

Figure 3. Schematic of MEMS speakers based on (**a**) piezoelectric, (**b**) electrodynamic, (**c**) electrostatic, and (**d**) thermoacoustic transduction mechanisms.

As shown in Figure 3a, piezoelectric MEMS speakers work on the flexural vibration of the piezoelectric diaphragm. When an AC voltage is applied across the piezoelectric film sandwiched by two metal electrodes, an in-plane strain will be generated based on the converse piezoelectric effect, thus causing the out-of-plane vibration of the diaphragm. The relation between the in-plane strain ε and the applied electric field E can be expressed by [21]:

$$\varepsilon = d_{31}E \tag{5}$$

where d_{31} is the piezoelectric constant of the employed piezoelectric film.

For electrodynamic MEMS speakers, the acoustic diaphragm is actuated by electromagnetic (Lorentz) force. As shown in Figure 3b, when the current flows through coils, Lorentz force will be generated due to the interaction between the external magnetic field and the electric current, thus bending the acoustic diaphragm. For a planar concentric coil with N turns carrying an electric current I, the Lorentz force $F_{Lorentz}$ generated by a magnetic field with a flux density B can be expressed as [2]:

$$F_{Lorentz} = I \int_0^l \vec{B} d\vec{l} = \sum_{i=1}^{N} 2\pi I R_i B_i \tag{6}$$

where l is the length of the coil, R_i is the radius of the ith turn, and B_i is the radial component of the magnetic flux density on the coil plane corresponding to the ith turn.

Electrostatic MEMS speakers are driven by the electrostatic force between two conductive plates. As shown in Figure 3c, the acoustic diaphragm is suspended over the substrate by a small gap d. Considering this structure as a parallel-plate capacitor with flat and rigid electrodes for simplification, the electrostatic force exerted on the diaphragm under an AC driving voltage V_{in} and a DC bias V_{DC} is given by [39]:

$$F_E = \frac{1}{2}\epsilon A\left(\frac{V_{in} + V_{DC}}{d}\right)^2 \tag{7}$$

where ϵ is the electric permittivity of air and A is the area of the diaphragm. Advanced models considering the bending of the plate and pull-in limitations are presented in [40,41].

Different from the mechanical vibration sound generators described above, thermoacoustic MEMS speakers emit sound by the thermoacoustic effect, which converts the Joule heat into sound. As shown in Figure 3d, when an AC current is applied to a conductive film, the film will be heated and exchange the thermal energy with the surrounding air, causing the periodic contraction and expansion of the air, thus generating sounds. The root-mean-square sound pressure amplitude produced by a thermoacoustic thin film speakers can be derived as [42]:

$$p_{rms} = \frac{\sqrt{\alpha}\rho_0}{2\sqrt{\pi}T_0} \cdot \frac{1}{r} \cdot P_{in} \cdot \frac{\sqrt{f}}{C_s} \cdot M \tag{8}$$

where ρ_0, α, and T_0 are the mass density, thermal diffusivity, and temperature of the ambient gas, respectively, r is the distance between the thin film conductor and the listener, P_{in} is the input power, f is the frequency of the sound, C_s is the heat capacity per unit area of the thin film conductor, and M is a frequency-related factor.

2.3. Modeling

The acoustic performance of MEMS speakers is dependent on many design parameters, including material properties, device structures, and acoustic enclosure designs. Lumped element modeling (LEM) and finite element analysis (FEA) can be used to effectively predict the acoustic performance of MEMS speakers and optimize the designs. For example, Neumann Jr. et al. presented CMOS-MEMS diaphragms for acoustic actuation based on electrostatic force, and developed a simplified acoustic model to investigate the effects of the dimensional parameters of the diaphragms [43]. Huang et al. studied the sound pressure response of miniaturized moving-coil loudspeakers using an equivalent circuit method (ECM), which can simulate the electrical, mechanical, and acoustical responses and optimize the device designs [44]. These methods are also called electro-mechano-acoustical modeling, or lumped element modeling (LEM), which can be applied to study the effects of different acoustic enclosures and model the performances of MEMS speakers based on different transduction mechanisms [16,45]. LEM is a simple and efficient tool for designing and analyzing multiphysics systems as well as for predicting their responses. In this method, the representation of spatially distributed physical systems is simplified by using a set of lumped elements when the length scale of the device is much smaller than the wavelength of the governing physical phenomenon. Since the acoustic wavelengths (34.3–343 mm for 1–10 kHz) for MEMS speakers are much greater than their sizes (1–10 mm), LEM is applicable.

Typically, MEMS speakers are packaged in enclosures with a front cover, a back chamber, and vent holes. Figure 4a illustrates a simplified structure of a MEMS speaker in a package. The LEM of this device is shown in Figure 4b, representing a multiphysics system consisting of electrical, mechanical, and acoustical energy domains. In the electrical domain, the effort and flow are voltage (in V) and current (in A), respectively. The electrical and mechanical domains are coupled by a transformer or a gyrator that models the transduction mechanism of the MEMS speaker. In the mechanical domain, the effort represents the force (in N) that actuates the vibrating diaphragm while the flow represents the velocity of the diaphragm (in m/s). The acoustical domain is coupled to the mechanical domain by the effective area of the diaphragm. Thus, the effort and flow in the acoustical domain correspond to the pressure (in Pa) and volume velocity (in m^3/s), respectively. The lumped elements sharing the same effort are connected in parallel, while those sharing the same flow are connected in series.

In the electrical domain, the electrical input impedance of the MEMS speaker is modeled as Z_e, which can be resistance and inductance from the wires and coils for electro-dynamic MEMS speakers, or capacitance and resistance for electrostatic MEMS speakers and piezoelectric MEMS speakers. The electrical domain is coupled to the mechanical domain by a transformer (or a gyrator), representing the energy transformation from the electromagnetic force, electrostatic force, or the piezoelectric force.

In the mechanical domain, the vibrating diaphragm is modeled as a mass-spring-damper system, governed by the following equation:

$$M_d \frac{d^2 w}{dt^2} + R_d \frac{dw}{dt} + \frac{w}{C_d} = F_t \tag{9}$$

where w is the vibration amplitude of the diaphragm, F_t is the total force applied on the diaphragm, and M_d, C_d, and R_d are the equivalent mass, compliance, and damping of the diaphragm, respectively. The mechanical and acoustical domains are coupled with the effective area of the diaphragm, which converts the actuation force to the acoustic pressure. Two separate transformers are used to account for the front and the back sides of the diaphragm [44].

In the acoustical domain, the air in an acoustic chamber with a volume V_a can be modeled as an acoustic compliance C_a that is readily given by [44]:

$$C_a = \frac{V_a}{\rho_a c^2} \tag{10}$$

where ρ_a and c are the air density and the sound speed, respectively. Therefore, the front volume and the back chamber can be modeled as acoustic compliances $C_{f,v}$ and $C_{b,c}$, respectively. The air flow inside narrow spaces can be modeled as acoustic resistances and masses, such as $R_{f,h}$ and $M_{f,h}$ of the acoustic holes in the front side and $R_{b,v}$ and $M_{b,v}$ of the backside vent. The acoustic radiation impedance of the diaphragm is also approximated as the acoustic resistance $R_{f,rad}$ ($R_{b,rad}$) and mass $M_{f,rad}$ ($M_{b,rad}$). Details of the calculation of these lumped elements are described in [44].
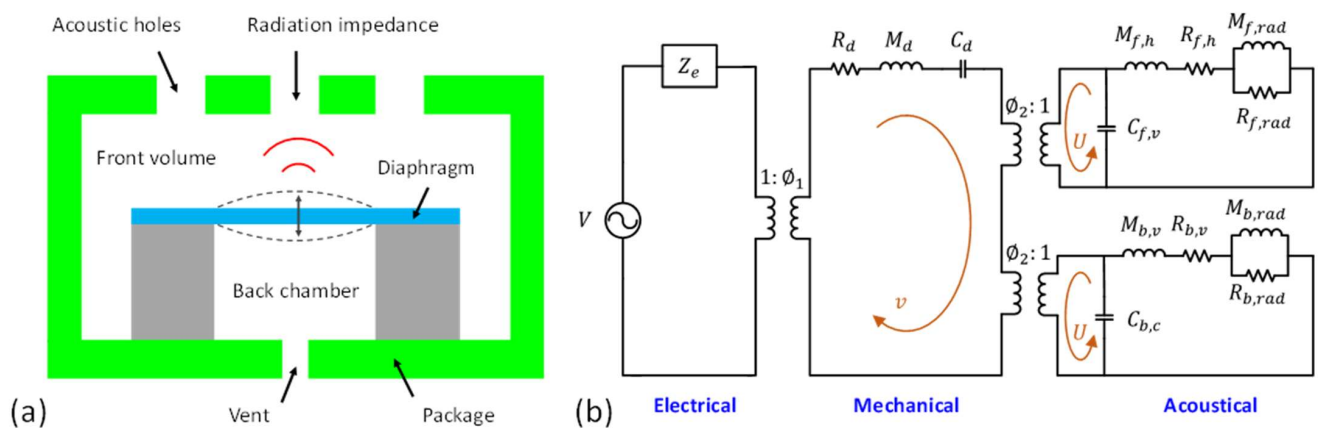
Figure 4. Lumped element model (LEM) of a packaged MEMS speaker: (**a**) illustration of the structures and (**b**) the equivalent circuit in multiple domains.

By solving the equivalent circuit in the LEM, the volume velocity U generated by the MEMS speaker in the acoustical domain can be obtained. Thus, by assuming the MEMS speaker as a point source at a far-field distance r (much larger than the Rayleigh distance), the sound pressure output P can be calculated as [45]:

$$P = j\frac{k\rho_a c}{2\pi r} \cdot U \cdot e^{j(\omega t - kr)} \tag{11}$$

where k and ω are the wave number and angular frequency of the acoustic wave, respectively. Here, it is worthy of note that the MEMS speaker is considered as a monopole mounted on a baffle plate for far-field calculation. Since the plate restricts the acoustic radiation only to the forward hemisphere, the pressure is twice that of a free radiation without a baffle plate [46].

The LEM has been widely applied to predict the dynamic responses of MEMS speakers, especially at low-frequency regions or in the neighborhood of the fundamental resonant frequency due to its simplicity [16,47]. However, the LEM is not sufficient to model higher order resonant modes and incapable to well predict the high frequency responses of MEMS speakers. Therefore, LEM is often used together with FEA to calculate the dynamic responses [5,44,48], analyze the enclosure designs [45,49], and optimize the diaphragm structural designs.

3. Development of MEMS Speakers

The study of MEMS speakers started in the late 1990s. Since then, significant progress has been made to develop MEMS speakers based on different transduction mechanisms, especially on piezoelectric, electrodynamic, and electrostatic transduction. To achieve a small size, high output sound pressure, and flat frequency response, various materials, structure designs, and fabrication techniques have been employed. In this section, the development of MEMS speakers will be reviewed based on their transduction mechanisms.

3.1. Piezoelectric MEMS Speakers

3.1.1. Design and Fabrication of Piezoelectric MEMS Speakers

Piezoelectric actuation, with the advantages of small driving voltage and large actuation force, has been widely used in many MEMS devices, including ink-jet printer heads [50], MEMS scanning mirrors [51], ultrasonic motors [52], RF resonators [53], and acoustic generators [54]. Among them, piezoelectric MEMS speakers are important applications and are attracting more and more interest. Piezoelectric MEMS speakers based on different piezoelectric materials, such as zinc oxide (ZnO), aluminum nitride (AlN), and lead zirconate titanate (PZT), have been presented for hearing aid or earphone applica-

tions [35,55,56]. Piezoelectric MEMS speakers mainly consist of a piezoelectric vibration diaphragm and an acoustic cavity. Typical vibration diaphragms can be designed as beam-like piezoelectric actuators [57] (Figure 5a), fully clamped diaphragms with piezoelectric layers embedded [12] (Figure 5b), or partially clamped diaphragms surrounded by piezo-electric actuators [21] (Figure 5c). Various piezoelectric MEMS speakers based on different designs have been demonstrated [12,58].
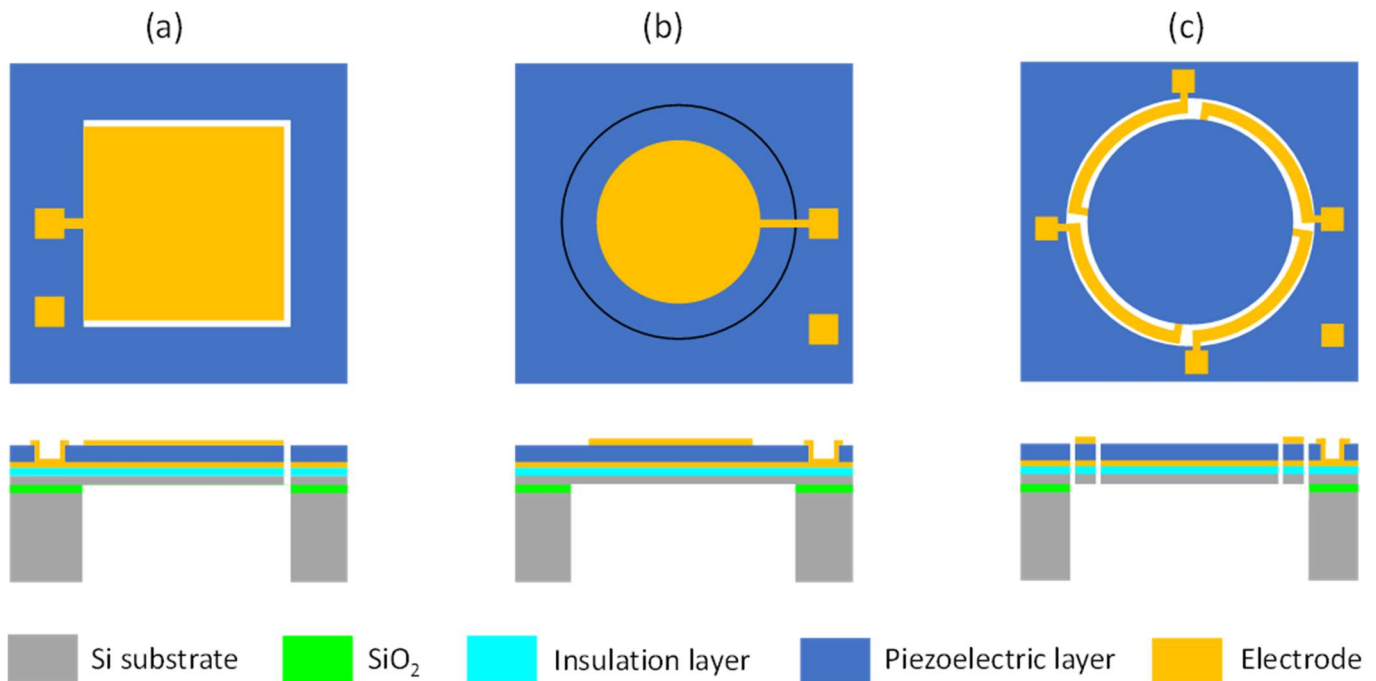

Figure 5. Schematics of typical structures of piezoelectric MEMS speakers in top view (**top**) and cross-sectional view (**bottom**). (**a**) Beam-like piezoelectric actuator. (**b**) Fully clamped diaphragm with piezoelectric layer embedded. (**c**) Partially clamped diaphragm surrounded by piezoelectric actuators.

The fabrication process of piezoelectric MEMS speakers with various structures can be different, depending on whether the diaphragm needs to be released from both sides (Figure 5a,c) or the backside only (Figure 5b), but their general steps are similar. Here, an example for the design of MEMS speakers with a partially clamped diaphragm (Figure 5c) is presented to illustrate the typical fabrication process. As shown in Figure 6, firstly, an insulation layer (Si_xN_y or SiO_2), a bottom electrode layer, and a piezoelectric layer are deposited in sequence on a silicon-on-insulator (SOI) substrate (Figure 6a). After that, the piezoelectric layer is patterned by wet etching or reactive ion etching (RIE) to expose the bottom electrode [59,60] (Figure 6b). Next, a top electrode is deposited and patterned (Figure 6c). After that, RIE is used to define a diaphragm and a set of piezoelectric actuators on the front side (Figure 6d). Subsequently, the acoustic cavity is defined on the backside with a two-sided photolithography and formed by the deep reactive ion etching (DRIE) of silicon or wet etching with KOH (Figure 6e). The buried oxide layer is used as the etch stop and finally removed by RIE or vapor hydrofluoric acid to release the moveable structures (Figure 6f). For the fabrication of fully clamped diaphragms in Figure 5b, the process step shown in Figure 6d can be skipped.

In the design and fabrication of piezoelectric MEMS speakers, the material of the piezoelectric layer is important as it will affect the selection of the fabrication method and the performance of the fabricated devices. Next, the piezoelectric materials for making MEMS speakers will be discussed.
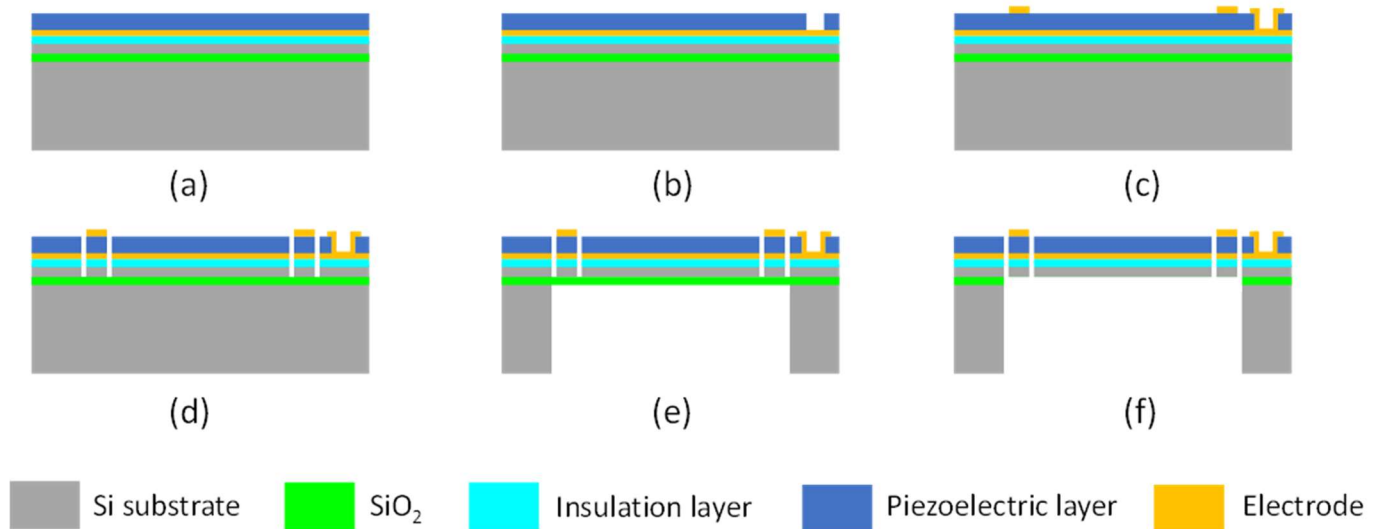
Figure 6. Typical fabrication process flow of a piezoelectric MEMS speaker. (**a**) Deposit insulation, electrode, and piezoelectric layers. (**b**) Pattern the piezoelectric layer. (**c**) Deposit and pattern the top electrode layer. (**d**) Reactive ion etching to define the diaphragm and piezoelectric actuators. (**e**) Etch the acoustic cavity. (**f**) Release the moveable structures.

3.1.2. Piezoelectric Materials

Lead zirconate titanate (PZT) ceramics, single-crystal lithium niobate (LiNbO$_3$), and single-crystal lead magnesium niobate-lead titanate (PMN-PT) are widely used bulk piezoelectric materials with high piezoelectric coefficients and electromechanical coupling factors for piezoelectric transducers [61]. However, how to thin down these materials remains an issue in fabricating piezoelectric MEMS devices. With the advancement of thin film deposition technologies, piezoelectric thin films including ZnO, AlN, and PZT can be fabricated by sputtering or sol-gel methods, which have been applied to fabricate piezoelectric MEMS devices, such as microspeakers [62,63]. Among these materials, ZnO is one of the most commonly used for making piezoelectric thin film devices such as film bulk acoustic wave resonators (FBAR), surface acoustic wave (SAW) resonators, piezoelectric micromachined ultrasonic transducers (pMUTs), and microspeakers in early years. ZnO-based piezoelectric MEMS speakers have been developed as early as in 1996, when Lee et al. fabricated a piezoelectric cantilever transducer that worked both as a microphone and a microspeaker [58]. In their system, the $2000 \times 2000 \times 4.5 \ \mu m^3$ piezoelectric cantilever was fabricated based on a 0.5 μm-thick ZnO layer with the magnetron sputtering method. In 2003, Ko et al. presented a piezoelectric microspeaker based on a clamped $3000 \times 3000 \times 3 \ \mu m^3$ diaphragm. This micromachined transducer also has a thin ZnO film as the piezoelectric layer, which is deposited on a membrane of low-stress silicon nitride of 1.5 μm [64].

Another type of piezoelectric material, AlN, has also been well studied and characterized in the past few decades. A thin film of AlN is normally deposited by the reactive magnetron sputtering method. Sputtered AlN thin films have better chemical and thermal stability than ZnO. The lower conductivity of AlN compared to ZnO also results in lower power loss [65]. With these advantages, AlN has also been a good candidate for fabricating the piezoelectric layer of MEMS speakers. In 2007, Seo et al. presented piezoelectric microspeakers with circular-type and cross-type electrode configurations based on a 0.5-μm-thick AlN film [36]. With a diaphragm size of 4×4 mm^2, the AlN-based microspeakers achieved good acoustic performance with a high sound pressure level (SPL).

However, it is challenging to sputter ZnO and AlN with controlled properties. Their morphology and crystalline quality will highly affect the piezoelectric constants of materials. In a fabrication process, the sputtering rate and residual stress are dependent on the sputtering condition and film thicknesses [66,67]. Sputtering with heated substrates (above 300 °C) have been reported with large residual stresses [35,68], which will wrinkle the

diaphragm of fabricated piezoelectric MEMS speakers and affect the sound pressure output. It is possible to deal with such residual stress problem by adding a stress compensation layer or fabricating dome-shaped diaphragms to reduce the effect of the residual stress. For example, in 2000, Han et al. reported dome-shaped piezoelectric MEMS speakers built on 1.5-μm-thick Parylene diaphragms, which can easily release the residual stress through volumetric shrinkage or expansion [69]. In 2009, Yi et al. reported piezoelectric AlN MEMS speakers with improved performance by controlling the residual stress of the compressively stressed diaphragm using Si_xN_y films [35]. The results revealed that the SPLs of the piezoelectric AlN microspeakers were increased by more than 10 dB when the residual stresses became more compressive, especially at the low frequency region.

Other limitations of sputtering ZnO and AlN thin films include low deposition rates (tens of nm/min), small film thicknesses, and small piezoelectric constants [67,70]. The lower value of piezoelectric constants will directly limit the vibration amplitude of a piezoelectric diaphragm and lead to poor acoustic performance. By contrast, PZT thin films have greater piezoelectric constants and are favorable for the applications of piezoelectric actuation. The sputtering and sol-gel methods have also been employed to deposit PZT thin films with typical thicknesses of 0.5–2 μm, which can be applied to a wide range of applications [63]. For example, in 2009, Cho et al. fabricated a piezoelectric MEMS speaker based on a sol-gel PZT thin film with a thickness of 700 nm [11]. The fabricated MEMS speaker had a circular diaphragm with a diameter of 2 mm, which achieved SPLs of 79 dB at 1 kHz, 87 dB at 5 kHz, and 90 dB at 10 kHz under a driving voltage of 13 V. However, sputtered and sol-gel PZT films also suffer from residual stresses and limited thicknesses. Thicker sol-gel PZT films require multiple coatings and high temperature annealing, which will cause serious stress issues. Moreover, since the piezoelectric properties of deposited thin films are largely dependent on the crystal orientation and substrate condition, proper buffer layers are required to prevent the material interdiffusion and oxidation and help to obtain good piezoelectric properties with lower residual stress.

The material properties of these commonly used piezoelectric thin films and the commercial ceramic PZT are summarized in Table 1. Since most of piezoelectric MEMS speakers work on the d_{31} mode of the piezoelectric layer, only the d_{31} piezoelectric constant is listed in the table for comparison. Among these materials, AlN thin films have the smallest piezoelectric constant, while PZT thin films exhibit the highest piezoelectric constant, which is about 10 to 20 times greater than that of ZnO thin films. However, the piezoelectric constant of PZT films also vary in a wide range, dependent on the film thickness, deposition, and poling conditions. In particular, the piezoelectric coefficient of the commercial ceramic PZT (e.g., PZT-5H) can reach 300 pm/V [71], which makes it a promising candidate for the construction of piezoelectric transducers.

Table 1. Material properties of commonly used piezoelectric thin films and the commercial ceramic PZT [21,71–76].

Property	ZnO	AlN	Sol-Gel PZT	Sputtered PZT	Ceramic PZT-5H
Density (kg/m^3)	5700	3260	7700	7700	7800
Young's modulus (GPa)	98.6	283	96	96	50
Dielectric constant	8.8	8.5–10.7	650–1470	400–980	3400
Piezoelectric constant d_{31} (pm/V)	3.9–5.5	2–2.6	23–76	45–102	270–300

3.1.3. Approaches to Improve SPLs

Although a large number of piezoelectric MEMS speakers have been demonstrated based on various piezoelectric thin films with promising results, inadequate sound pressure level (SPL) outputs and non-flat frequency responses are common challenges of these devices. High SPLs of over 90 dB were achieved in a few piezoelectric MEMS speakers, but they were measured either in canals or ear simulators or at high-frequency resonances. Piezoelectric MEMS speakers with high SPLs (90 dB or above) over wide frequency ranges, especially in open air and low-frequency range, are needed for broader applications such

as mobile phones, laptops, wearable electronics, and Internet of Things (IoT) devices. Therefore, several approaches have been proposed to improve the SPLs of piezoelectric MEMS speakers in terms of materials and fabrication processes and structure designs, which will be reviewed in the following.

Materials and Fabrication Processes

As discussed in Section 3.1.1, the commonly used piezoelectric thin films of ZnO and AlN deposited by sputtering or sol-gel methods suffer from large residual stresses and limited thickness. For sputtered or sol-gel PZT, their obtained piezoelectric constants are also not comparable with those of bulk piezoelectric crystals or ceramics. As illustrated in Table 1, the piezoelectric constant of ceramic PZT is over four times greater than that of sputtered or sol-gel PZT. Thus, ceramic PZT was gradually employed in fabricating the piezoelectric layer of MEMS speakers with particular fabrication process to thin down this material. In 2009, Kim et al. thinned ceramic PZT down to around 40 μm and fabricated piezoelectric MEMS speakers based on it, and they measured an SPL of 90 dB ($\pm$5 dB) in the audible frequency range under a 32-V_{pp} drive at 1 cm away from the MEMS speaker in an anechoic box [17]. The fabricated MEMS speaker also exhibited a total harmonic distortion (THD) of less than 15% from 400 Hz to 8 kHz. However, the acoustic diaphragm was as large as 20 mm $\times$ 18 mm.

Since the resonant frequency of a diaphragm is affected by its area and thickness, scaling down the diaphragm size requires a thinner piezoelectric layer to maintain a proper resonant frequency. In 2020, Wang et al. presented a piezoelectric MEMS speaker based on thin ceramic PZT [16]. By using wafer bonding and chemical mechanical polishing techniques, ceramic PZT was thinned down to only 5 μm and applied to fabricate MEMS speakers. An optical image of the fabricated MEMS speaker and a cross-section SEM image of the device layers are shown in Figure 7a1,a2. Thin ceramic PZT not only exhibits much greater piezoelectric constants than sol-gel or sputtered PZT thin films but also has a wider range of thicknesses, thus allowing the scaling of diaphragms within size restrictions for different applications. With a 6 mm diameter diaphragm, the fabricated MEMS speaker achieved a maximum SPL of 119 dB measured at 1 cm under a 10-V_{pp} drive, as shown in Figure 9a [16].
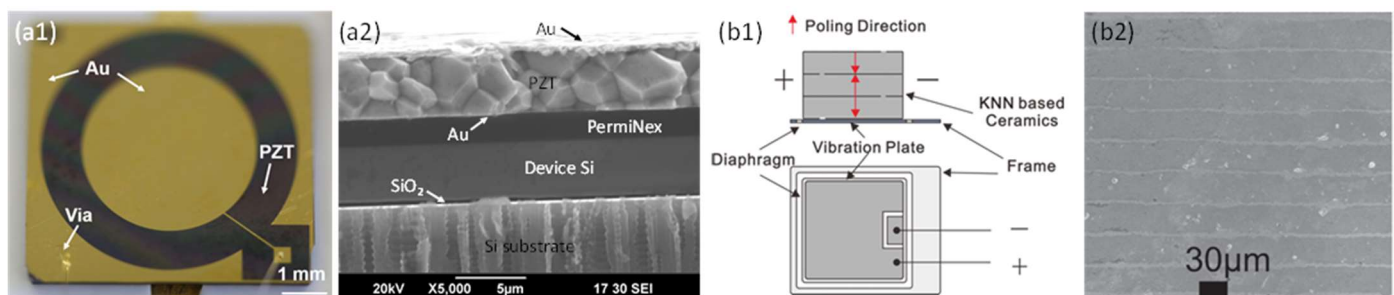

Figure 7. Piezoelectric MEMS speakers based on new materials: (**a1**) optical image and (**a2**) cross-section SEM image of a thin ceramic PZT-based MEMS speaker (Reproduced with permission from Elsevier [16]); (**b1**) schematic and (**b2**) cross-section SEM image of multilayer ceramics of a KNN ceramics-based MEMS speaker (Reproduced with permission from IOP [77]).

Furthermore, lead-free piezoelectric ceramics with high piezoelectric constants have also been explored for fabricating piezoelectric MEMS speakers. For example, in 2014, Gao et al. fabricated piezoelectric MEMS speakers using potassium sodium niobate ((K,Na)NbO$_3$, KNN)-based multilayer piezoelectric ceramics [77]. They employed a tape casting and cofiring process and used Ag–Pd alloys as an inner electrode. A schematic of the multilayer ceramics based piezoelectric MEMS speaker and a cross-section SEM image of the multilayer KNN-based ceramics are shown in Figure 7b1,b2, respectively. With a form factor of 23 $\times$ 27 $\times$ 0.6 mm^3, using three layers of 30-μm-thick KNN-based ceramics,

the fabricated MEMS speakers showed an average SPL of 87 dB from 1 kHz to 20 kHz measured at 3.16 cm under a 5-V_{rms} drive.

Structure Designs

As illustrated in Section 2.1, the output SPL of a MEMS speaker is directly determined by the frequency, area, and displacement of its diaphragm. Increasing the out-of-plane displacement of piezoelectric diaphragms is an effective approach to improve SPLs, especially at low frequency, as a much larger displacement is required at low frequency to achieve the same SPL at high frequency. Therefore, various designs of piezoelectric MEMS speakers have been proposed to improve their SPLs by changing the diaphragm structures, electrode configurations, or using an array form to enhance their acoustic performance.

Diaphragm Structures

In 2018, Stoppel et al. demonstrated a piezoelectric MEMS speaker based on a 2-μm-thick sputtered PZT with two open cuts on a square diaphragm (4×4 mm^2) for in-ear applications, as shown in Figure 8a [18]. Without a closed diaphragm, four individual actuators are mechanically decoupled from each other and thus can achieve larger out-of-plane displacements. The measurement in an ear simulator showed a high SPL of above 81 dB from 20 Hz and above 100 dB from 4.7 kHz to 15.8 kHz under a 2-V_{pp} drive, as shown in Figure 9b. The measured THD was less than 2% at most frequencies, except for the subharmonics of the resonance frequency, where the THD was increased to 7%.

In 2020, Cheng et al. presented a piezoelectric MEMS speaker with enhanced SPL by designing suspension-spring actuators with a dual-electrode driving [21]. As shown in Figure 8b, the designed MEMS speaker consisted of a circular moveable diaphragm and four flexible spring actuators. Dual-curve spring actuators with dual-electrode driving were utilized to achieve larger displacements than single-curve spring actuators under the same form factor. Measurements in a 3-cm-long tube showed a maximum SPL of 90.1 dB at the resonance of 1.85 kHz under a 2-V_{pp} drive, which was 28 dB higher than the SPL of a fully clamped diaphragm speaker at the same frequency (Figure 9c). The measured THD of the dual-curve spring device was also lower than those of the clamped diaphragm devices, which was less than 2% at most frequencies and low than 8% at the resonant or harmonic frequencies.

In addition to employing unsealed vibration diaphragms with large displacements, Wang et al. proposed a rigid–flexible vibration coupling mechanism in 2021. By depositing a Parylene film on a pre-etched diaphragm, the fabricated MEMS speaker can maintain large displacements of the unsealed diaphragms without acoustic loss. Measurement in an ear simulator under a 2-V drive showed SPLs can exceed 59 dB from 250 Hz to 20 kHz, with the maximum value of 101.2 dB obtained at the resonance of 6.7 kHz [78].

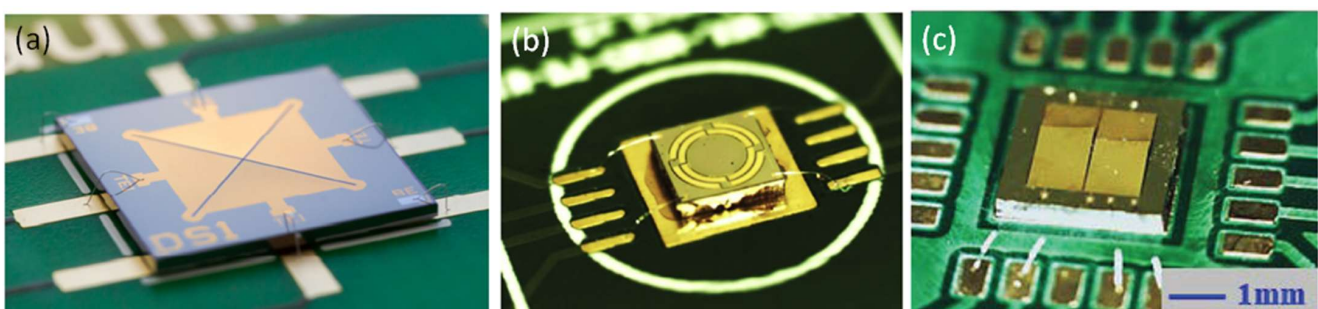**Figure 8.** Optical images of piezoelectric MEMS speakers with novel structural designs. (**a**) A diaphragm with two open cuts (Reproduced with permission from IEEE [18]). (**b**) A diaphragm with suspension-spring actuators (Reproduced with permission from Elsevier [21]). (**c**) A diaphragm formed by four piezoelectric cantilevers with different dimensions (Reproduced with permission from IEEE [79]).
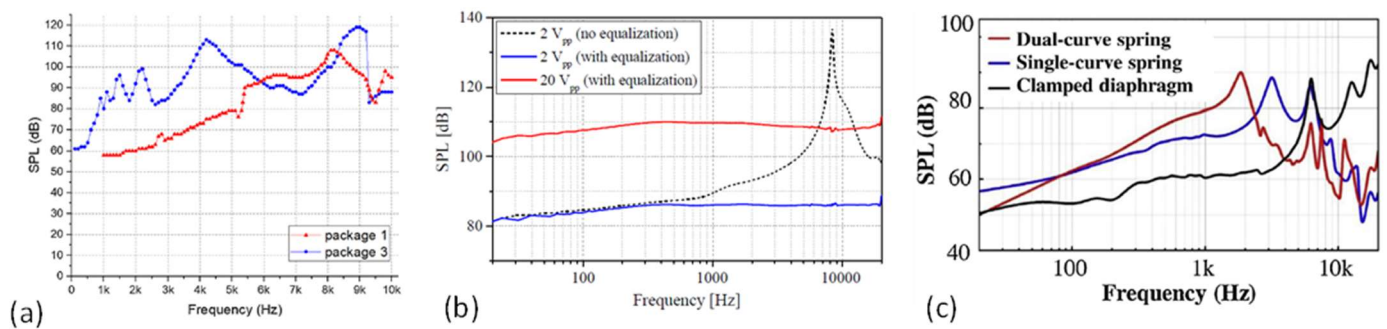
Figure 9. Measured frequency response of piezoelectric MEMS speakers with (**a**) a thin ceramic PZT-based diaphragm in free field at 1 cm (Reproduced with permission from Elsevier [16]), (**b**) a square diaphragm with open cuts in an ear simulator (Reproduced with permission from IEEE [18]), and (**c**) dual-curve spring actuators in a 3-cm-long tube (Reproduced with permission from Elsevier [21]).

To improve SPLs over a broad frequency range, in 2021, Wang et al. proposed a cantilever array design with an in-phase/out-of-phase hybrid driving method to realize a broadband piezoelectric MEMS speaker [79]. As shown in Figure 8c, the device consisted of four piezoelectric cantilevers with different dimensions, the four resonance frequencies of which contribute to the broadband performance of the MEMS speaker. In this device, in order to avoid the sound pressure cancellation due to the large phase shifts around the resonances of the cantilevers, a hybrid drive voltage with a combination of both in-phase and out-of-phase signals was applied to ensure that the cantilevers vibrate in the same direction. Measurements showed a broadband frequency response from 100 Hz to 10 kHz with an SPL of 70 dB or higher and a maximum SPL of 110 dB at 1.54 kHz in an ear simulator under a 2-V_{pp} drive.

Electrode Configurations

Efforts have also been devoted to improving the SPLs of MEMS speakers by the special design of electrode configurations. Electrode configurations on piezoelectric diaphragms are important as they largely determine the excitation mode, vibration displacement, and electromechanical coupling efficiency. As introduced in Section 2.1, most piezoelectric MEMS speakers work on the d_{31} flexural vibration mode of piezoelectric diaphragms with the electrical field applied in the thickness direction and the strain generated in the lateral directions. In addition to the d_{31} vibration mode, piezoelectric materials can also be excited in the d_{33} mode with the applied electrical field and the generated stain in the same direction, typically in the thickness direction. Typically, the magnitude of the d_{33} constant of a piezoelectric material is roughly two times larger than that of the d_{31} constant. Therefore, by proper electrode configurations, the d_{33} mode of piezoelectric diaphragms can be excited with larger out-of-plane displacements than the d_{31} mode. In 2015, Kim et al. presented a piezoelectric MEMS speaker based on the d_{33} mode PMN-PT single crystal diaphragm with a circular inter-digitated electrode (IDE) configuration and studied the effects of the patterned electrodes on the acoustic characteristics of the MEMS speaker [23]. A single crystal PMN-PT was thinned down to 10 µm to form an 8.5 mm diameter diaphragm by grinding, polishing, and inductively-coupled-plasma (ICP) etching, followed by metallization with circular IDE patterns on the top, as shown in Figure 10a. Measurements showed improved SPL with increasing area of the patterned IDE. With an 8 mm diameter IDE, the MEMS speaker showed an average SPL of above 70 dB from 1 kHz to 10 kHz and a maximum SPL of around 100 dB at 1 cm under a 5-V_{rms} drive.

In addition to the IDE configuration that can excite the piezoelectric d_{33} mode for SPL improvement, dual-electrode configuration has been investigated to improve the SPLs of piezoelectric MEMS speakers working on the d_{31} mode. In 2020, Tseng et al. presented a piezoelectric MEMS speaker with the SPL improved by dual-electrode driving [56]. The schematic of the designed MEMS speaker is shown in Figure 10b, where the square

diaphragm consists of four triangular plates whose vibrations are synchronized by a connection mass. The low frequency response can be enhanced by reducing the size of the gaps between the triangular plates. Each triangular plate can be driven by an inner electrode and an outer electrode with a 180° phase difference to actuate the piston mode of the diaphragm to increase the SPL. Measurements showed a SPL enhancement of 9.5 dB under the dual-electrode driving in comparison with the single (inner or outer) electrode driving.
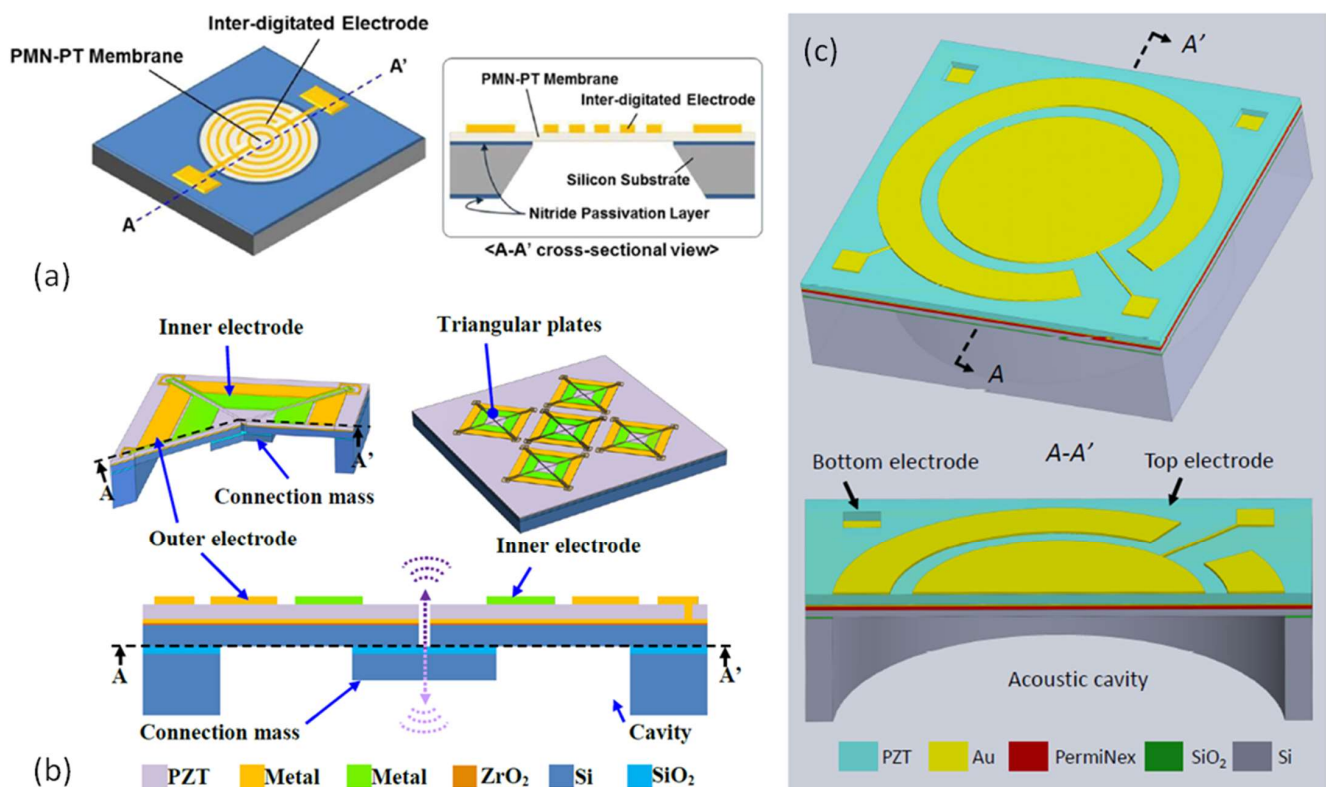

Figure 10. Schematics of piezoelectric MEMS speakers with special electrode configurations. (**a**) Circular inter-digitated electrode (Reproduced with permission from Springer [23]). (**b**) Triangular plates with dual electrode (Reproduced with permission from IEEE [56]). (**c**) Circular diaphragm with dual electrode (Reproduced with permission from IEEE [24]).

In addition to the 180° out-of-phase, other phase differences in dual-electrode driving and their influences on the SPL improvement of piezoelectric MEMS speakers have been studied. In 2021, Wang et al. presented a ceramic PZT-based piezoelectric MEMS speaker with the SPL improved by dual-electrode driving and studied the effects of the phase difference at different frequencies [24]. As shown in Figure 10c, the reported MEMS speaker consists of an inner circular electrode and an outer ring-shaped electrode. By applying sine waves on these two electrodes with a phase difference tuned from 0° to 360° in the experiments, the measurement results revealed that the SPL changed significantly with the phase difference and was frequency dependent, peaking at different phase differences for different frequencies. With the optimal phase differences, a 2–10 dB SPL improvement can be achieved in the frequency band spanning from 600 Hz to 10 kHz, compared with the single-electrode driving method.

Array Structures

Another approach to improve the SPLs of the piezoelectric MEMS speakers is using digital sound reconstruction or speaker arrays. Different from traditional sound generation techniques that rely on the vibration amplitudes and frequencies of a single or a few diaphragms to achieve high SPL at specific frequencies, digital sound reconstruction generates loud sound by adding the outputs of a large number of speaker pixels that can

be excited individually by signals with different frequency compositions [80]. Typically, a speaker array containing 2^n speaker pixels is used in digital sound reconstruction, where *n* is the bit number, and each pixel contributes a small amount of sound pressure in the system. In 2015, Casset et al. implemented digital sound reconstruction with piezoelectric MEMS speaker arrays [81]. Figure 11a shows the fabricated speaker array packaged on an electronic board. With a chip size of 4×4 cm^2, the speaker array contains 256 piezoelectric diaphragms based on a 2-μm sol-gel PZT film. The output SPL of the speaker array reached over 100 dB at 13 cm. In 2016, Arevalo et al. increased the bit number and presented a 10-bit (1024 elements) piezoelectric MEMS speaker array with a chip size of 2.3×2.3 cm^2 [82]. An optical image of part of the speaker array is shown in Figure 11b. The characterization results demonstrated the potential of piezoelectric MEMS loudspeaker arrays for digital sound reconstruction, but more efforts are still needed to optimize the design for better acoustic performances.
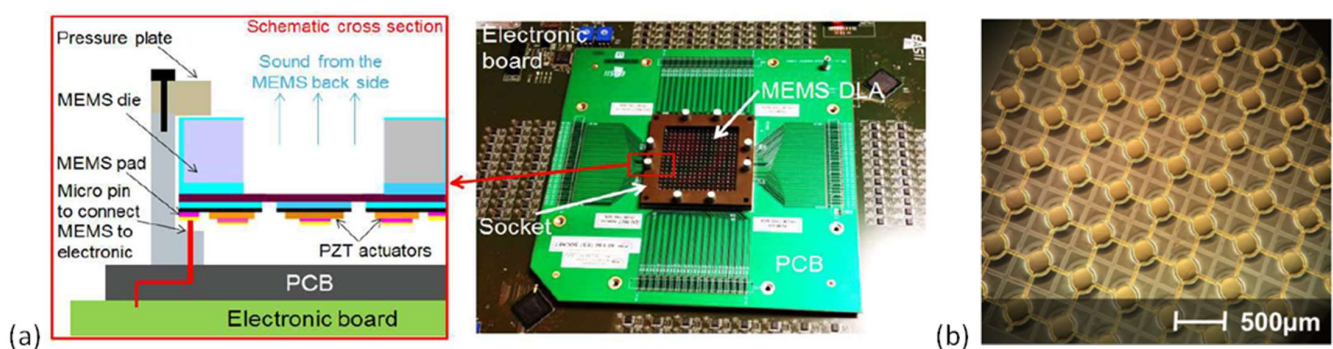

Figure 11. Schematic and optical images of piezoelectric MEMS speaker arrays: (**a**) a 256-speaker array packaged on an electronic board (Reproduced with permission from Elsevier [81]) and (**b**) part of a 1024-speaker array (Reproduced with permission from IEEE [82]).

3.1.4. Summary of Piezoelectric MEMS Speakers

Piezoelectric MEMS speakers are reviewed above from piezoelectric materials, fabrication techniques, and approaches to improve SPLs. Table 2 summarizes the key results of these piezoelectric MEMS speakers. As shown in the table, sol-gel and sputtered PZT films are popular piezoelectric materials for fabricating piezoelectric MEMS speakers due to their higher piezoelectric constants than those of ZnO or AlN films. Piezoelectric MEMS speakers based on sol-gel or sputtered PZT films with thicknesses of 1–2 μm typically have diaphragm sizes of no more than 4 mm and can generate high SPLs over 90 dB in tubes or ear simulators for in-ear applications. With optimized structure designs, their SPLs can be significantly improved to reach maximum values over 110 dB under small driving voltages. Moreover, piezoelectric MEMS speakers based on ceramic PZT or single-crystal PMN-PT can generate high SPLs in open air, which have potential applications in consumer electronics such as cell phones or laptops. Bulk ceramic PZT or PMN-PT with superior piezoelectric properties can be thinned down to 5–40 μm for fabricating piezoelectric MEMS speakers, which enables larger diaphragm designs ranging from 6 mm to 2 cm and high SPLs of over 100 dB at 1 cm in open air.

Table 2. Key results of different piezoelectric MEMS speakers.

	Ref.	Piezoelectric Layer	Diaphragm Size	1st Resonant Frequency	Maximum SPL	Driving Voltage	Note
In-coupler measurement	[58]	0.5 μm ZnO	2 mm length (square)	890 Hz	100 dB at 4.8 kHz	12 V_{PP}	Measured into a 2 cm^3 coupler
	[21]	1 μm sputtered PZT	1.13 mm diameter (central part)	1.85 kHz	90.1 dB at 1.85 kHz	2 V_{PP}	Measured in a 3-cm-long tube
	[78]	1 μm sputtered PZT	2 mm side length (hexagon)	6.7 kHz	101.2 dB at 6.7 kHz	2V (unspecified)	Measured in an ear simulator
	[79]	2 μm sputtered PZT	4 mm^2 (rectangle)	1.54 kHz	110 dB at 1.54 kHz	2 V_{PP}	Measured in an ear simulator
	[56]	2 μm sputtered PZT	3.24 mm^2 (four triangles)	~6 kHz	118.1 dB at 11.9 kHz	2 V_{PP}	5-speaker array, measured in an ear simulator
	[18]	2 μm sputtered PZT	4 mm length (square)	8.3 kHz	138 dB at 8.3 kHz	2 V_{PP}	Measured in an ear simulator
Free-field measurement	[64]	0.5 μm ZnO	3 mm length (square)	7.3 kHz	83.1 dB at 13.3 kHz	30 V_{PP}	Measured at 1 cm
	[10]	0.5 μm ZnO	5 mm length (square)	2.92 kHz	92.4 dB at 2.92 kHz	6 V_{PP}	Measured at 2 mm
	[36]	0.5 μm AlN	4 mm length (square)	–	100 dB at 10 kHz	20 V_{PP}	Measured at 3 mm
	[35]	0.5 μm AlN	–	–	104 dB at 3 kHz	20 V_{PP}	Device in a 4 cm^3 package, measured at 1 cm
	[11]	0.7 μm sol-gel PZT	2 mm diameter	–	90 dB at 10 kHz	13 V (unspecified)	Measured at 1 cm
	[81]	2 μm sol-gel PZT	2.6 mm diameter	18 kHz	~110 dB	8 V (unspecified)	256-speaker array, measured at 13 cm
	[16]	5 μm ceramic PZT	6 mm diameter	4.3 kHz	119 dB at 9 kHz	10 V_{PP}	Measured at 1 cm
	[17]	40 μm ceramic PZT	18 mm × 20 mm	0.49 kHz	~106 dB at 5.5 kHz	32 V_{PP}	Measured at 1 cm
	[23]	10 μm PMN-PT	8.5 mm diameter	1.4–1.84 kHz	~100 dB at 6.5 kHz	$10\sqrt{2}$ V_{PP}	Measured at 1 cm

3.2. Electrodynamic MEMS Speakers

Electrodynamic MEMS speakers have been developed based on electromagnetic actuation, which is the most widely used actuation mechanism in classical speakers. Electrodynamic MEMS speakers have advantages of high power density, low driving voltage, and linear responses. Efforts have been devoted to the development of electrodynamic MEMS speakers with integrated magnetic materials and small form factors at low cost,

while improving their sound performances. However, the full integration of magnetic materials to realize electrodynamic MEMS speakers is still challenging.

In 2004, Cheng et al. presented an electrodynamic MEMS speaker for hearing instruments. The device was fabricated with a low temperature process using an electroplated Ni/Fe soft magnet, which was suitable for post-CMOS processing and potential integration with electronic circuits [5]. A schematic of the designed MEMS speaker is shown in Figure 12a, which has a chip size of 5 mm × 5 mm and consists of a micromachined polymer diaphragm on a silicon wafer, a single-curve Cu coil, an electroplated Ni/Fe soft magnet, and a permanent magnet mounted on the backside. The frequency responses of the fabricated device are measured in air and in a 2-cc coupler with results shown in Figure 13a. At a low driving voltage of 1.5 V, the MEMS speaker generated a maximum SPL of 93 dB at 5 kHz in a 2-cc volume. This work provided a concept and process for micromachining electrodynamic MEMS speakers. Following that, several electrodynamic MEMS speakers have been reported for lower power consumption, high-level integration process, and improved SPL and sound quality.

In 2009, Chen et al. presented an electrodynamic MEMS speaker with improved power efficiency through incorporating Ni nano-composites into Cu to make the voice coil [83]. A cross-sectional view of the MEMS speaker structure is illustrated in Figure 12b, where the coil is made of a Cu–Ni composite by mixing Ni nano-powders with alkaline non-cyanide-based copper-plating solution in a colloidal bath. The frequency responses of the fabricated MEMS speakers driven by the Cu–Ni nanocomposite and pure Cu coils are measured and compared, as shown in Figure 13b. The experimental results showed that the MEMS speaker with a Cu–Ni composite coil can averagely provide about 40% power savings than the one with a Cu coil for the same SPL output at 70 dB.
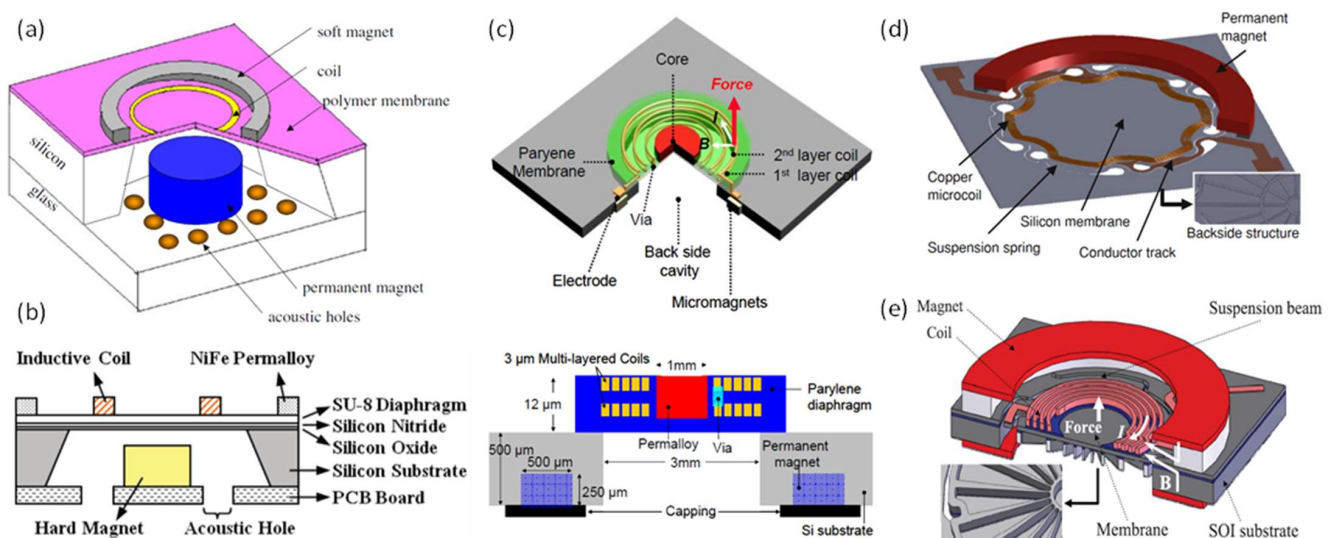

Figure 12. Electrodynamic MEMS speakers: (**a**) typical structure of an electrodynamic MEMS speaker (Reproduced with permission from IOP [5]), (**b**) cross-sectional view of a low-power electrodynamic MEMS speaker with Cu–Ni nanocomposite coil synthesized (Reproduced with permission from Journal of IEEE [83]), (**c**) schematic (top) and cross-sectional view (bottom) of a fully integrated electrodynamic MEMS speaker (Reproduced with permission from IEEE [15]), (**d**) schematic of an electrodynamic MEMS speaker with a rigid silicon diaphragm (Reproduced with permission from Springer [84]), (**e**) schematic of an electrodynamic MEMS speaker showing a rigid silicon diaphragm and the optimized configuration of coil and two face-to-face magnets (Reproduced with permission from IEEE [2]).

As shown in Figure 12a,b, most electrodynamic MEMS speakers require the assembly of a bulky permanent magnet, which will not only increase the overall footprint of the device but also add challenge to the batch fabrication process and precise alignment of the magnet to the diaphragm coil. In order to address this issue, in 2009, Je et al. presented a fully-integrated electrodynamic MEMS speaker with an IC process-compatible micromachined permanent magnets for hearing aid applications [15]. A schematic and cross-sectional view of the presented MEMS speaker is shown in Figure 12c, where a Parylene diaphragm containing embedded multi-turn coils and a soft magnet core is suspended over an acoustic cavity. A rare earth Nd–Fe–B magnetic powder was mixed into a wax binder and dispensed into pre-etched trenches to form the permanent ring micromagnet. The fabricated MEMS speaker produced a 0.64 μm peak displacement at 1 kHz with a 46-mW power consumption. Referring to Figure 2, the achieved displacement is too small to generate sufficient SPLs by a diaphragm with a diameter of 3 mm. Although this work demonstrated the feasibility of fabricating fully integrated electrodynamic MEMS speakers, further design optimization is required to improve the displacement and acoustic performance.

Most MEMS speakers use clamped polymer diaphragms, such as polyimide, Parylene, and SU-8, for flexural vibration and sound generation, whose small mass is in favor of power efficiency and large deflection. However, the flexible nature of polymer diaphragms will lead to dynamic deformations and numerous structural modes within the audio frequency band, thus inducing sound distortion and non-flat frequency response. From 2012 to 2013, to improve the sound performance of electrodynamic MEMS speakers, Shahosseini et al. proposed novel electrodynamic MEMS speakers based on rigid silicon diaphragms and optimized structural designs [2,6,37,84]. The rigid silicon diaphragms were designed with radial ribbed structures for increased stiffnesses and reduced masses, thus enhancing the piston mode vibration for the sound generation and shifting other modes out of the audio frequency band. Figure 12d shows the structure of such an electrodynamic MEMS speaker, where the silicon diaphragm was connected to the substrate by a set of flexible springs to provide out-of-plane displacements [84]. A 14-turn Cu coil was shaped in a special geometry to prevent the damage near the springs' clamp areas, and it was located as close as possible to the permanent magnet to maximize the electromagnetic force. The same research group also investigated the distribution of the magnetic flux density under different configurations of the permanent magnets. In 2013, another electrodynamic MEMS speaker with an optimized microcoil configuration and two face-to-face magnets has been developed [2], as shown in Figure 12e. The fabricated MEMS speaker had a circular diaphragm with a diameter of 15 mm and generated a SPL of around 80 dB at 10 cm starting from 300 Hz to over 20 kHz, as shown in Figure 13c.
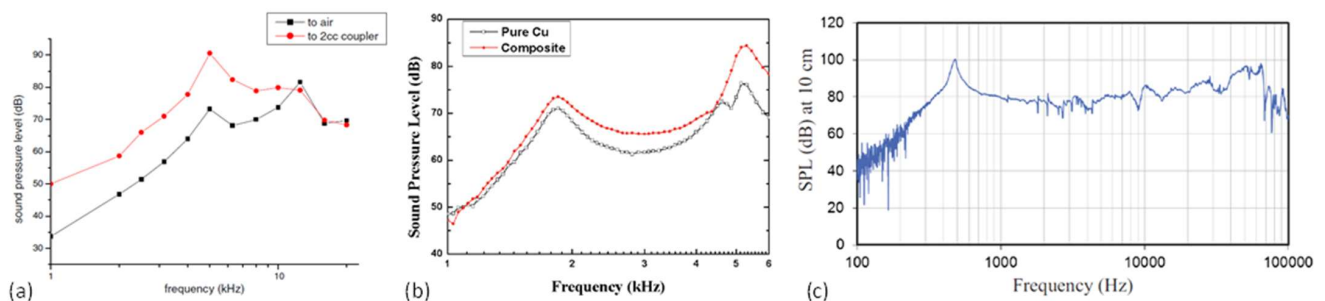

Figure 13. Typical frequency responses of electrodynamic MEMS speakers based on (**a**) a polymer diaphragm and a Cu coil and measured at 2 cm in air and in a 2-cc coupler (Reproduced with permission from IOP [5]), (**b**) pure Cu and Cu–Ni composite coils and measured in a 2-cc coupler (Reproduced with permission from IEEE [83]), (**c**) a rigid silicon diaphragm and measured at 10 cm in air (Reproduced with permission from IEEE [2]).

Table 3 summarizes the key results of these electrodynamic MEMS speakers. As shown in the table, electrodynamic MEMS speakers based on polymer diaphragms typically have small size and low power consumption but limited SPLs. Their maximum SPLs are around 100 dB or less measured in 2-cc couplers or ear simulators. By contrast, electrodynamic MEMS speakers with rigid silicon diaphragms can generate loud sound in open air at large distance but suffer from large diaphragm size and high power consumption.

Table 3. Key results of different electrodynamic MEMS speakers.

Ref	Diaphragm Material	Diaphragm Size	Maximum SPL	Power Consumption	Note
[5]	Polyimide	3.5 mm diameter	93 dB at 5 kHz	320 mW	Measured in a 2 cm^3 volume
[7]	Polyimide	3 mm diameter	106 dB at 1 kHz	0.13 mW	Calculated based on the displacement
[85]	Polyimide	2.5 mm diameter	90 dB at 1,5,10 kHz	–	Measured in a sealed 1500 mm^3 silicone tube
[83]	SU-8	-	Around 85 dB at 5.2 kHz	–	Measured in a 2 cm^3 volume
[86]	PDMS	3.5 mm diameter	106 dB at 1 kHz	1.76 mW	Measured in a 2 cm^3 volume
[2]	Silicon	15 mm diameter	80 dB at 0. 33 kHz	0.5 W	Measured at 10 cm

3.3. Electrostatic MEMS Speakers

MEMS speakers based on electrostatic actuation have also been proposed, which typically consist of parallel or lateral plate actuators. The advantages of such speakers include easy fabrication, high electromechanical efficiency, and relatively flat frequency response. In this section, the recent designs of electrostatic MEMS speakers based on different diaphragm materials will be introduced first. Then, the approaches to improve SPLs of electrostatic MEMS speakers while balancing the design constraints will be reviewed in detail.

3.3.1. Devices with Different Diaphragm Materials

Electrostatic MEMS speakers have been demonstrated based on different diaphragm materials [8,39,87,88]. In 2005, Kim et al. reported an electrostatic MEMS speaker based on a Parylene thin diaphragm. As the cross-sectional SEM image shown in Figure 14a1, the speaker contains two separated chambers on the top and bottom, respectively, which enables bi-directional actuation by electrostatic forces [88]. Figure 14a2 shows the measured frequency response of the speaker. With a diaphragm size of 2×2 mm^2, the fabricated device generated high SPLs of 113.4 dB at 7.68 kHz and 98.8 dB at 13.81 kHz, which were measured at a distance of 1 cm under a driving voltage of 150 V. In 2007, Roberts et al. presented an electrostatically driven touch-mode MEMS speaker based on poly-SiC diaphragms with a diameter of 800 μm, which was robust and operable in harsh environments [8]. Figure 14b1 shows the SEM image of the suspended poly-SiC diaphragm of the fabricated device. At a distance of 1 cm, a maximum SPL of 73 dB was obtained at 16.59 kHz under a driving voltage of 200 V (Figure 14b2). Another material, graphene, has also been explored for fabricating high-quality broad-band audio speakers due to its extremely low mass density and high mechanical strength. In 2013, Zhou et al. presented a miniaturized electrostatic speaker based on a 30 nm thin graphene diaphragm and demonstrated a broad frequency response from 20 Hz to 20 kHz with the performance matching or surpassing a commercial magnetic coil speaker [39].
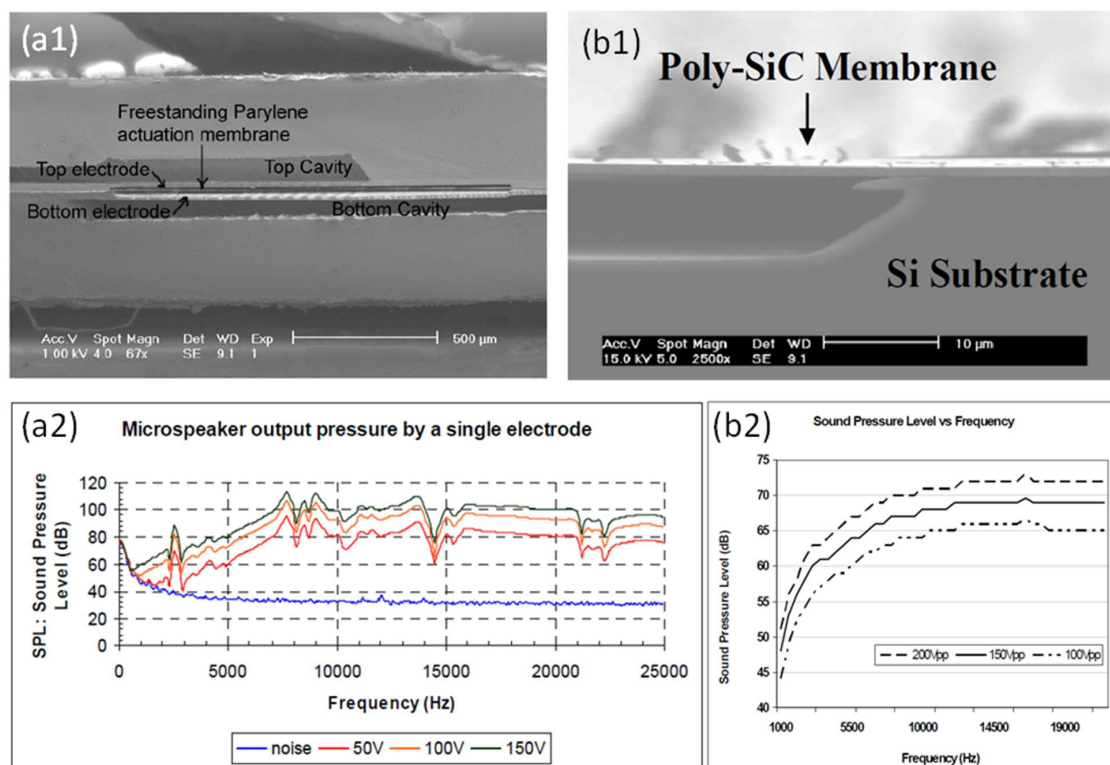
Figure 14. Electrostatic MEMS speakers based on different diaphragm materials: (**a1**) SEM image and (**a2**) measured frequency response of a bi-directional MEMS speaker with a Parylene diaphragm (Reproduced with permission from IEEE [88]). (**b1**) SEM image and (**b2**) measured frequency response of a touch-mode MEMS speaker with a poly-SiC diaphragm (Reproduced with permission from IEEE [8]).

3.3.2. Approaches to Improve SPLs

Most electrostatic MEMS speakers are based on the conventional parallel plate structures and have low SPLs due to the small deflections of their diaphragms, which is a direct result of the balance between the electrostatic force and the mechanical restoring force. In order to overcome the limitation of low SPLs, large electrostatic forces need to be generated. As introduced in Section 2.2 and shown in Equation (7), large electrostatic forces require high driving voltages and small separation gaps. However, the small separation gap will limit the deflection range of the diaphragm and generate large squeeze film air damping [89]. Moreover, the driving voltage must be reasonably lower than the pull-in voltage of the parallel plates to ensure a good reliability. Therefore, tradeoffs have to be made among the electrostatic force, the separation gap between the parallel plates, and the driving voltage to increase the SPLs of electrostatic MEMS speakers.

To generate considerable SPLs and balance the above-mentioned constraints, several approaches in terms of device structure and driving voltage have been applied in the development of electrostatic MEMS speakers [9,38,90–92]. One approach to improve SPL is to use multiple speakers, i.e., array structures. In 2016, Arevalo et al. presented an electrostatic MEMS speaker array for digital sound generation, where each of the individual MEMS speakers had a hexagonal diaphragm connected to an outer hexagonal ring by tethers (Figure 15a) [91]. This work demonstrated the feasibility of generating sounds with electrostatic MEMS speaker arrays but lacked acoustical characterization results.

Different from conventional MEMS speakers that work on the out-of-plane deflection of a diaphragm, Kaiser et al. proposed a novel structure design in 2019, which consisted of in-plane bending electrostatic actuators working in air chambers based on the so-called nanoscopic electrostatic drive (NED) technology, as shown in Figure 15b [9,48]. Utilizing the curvy geometric shape of the moving beams, electrostatic forces are translated into

lateral forces and cause the bending of the beams. Therefore, high SPLs can be reached by the large deflection of the beams in the air chambers and a large number of beams in one chip, without the limitation of small separation gaps between electrodes [93]. This novel structure utilized the chip's bulk volume rather than the surface to generate sound pressures. Figure 15c shows an optical image of such a fabricated electrostatic MEMS speaker with in-plane actuators. The acoustic measurement in an ear simulator showed a SPL of 69 dB at 500 Hz with a THD of 4.4%. The maximum SPL reached 104 dB at 11.4 kHz.

In 2020, Garud et al. designed and fabricated a MEMS speaker with peripheral electrostatic actuation [38]. Figure 15d shows the schematic of the designed electrostatic MEMS speaker, where the clamped circular diaphragm has a peripheral electrode configuration that can mitigate the squeeze film damping effect and increase the pull-in voltage. The simulation results showed that as the peripheral electrode width was reduced from 100% (full electrode coverage) to 10%, the pull-in voltage and the vibration amplitude of the diaphragm could be increased by a factor up to 40 and 80, respectively.

To reduce or eliminate the DC bias of electrostatic MEMS speakers, electrets embedded with quasi-permanent electrical charges have been integrated within the electrode structures. In 2020, Sano et al. presented an electret-augmented electrostatic MEMS speaker and demonstrated its sound generation under low driving voltages [92]. The schematic and an SEM image of the fabricated MEMS speaker are shown in Figure 15e,f, respectively. By integrating the electrets into the MEMS speaker, the built-in electrical potential is equivalent to an external DC bias, thus resulting in an increased displacement or a reduced bias voltage. The characterization result showed that a -10 V electret-augmented electrostatic MEMS speaker reached a maximum SPL of 50 dB at 1.5 cm under a 5-V_{pp} AC driving voltage.
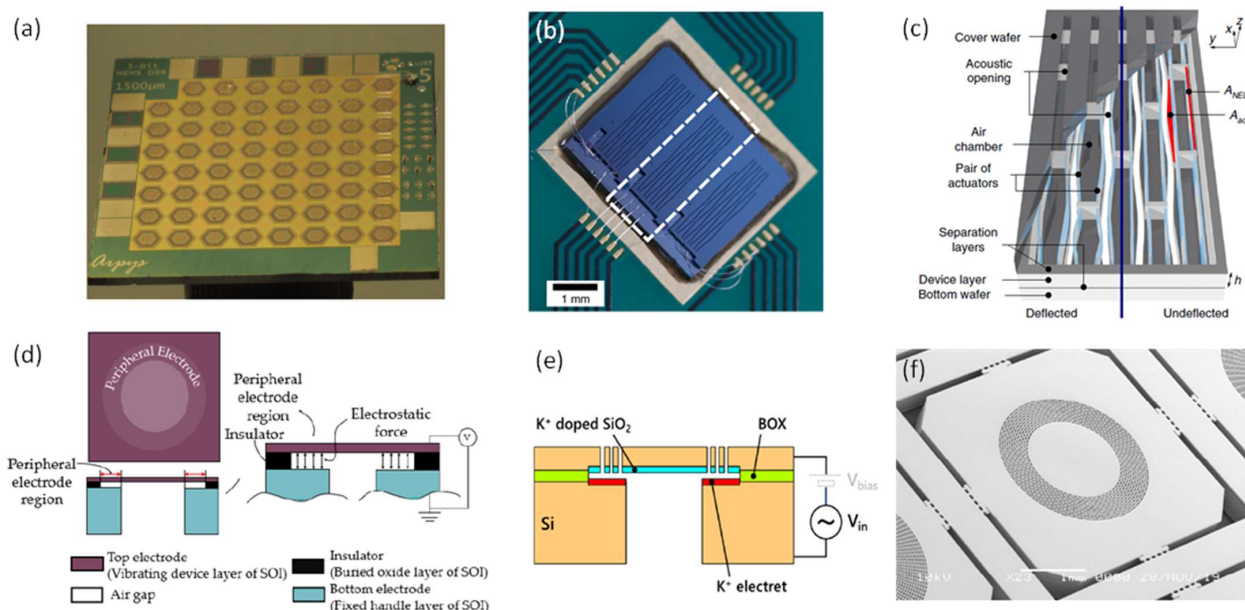

Figure 15. Electrostatic MEMS speakers with special designs: (**a**) optical image of a MEMS speaker array (Reproduced with permission from IEEE [91]), (**b**) schematic and (**c**) optical image of a MEMS speaker with in-plane bending electrostatic actuators working in air chambers (Reproduced with permission from Nature Portfolio [9]), (**d**) schematic of a peripheral electrode configuration (Reproduced with permission from IEEE [38]), (**e**) schematic and (**f**) the corresponding SEM image of an electret-augmented MEMS speaker (Reproduced with permission from MDPI AG [92]).

Table 4 summarizes the representative electrostatic MEMS speakers reported in the literature. It can be seen that electrostatic MEMS speakers typically require high driving voltage and large DC bias to generate considerable diaphragm deflection. Most of electrostatic MEMS speakers have small separation gaps (1–8 μm) and limited sound pressure output. High SPLs are generally obtained only at the high-frequency range.

Table 4. Key results of different electrostatic MEMS speakers.

Ref	Diaphragm Size	Electrode Separation	Maximum SPL	Driving Voltage	Note
[92]	2 mm diameter	2 μm, peripheral electrode	50 dB at around 35 kHz	AC 5 V_{pp}	Measured at 1.5 cm
[8]	0.8 mm diameter	8 μm, touch mode	73 dB at 16.59 kHz	AC 200 V_{pp}	Measured at 1 cm
[38]	3.1 mm diameter	1 μm, peripheral electrode	75–78 dB at above 10 kHz	DC 30 V + AC 30 V	Measured at 1 cm
[88]	2 mm length (square)	7.5 μm	113.4 dB at 7.68 kHz	AC 150 V	Measured at 1 cm
[9]	-	-	104 dB at 11.4 kHz	DC 40 V + AC 10 V_{pp}	Measured in an ear simulator

3.4. Thermoacoustic MEMS Speakers

In addition to the above reviewed three major types of transduction mechanisms, thermoacoustic transduction also has potential to be applied for making MEMS speakers. Several thermoacoustic loudspeakers have been developed based on carbon nanotube or graphene with research efforts focused on increasing the sound pressure output and reducing the power consumption [13,14,42,94,95].

In 2008, Xiao et al. found that thin carbon nanotube films emitted sound when a current in audio frequency was applied, which could be attributed to the thermoacoustic effect [42]. Based on this finding, they successfully fabricated thermoacoustic speakers with A4 paper sizes and cylindrical shapes (9 cm diameter and 8.5 cm height) based on one-layer or four-layer carbon nanotube thin films, which could generate over 70 dB SPLs at 5 cm starting from 1 kHz, with an input power of 3 Watts. Figure 16a shows the photograph of a fabricated thermoacoustic speaker with an A4 paper size. This work demonstrated the feasibility of developing thermoacoustic speakers using carbon nanotube films. However, it required a large device size and a high power consumption to generate high SPLs.
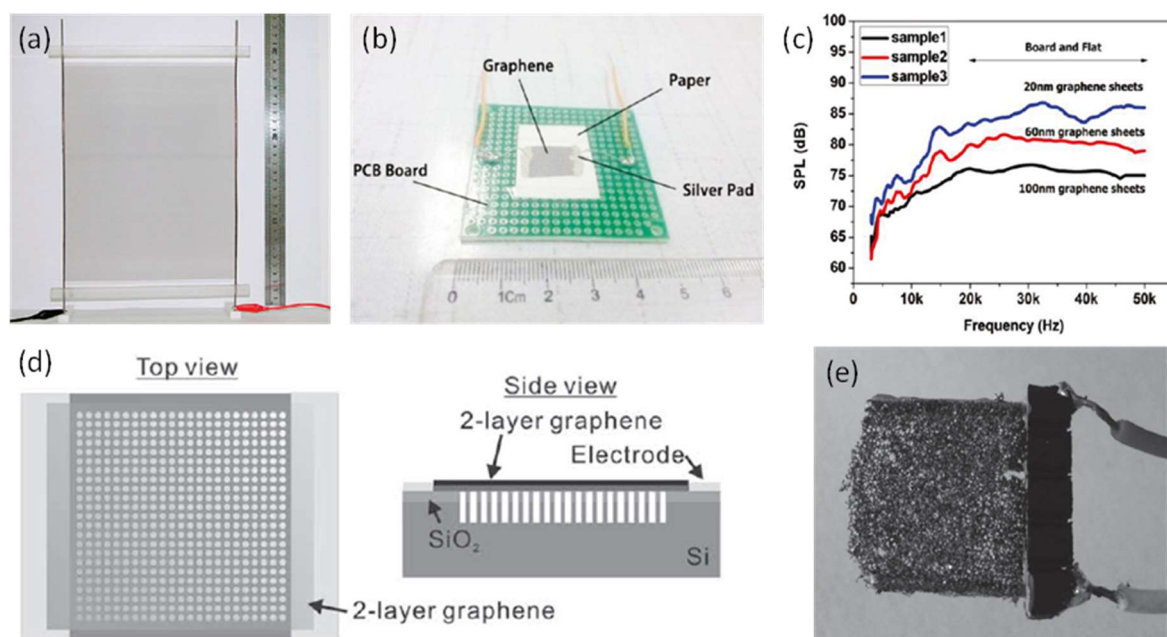

Figure 16. Thermoacoustic speakers: (**a**) photograph of a A4 paper size carbon nanotube thin film thermoacoustic speaker (Reproduced with permission from ACS [42]), (**b**) photograph and (**c**) measured SPL of a graphene-on-paper speaker (Reproduced with permission from ACS [94]), (**d**) schematic of a graphene speaker on a patterned substrate (Reproduced with permission from Wiley [14]), (**e**) photograph of a graphene foam speaker (Reproduced with permission from Wiley [13]).

In 2011, Tian et al. observed thermoacoustic effect on graphene and demonstrated graphene-on-paper speakers [94]. As shown in Figure 16b, the fabricated thermoacoustic speaker had a 1 cm × 1 cm graphene sheet, which was placed on a piece of paper and connected to a printed circuit board (PCB) using silver ink. Graphene sheets with thicknesses of 20 nm, 60 nm, and 100 nm were used to fabricate speakers. Figure 16c shows the SPL curves with the input power density normalized to 1 W/cm^2, which indicated that thinner graphene sheets produced higher SPLs and the SPL of 20 nm graphene sheets reached 85 dB at 5 cm with the frequency increased to over 15 kHz.

The sound performance of thermoacoustic speakers has been further studied and optimized in terms of substrate material and structure design. For example, in 2012, Suk et al. studied thermoacoustic sound generation with graphene on different substrates, including glass, polyethylene terephthalate (PET), and polydimethylsiloxane (PDMS) [14]. The substrate effect was also investigated by transferring graphene onto patterned substrates with different porosities, as shown in Figure 16d. The experiments revealed that graphene on the substrates with lower thermal effusivity and higher porosity exhibited better sound performances. In 2015, Fei et al. presented a low-voltage driven thermoacoustic speaker based on graphene foam synthesized by the nickel-template chemical vapor deposition (CVD) method [13]. A photograph of the fabricated free-standing graphene foam speaker is shown in Figure 16e. Benefited from high thermal conductivity and low in-plane resistance of the 3D graphene foam, the speaker generated a SPL of around 50 dB at 3 cm and 10 kHz with a power consumption of only 0.1 W.

In summary, thermoacoustic speakers made of carbon nanotubes or graphene films have advantages of simple structure, light weight, and easy fabrication. The transparent and stretchable nature of carbon nanotube or graphene films also makes it possible to fabricate them into any shape and size, freestanding or on any insulating surfaces, showing great potentials to be applied for developing thermoacoustic MEMS speakers. However, current thermoacoustic speakers require large size (1–4 cm) and high power consumption (0.1–3 W) to generate adequate sound pressure output.

4. Comparison of Different MEMS Speakers

As reviewed in Sections 2 and 3, MEMS speakers have been demonstrated based on piezoelectric, electrodynamic, electrostatic, and thermoacoustic transduction mechanisms, showing great potentials for various applications including hearing instruments and portable electronic devices. Among them, piezoelectric MEMS speakers and electrodynamic MEMS speakers are the dominant types of MEMS speakers, which have been extensively studied and reported in a vast amount of the literature. Piezoelectric MEMS speakers have advantages of relatively large driving force and high sound pressure output over other MEMS speakers. High SPLs of over 90 dB have been achieved by several piezoelectric MEMS speakers either in ear simulators or in open air. Piezoelectric thin films including ZnO, AlN, PZT, and PMN-PT have been fabricated either by deposition or thinning down bulk materials and applied for fabricating piezoelectric MEMS speakers. However, most current piezoelectric MEMS speakers suffer from non-flat frequency responses due to the resonance behavior of diaphragms. The nonlinearity and hysteresis of piezoelectric materials are also drawbacks of piezoelectric MEMS speakers.

By contrast, electrodynamic MEMS speakers with quasi-linear behaviors are favorable for high-fidelity sound reconstruction. Low power consumption and large mechanical displacements are also advantages of electrodynamic MEMS speakers. Several electrodynamic MEMS speakers have been developed based on polymer diaphragms or rigid silicon diaphragms, with SPLs of 60–100 dB obtained in 2-cc couplers for in-ear applications. However, the requirement of permanent magnets for electrodynamic MEMS speakers not only increases the overall size of devices but also makes the full integration complicated and challenging.

In comparison, electrostatic MEMS speakers do not require complicated fabrication processes but suffer from small displacements, very high driving voltages, and pull-in

limitations. Several approaches, such as nanoscopic electrostatic drive (NED) technology, have been proposed to balance the driving voltage, pull-in limitation, and displacement of the diaphragm. Improved SPLs and low THDs have been obtained on these electrostatic MEMS speakers.

Compared with piezoelectric, electrodynamic, and electrostatic MEMS speakers, thermoacoustic MEMS speakers are special acoustic devices that do not rely on mechanical vibration of diaphragms to generate sounds. Therefore, there are no resonant peaks in the frequency response of thermoacoustic MEMS speakers. High transparency, high stretchability, and easy fabrication into any sizes and shapes are the advantages of thermoacoustic MEMS speakers. However, current thermoacoustic speakers all require much larger sizes to achieve comparable SPLs of piezoelectric or electrodynamic MEMS speakers. Large power consumption is another concern of thermoacoustic MEMS speakers.

In common, all these MEMS speakers are required to improve their SPLs at specific frequencies to satisfy a wider range of applications. Approaches have been proposed and demonstrated on these MEMS speakers with improved SPLs, including applying new materials and fabrication processes, designing novel structures and special electrode configurations, and using large speaker arrays.

5. Summary and Outlook

In summary, MEMS speakers have been reviewed in terms of the theory, modeling, transduction mechanisms, and development history in this article. Four types of MEMS speakers, working on piezoelectric, electrodynamic, and electrostatic actuation and the thermoacoustic effect have been introduced; their respective development milestones, performances, advantages, and limitations are also discussed. Approaches to improve the SPLs of MEMS speakers including special structures, new materials, electrode configurations, and speaker arrays are highlighted and discussed, especially for piezoelectric MEMS speakers.

In the future, the SPLs of all types of MEMS speakers will continue to be improved by the incorporation of new materials, novel fabrication techniques, and optimized device and enclosure designs, as well as with deeper understandings of their modeling. In addition to SPLs, fabrication challenges, frequency response, sound quality, and power consumption will also be taken into account. Particularly, piezoelectric MEMS speakers will be extensively investigated to obtain flat frequency responses. Electrodynamic MEMS speakers will be further studied with electroacoustic efficiency improved and permanent magnets fully integrated in batch processes. Electrostatic MEMS speakers, with efforts in reducing driving voltages, and high-level integration with electronic circuits, may find broader applications, especially in digital sound reconstruction. Finally, thermoacoustic MEMS speakers will continue to be explored with efforts to reduce the device size and power consumption. Thereby, MEMS speakers are expected to become a promising candidate not only in the in-ear applications but also in a wide range of consumer electronics.

References

1. Hwang, S.-M.; Lee, H.-J.; Hong, K.-S.; Kang, B.-S.; Hwang, G.-Y. New development of combined permanent-magnet type microspeakers used for cellular phones. *IEEE Trans. Magn.* **2005**, *41*, 2000–2003. [CrossRef]
2. Shahosseini, I.; Lefeuvre, E.; Moulin, J.; Martincic, E.; Woytasik, M.; Lemarquand, G. Optimization and microfabrication of high performance silicon-based MEMS microspeaker. *IEEE Sens. J.* **2013**, *13*, 273–284. [CrossRef]
3. Shahosseini, I.; Lefeuve, E.; Mrtincic, E.; Woytasik, M.; Moulin, J.; Megherbi, S.; Ravaud, R.; Lemarquand, G. Microstructured silicon membrane with soft suspension beams for a high performance MEMS microspeaker. *Microsyst. Technol.* **2012**, *18*, 1791–1799. [CrossRef]
4. Lee, C.-M.; Kwon, J.-H.; Kim, K.-S.; Park, J.-H.; Hwang, S.-M. Design and analysis of microspeakers to improve sound characteristics in a low frequency range. *IEEE Trans. Magn.* **2010**, *46*, 2048–2051. [CrossRef]
5. Cheng, M.C.; Huang, W.S.; Huang, S.R.S. A silicon microspeaker for hearing instruments. *J. Micromech. Microeng.* **2004**, *14*, 859–866. [CrossRef]
6. Lemarquand, G.; Ravaud, R.; Shahosseini, I.; Lemarquand, V.; Moulin, J.; Lefeuvre, E. MEMS electrodynamic loudspeakers for mobile phones. *Appl. Acoust.* **2012**, *73*, 379–385. [CrossRef]
7. Je, S.S.; Rivas, F.; Diaz, R.E.; Kwon, J.; Kim, J.; Bakkaloglu, B.; Kiaei, S.; Fellow; Chae, J. A compact and low-cost MEMS loudspeaker for digital hearing aids. *IEEE Trans. Biomed. Circuits Syst.* **2009**, *3*, 348–358. [CrossRef]
8. Roberts, R.C.; Du, J.; Ong, A.K.; Li, D.; Zorman, C.A.; Tien, N.C. Electrostatically driven touch-mode poly-SiC micro speaker. In Proceedings of the SENSORS, 2007 IEEE, Atlanta, GA, USA, 28–31 October 2007; pp. 284–287. [CrossRef]
9. Kaiser, B.; Langa, S.; Ehrig, L.; Stolz, M.; Schenk, H.; Conrad, H.; Schenk, H.; Schimmanz, K.; Schuffenhauer, D. Concept and proof for an all-silicon MEMS micro speaker utilizing air chambers. *Microsyst. Nanoeng.* **2019**, *5*, 1–11. [CrossRef] [PubMed]
10. Yi, S.H.; Kim, E.S. Micromachined piezoelectric microspeaker. *Jpn. J. Appl. Phys. Part 1 Regul. Pap. Short Notes Rev. Pap.* **2005**, *44*, 3836–3841. [CrossRef]
11. Cho, I.J.; Jang, S.; Nam, H.J. A Piezoelectrically actuated mems speaker with polyimide membrane and thin film Pb(Zr,Ti)O$_3$(PZT) actuator. *Integr. Ferroelectr.* **2009**, *105*, 27–36. [CrossRef]
12. Wang, H.; Li, M.; Yu, Y.; Chen, Z.; Ding, Y.; Jiang, H.; Xie, H. A piezoelectric MEMS loud speaker based on ceramic PZT. In Proceedings of the 2019 20th International Conference on Solid-State Sensors, Actuators and Microsystems & Eurosensors XXXIII (TRANSDUCERS & EUROSENSORS XXXIII), Berlin, Germany, 23–27 June 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 857–860.
13. Fei, W.; Zhou, J.; Guo, W. Low-voltage driven graphene foam thermoacoustic speaker. *Small* **2015**, *11*, 2252–2256. [CrossRef]
14. Suk, J.W.; Kirk, K.; Hao, Y.; Hall, N.A.; Ruoff, R.S. Thermoacoustic sound generation from monolayer graphene for transparent and flexible sound sources. *Adv. Mater.* **2012**, *24*, 6342–6347. [CrossRef] [PubMed]
15. Je, S.-S.; Wang, N.; Brown, H.C.; Arnold, D.P.; Chae, J. An electromagnetically actuated microspeaker with fully-integrated wax-bonded Nd-Fe-B micromagnets for hearing aid applications. In Proceedings of the TRANSDUCERS 2009—2009 International Solid-State Sensors, Actuators and Microsystems Conference, Denver, CO, USA, 21–25 June 2009; IEEE: Piscataway, NJ, USA, 2009; pp. 885–888.
16. Wang, H.; Chen, Z.; Xie, H. A high-SPL piezoelectric MEMS loud speaker based on thin ceramic PZT. *Sens. Actuators A Phys.* **2020**, *309*, 112018. [CrossRef]
17. Kim, H.J.; Koo, K.; Lee, S.Q.; Park, K.H.; Kim, J. High performance piezoelectric microspeakers and thin speaker array system. *ETRI J.* **2009**, *31*, 680–687. [CrossRef]
18. Stoppel, F.; Mannchen, A.; Niekiel, F.; Beer, D.; Giese, T.; Wagner, B. New integrated full-range MEMS speaker for in-ear applications. In Proceedings of the 2018 IEEE Micro Electro Mechanical Systems (MEMS), Belfast, UK, 21–25 January 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1068–1071.
19. Park, K.-H.; Jiang, Z.-X.; Hwang, S.-M. Design and Analysis of a Novel Microspeaker with Enhanced Low-Frequency SPL and Size Reduction. *Appl. Sci.* **2020**, *10*, 8902. [CrossRef]
20. Liu, W.; Huang, J.; Shen, Y.; Cheng, J. Theoretical Modeling of Piezoelectric Cantilever MEMS Loudspeakers. *Appl. Sci.* **2021**, *11*, 6323. [CrossRef]
21. Cheng, H.H.; Lo, S.C.; Huang, Z.R.; Wang, Y.J.; Wu, M.; Fang, W. On the design of piezoelectric MEMS microspeaker for the sound pressure level enhancement. *Sens. Actuators A Phys.* **2020**, *306*, 111960. [CrossRef]
22. Stoppel, F.; Eisermann, C.; Gu-Stoppel, S.; Kaden, D.; Giese, T.; Wagner, B. Novel membrane-less two-way MEMS loudspeaker based on piezoelectric dual-concentric actuators. In Proceedings of the TRANSDUCERS 2017—19th International Conference on Solid-State Sensors, Actuators and Microsystems, Kaohsiung, Taiwan, 18–22 June 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 2047–2050.
23. Kim, H.J.; Yang, W.S. The effects of electrodes patterned onto the piezoelectric thin film on frequency response characteristics of PMN-PT MEMS acoustic actuators. *J. Electroceramics* **2015**, *35*, 45–52. [CrossRef]
24. Wang, H.; Feng, P.X.L.; Xie, H. A dual-electrode MEMS speaker based on ceramic PZT with improved sound pressure level by phase tuning. In Proceedings of the 2021 IEEE 34th International Conference on Micro Electro Mechanical Systems (MEMS), Gainesville, FL, USA, 25–29 January 2021; IEEE: Piscataway, NJ, USA, 2021; Volume 2021, pp. 701–704.
25. Usound GmbH. Available online: https://www.usound.com/home/ (accessed on 13 October 2021).
26. TDK Corporation. Available online: https://www.tdk.com/en/news_center/press/20190521_01.html (accessed on 13 October 2021).

27. Audio Pixels Limited. Available online: https://www.audiopixels.com.au/index.cfm/technology/ (accessed on 13 October 2021).
28. Arioso Systems GmbH. Available online: https://arioso-systems.com/ (accessed on 13 October 2021).
29. Chiang, H.-Y.; Huang, Y.-H. Vibration and sound radiation of an electrostatic speaker based on circular diaphragm. *J. Acoust. Soc. Am.* **2015**, *137*, 1714–1721. [CrossRef]
30. Ŝvec, J.G.; Granqvist, S. Tutorial and guidelines on measurement of sound pressure level in voice and speech. *J. Speech Lang. Hear. Res.* **2018**, *61*, 441–461. [CrossRef]
31. Fedtke, T.; Grason, L. Sound level calibration: Microphones, ear simulators, couplers, and sound level meters. *Semin. Hear.* **2014**, *35*, 295–311. [CrossRef]
32. Jax, P.; Vary, P. Bandwidth extension of speech signals: A catalyst for the introduction of wideband speech coding? *IEEE Commun. Mag.* **2006**, *44*, 106–111. [CrossRef]
33. Blevins, R.D. *Formulas for Natural Frequency and Mode Shape*; Van Nostrand Reinhold Company: New York, NY, USA, 1979; ISBN 9780442207106.
34. Cheng, H.-H.; Huang, Z.-R.; Wu, M.; Fang, W. Low frequency sound pressure level improvement of piezoelectric MEMS microspeaker using novel spiral spring with dual electrode. In Proceedings of the 2019 20th International Conference on Solid-State Sensors, Actuators and Microsystems & Eurosensors XXXIII (TRANSDUCERS & EUROSENSORS XXXIII), Berlin, Germany, 23–27 June 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 2013–2016.
35. Yi, S.; Ur, S.C.; Kim, E.S. Performance of packaged piezoelectric microspeakers depending on the material properties. In Proceedings of the 2009 IEEE 22nd International Conference on Micro Electro Mechanical Systems, Sorrento, Italy, 25–29 January 2009; IEEE: Piscataway, NJ, USA, 2009; pp. 765–768.
36. Seo, K.; Park, J.; Kim, H.; Kim, D.; Ur, S.; Yi, S. Micromachined piezoelectric microspeakers fabricated with high quality AlN thin film. *Integr. Ferroelectr.* **2007**, *95*, 74–82. [CrossRef]
37. Sturtzer, E.; Shahosseini, I.; Pillonnet, G.; Lefeuvre, E.; Lemarquand, G. High fidelity microelectromechanical system electrodynamic micro-speaker characterization. *J. Appl. Phys.* **2013**, *113*, 214905. [CrossRef]
38. Garud, M.V.; Pratap, R. A novel MEMS speaker with peripheral electrostatic actuation. *J. Microelectromechanical Syst.* **2020**, *29*, 592–599. [CrossRef]
39. Zhou, Q.; Zettl, A. Electrostatic graphene loudspeaker. *Appl. Phys. Lett.* **2013**, *102*, 223109. [CrossRef]
40. Melnikov, A.; Schenk, H.A.G.; Monsalve, J.M.; Wall, F.; Stolz, M.; Mrosk, A.; Langa, S.; Kaiser, B. Coulomb-actuated microbeams revisited: Experimental and numerical modal decomposition of the saddle-node bifurcation. *Microsyst. Nanoeng.* **2021**, *7*. [CrossRef]
41. Younis, M.I.; Abdel-Rahman, E.M.; Nayfeh, A. A reduced-order model for electrically actuated microbeam-based MEMS. *J. Microelectromechanical Syst.* **2003**, *12*, 672–680. [CrossRef]
42. Xiao, L.; Chen, Z.; Feng, C.; Liu, L.; Bai, Z.-Q.; Wang, Y.; Qian, L.; Zhang, Y.; Li, Q.; Jiang, K.; et al. Flexible, Stretchable, Transparent Carbon Nanotube Thin Film Loudspeakers. *Nano Lett.* **2008**, *8*, 4539–4545. [CrossRef]
43. Neumann, J.J.; Gabriel, K.J. CMOS-MEMS membrane for audio-frequency acoustic actuation. *Sens. Actuators A Phys.* **2002**, *95*, 175–182. [CrossRef]
44. Huang, J.H.; Her, H.-C.; Shiah, Y.C.; Shin, S.-J. Electroacoustic simulation and experiment on a miniature loudspeaker for cellular phones. *J. Appl. Phys.* **2008**, *103*, 033502. [CrossRef]
45. Chang, J.R.; Wang, C.N. Acoustical analysis of enclosure design parameters for microspeaker system. *J. Mech.* **2019**, *35*, 1–12. [CrossRef]
46. Blackstock, D.T. *Fundamentals of Physical Acoustics*; Wiley: New York, NY, USA, 2000.
47. Chiang, H.-Y.; Huang, Y.-H. Experimental modeling and application of push-pull electrostatic speakers. *J. Acoust. Soc. Am.* **2019**, *146*, 2619–2631. [CrossRef] [PubMed]
48. Spitz, B.; Wall, F.; Schenk, H.; Melnikov, A.; Pufe, W. Audio-transducer for in-ear-applications based on CMOS compatible electrostatic actuators. In Proceedings of the MikroSystemTechnik Kongress, Berlin, Germany, 28–30 October 2019.
49. Sun, P.; Xu, D.P.; Hwang, S.M. Design of microspeaker module considering added stiffness. *J. Mech. Sci. Technol.* **2014**, *28*, 1623–1628. [CrossRef]
50. Kim, B.H.; Lee, H.S.; Kim, S.W.; Kang, P.; Park, Y.S. Hydrodynamic responses of a piezoelectric driven MEMS inkjet print-head. *Sens. Actuators A Phys.* **2014**, *210*, 131–140. [CrossRef]
51. Zhu, Y.; Liu, W.; Jia, K.; Liao, W.; Xie, H. A piezoelectric unimorph actuator based tip-tilt-piston micromirror with high fill factor and small tilt and lateral shift. *Sens. Actuators A Phys.* **2011**, *167*, 495–501. [CrossRef]
52. Smith, G.L.; Rudy, R.Q.; Polcawich, R.G.; DeVoe, D.L. Integrated thin-film piezoelectric traveling wave ultrasonic motors. *Sens. Actuators A Phys.* **2012**, *188*, 305–311. [CrossRef]
53. Rinaldi, M.; Zuniga, C.; Chengjie, Z.; Piazza, G. Super-high-frequency two-port AlN contour-mode resonators for RF applications. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **2010**, *57*, 38–45. [CrossRef]
54. Wang, H.; Yu, Y.; Chen, Z.; Yang, H.; Jiang, H.; Xie, H. Design and fabrication of a piezoelectric micromachined ultrasonic transducer array based on ceramic PZT. In Proceedings of the 2018 IEEE SENSORS, New Delhi, India, 28–31 October 2018; IEEE: Piscataway, NJ, USA, 2018; Volume 2018, pp. 1–4.

55. Yi, S.; Yoon, M.; Ur, S. Piezoelectric microspeakers with high compressive ZnO film and floating electrode. *J. Electroceramics* **2009**, *23*, 295–300. [CrossRef]

56. Tseng, S.-H.; Lo, S.-C.; Wang, Y.-J.; Lin, S.-W.; Wu, M.; Fang, W. Sound pressure and low frequency enhancement using novel PZT MEMS microspeaker design. In Proceedings of the 2020 IEEE 33rd International Conference on Micro Electro Mechanical Systems (MEMS), Vancouver, BC, Canada, 18–22 January 2020; IEEE: Piscataway, NJ, USA, 2020; Volume 2020, pp. 546–549.

57. Ren, T.; Zhang, L.; Liu, L.; Li, Z. Design optimization of beam-like ferroelectrics-silicon microphone and microspeaker. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **2002**, *49*, 266–270. [CrossRef]

58. Lee, S.S.; Ried, R.P.; White, R.M. Piezoelectric cantilever microphone and microspeaker. *J. Microelectromechanical Syst.* **1996**, *5*, 238–242. [CrossRef]

59. Wang, H.; Godara, M.; Chen, Z.; Xie, H. A one-step residue-free wet etching process of ceramic PZT for piezoelectric transducers. *Sens. Actuators A Phys.* **2019**, *290*, 130–136. [CrossRef] [PubMed]

60. Jung, J.K.; Lee, W.J. Dry etching characteristics of Pb(Zr,Ti)O$_3$ films in CF$_4$ and Cl$_2$/CF$_4$ inductively coupled plasmas. *Jpn. J. Appl. Phys. Part. 1 Regul. Pap. Short Notes Rev. Pap.* **2001**, *40*, 1408–1419. [CrossRef]

61. Shung, K.K.; Cannata, J.M.; Zhou, Q.F. Piezoelectric materials for high frequency medical imaging applications: A review. *J. Electroceramics* **2007**, *19*, 139–145. [CrossRef]

62. Muralt, P. Recent progress in materials issues for piezoelectric MEMS. *J. Am. Ceram. Soc.* **2008**, *91*, 1385–1396. [CrossRef]

63. Muralt, P. PZT thin films for microsensors and actuators: Where do we stand? *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **2000**, *47*, 903–915. [CrossRef]

64. Ko, S.C.; Kim, Y.C.; Lee, S.S.; Choi, S.H.; Kim, S.R. Micromachined piezoelectric membrane acoustic device. *Sens. Actuators A Phys.* **2003**, *103*, 130–134. [CrossRef]

65. Shelton, S.; Chan, M.L.; Park, H.; Horsley, D.; Boser, B.; Izyumin, I.; Przybyla, R.; Frey, T.; Judy, M.; Nunan, K.; et al. CMOS-compatible AlN piezoelectric micromachined ultrasonic transducers. In Proceedings of the IEEE International Ultrasonics Symposium, Rome, Italy, 20–23 September 2009; pp. 402–405. [CrossRef]

66. Ait Aissa, K.; Achour, A.; Camus, J.; Le Brizoual, L.; Jouan, P.-Y.; Djouadi, M.-A. Comparison of the structural properties and residual stress of AlN films deposited by dc magnetron sputtering and high power impulse magnetron sputtering at different working pressures. *Thin Solid Films* **2014**, *550*, 264–267. [CrossRef]

67. Ababneh, A.; Schmid, U.; Hernando, J.; Sánchez-Rojas, J.L.; Seidel, H. The influence of sputter deposition parameters on piezoelectric and mechanical properties of AlN thin films. *Mater. Sci. Eng. B* **2010**, *172*, 253–258. [CrossRef]

68. Lim, W.T.; Lee, C.H. Highly oriented ZnO thin films deposited on Ru/Si substrates. *Thin Solid Film.* **1999**, *353*, 12–15. [CrossRef]

69. Han, C.-H.; Kim, E.S. Parylene-diaphragm piezoelectric acoustic transducers. In Proceedings of the Proceedings IEEE Thirteenth Annual International Conference on Micro Electro Mechanical Systems, Miyazaki, Japan, 23–27 January 2000; IEEE: Piscataway, NJ, USA, 2000; pp. 148–152.

70. Zhang, Y.; Du, G.; Liu, D.; Wang, X.; Ma, Y.; Wang, J.; Yin, J.; Yang, X.; Hou, X.; Yang, S. Crystal growth of undoped ZnO films on Si substrates under different sputtering conditions. *J. Cryst. Growth* **2002**, *243*, 439–443. [CrossRef]

71. CTS Incorporation. Available online: https://www.ctscorp.com/products/piezoelectric-move-products/speakers/ (accessed on 13 October 2021).

72. Li, J.; Wang, C.; Ren, W.; Ma, J. ZnO thin film piezoelectric micromachined microphone with symmetric composite vibrating diaphragm. *Smart Mater. Struct.* **2017**, *26*, 55033. [CrossRef]

73. Watanabe, S.; Fujiu, T.; Fujii, T. Effect of poling on piezoelectric properties of lead zirconate titanate thin films formed by sputtering. *Appl. Phys. Lett.* **1995**, *66*, 1481–1483. [CrossRef]

74. Cheng, J.-R.; Zhu, W.; Li, N.; Cross, L.E. Electrical properties of sol-gel-derived Pb(Zr$_{0.52}$Ti$_{0.48}$)O$_3$ thin films on a PbTiO$_3$-coated stainless steel substrate. *Appl. Phys. Lett.* **2002**, *81*, 4805–4807. [CrossRef]

75. Tsaur, J.; Wang, Z.J.; Zhang, L.; Ichiki, M.; Wan, J.W.; Maeda, R. Preparation and application of lead zirconate titanate (PZT) films deposited by hybrid process: Sol-gel method and laser ablation. *Jpn. J. Appl. Phys. Part. 1 Regul. Pap. Short Notes Rev. Pap.* **2002**, *41*, 6664–6668. [CrossRef]

76. Moriyama, M.; Totsu, K.; Tanaka, S. Sol–gel deposition and characterization of lead zirconate titanate thin film using different commercial sols. *Sens. Mater.* **2019**, *31*, 2497–2509. [CrossRef]

77. Gao, R.; Chu, X.; Huan, Y.; Sun, Y.; Liu, J.; Wang, X.; Li, L. A study on (K, Na) NbO$_3$ based multilayer piezoelectric ceramics micro speaker. *Smart Mater. Struct.* **2014**, *23*, 105018. [CrossRef]

78. Wang, Q.; Yi, Z.; Ruan, T.; Xu, Q.; Yang, B.; Liu, J. Obtaining high SPL piezoelectric MEMS speaker via a rigid-flexible vibration coupling mechanism. *J. Microelectromechanical Syst.* **2021**, *30*, 725–732. [CrossRef]

79. Wang, Y.-J.; Lo, S.; Hsieh, M.; Wang, S.; Chen, Y.; Wu, M.; Fang, W. Multi-way in-phase/out-of-phase driving cantilever array for performance enhancement of PZT MEMS microspeaker. In Proceedings of the 2021 IEEE 34th International Conference on Micro Electro Mechanical Systems (MEMS), Gainesville, FL, USA, 25–29 January 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 83–84.

80. Diamond, B.M.; Neumann, J.J.; Gabriel, K.J. Digital sound reconstruction using arrays of CMOS-MEMS microspeakers. In Proceedings of the Technical Digest. MEMS 2002 IEEE International Conference on Micro Electro Mechanical Systems, Las Vegas, NV, USA, 24 January 2002; IEEE: Piscataway, NJ, USA, 2002; Volume 1, pp. 292–295.

81. Casset, F.; Dejaeger, R.; Laroche, B.; Desloges, B.; Leclere, Q.; Morisson, R.; Bohard, Y.; Goglio, J.P.; Escato, J.; Fanget, S. A 256 MEMS Membrane Digital Loudspeaker Array Based on PZT Actuators. *Procedia Eng.* **2015**, *120*, 49–52. [CrossRef]

82. Arevalo, A.; Conchouso, D.; Castro, D.; Kosel, J.; Foulds, I.G. Piezoelectric transducer array microspeaker. In Proceedings of the IEEE 11th Annual International Conference on Nano/Micro Engineered and Molecular Systems (NEMS), Sendai, Japan, 17–20 April 2016; pp. 180–183. [CrossRef]

83. Chen, Y.C.; Liu, W.T.; Chao, T.Y.; Cheng, Y.T. An optimized Cu-Ni nanocomposite coil for low-power electromagnetic microspeaker fabrication. In Proceedings of the TRANSDUCERS 2009—2009 International Solid-State Sensors, Actuators and Microsystems Conference, Denver, CO, USA, 21–25 June 2009; pp. 25–28. [CrossRef]

84. Shahosseini, I.; Lefeuvre, E.; Moulin, J.; Woytasik, M.; Martincic, E.; Pillonnet, G.; Lemarquand, G. Electromagnetic MEMS microspeaker for portable electronic devices. *Microsyst. Technol.* **2013**, *19*, 879–886. [CrossRef]

85. Majlis, B.Y.; Sugandi, G.; Noor, M.M. Compact electrodynamics MEMS-speaker. In Proceedings of the 2017 China Semiconductor Technology International Conference (CSTIC), Shanghai, China, 12–13 March 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 1–3.

86. Chen, Y.C.; Cheng, Y.T. A low-power milliwatt electromagnetic microspeaker using a PDMS membrane for hearing aids application. In Proceedings of the IEEE 24th International Conference on Micro Electro Mechanical Systems, Cancun, Mexico, 23–27 January 2011; pp. 1213–1216. [CrossRef]

87. Murarka, A.; Lang, J.H.; Bulovic, V. Printed membrane electrostatic MEMS microspeakers. In Proceedings of the IEEE 29th International Conference on Micro Electro Mechanical Systems (MEMS), Shanghai, China, 24–28 January 2016; pp. 1118–1121. [CrossRef]

88. Kim, H.; Astle, A.A.; Najafi, K.; Bernal, L.P.; Washabaugh, P.D.; Cheng, F. Bi-directional electrostatic micro speaker with two large-deflection flexible membranes actuated by single/dual electrodes. In Proceedings of the SENSORS, 2005 IEEE, Irvine, CA, USA, 30 October–3 November 2005; pp. 89–92. [CrossRef]

89. Bao, M.; Yang, H. Squeeze film air damping in MEMS. *Sens. Actuators A Phys.* **2007**, *136*, 3–27. [CrossRef]

90. Glacer, C.; Dehé, A.; Tumpold, D.; Laur, R. Silicon microspeaker with out-of-plane displacement. In Proceedings of the IEEE International Conference on Nano/micro Engineered & Molecular Systems, Waikiki Beach, HI, USA, 13–16 April 2014.

91. Arevalo, A.; Castro, D.; Conchouso, D.; Kosel, J.; Foulds, I.G. Digital electrostatic acoustic transducer array. In Proceedings of the IEEE 11th Annual International Conference on Nano/Micro Engineered and Molecular Systems (NEMS), Sendai, Japan, 17–20 April 2016; pp. 225–228. [CrossRef]

92. Sano, C.; Ataka, M.; Hashiguchi, G.; Toshiyoshi, H. An electret-augmented low-voltage MEMS electrostatic out-of-plane actuator for acoustic transducer applications. *Micromachines* **2020**, *11*, 267. [CrossRef] [PubMed]

93. Conrad, H.; Schenk, H.; Kaiser, B.; Langa, S.; Gaudet, M.; Schimmanz, K.; Stolz, M.; Lenz, M. A small-gap electrostatic micro-actuator for large deflections. *Nat. Commun.* **2015**, *6*, 10078. [CrossRef] [PubMed]

94. Tian, H.; Ren, T.-L.; Xie, D.; Wang, Y.-F.; Zhou, C.-J.; Feng, T.-T.; Fu, D.; Yang, Y.; Peng, P.-G.; Wang, L.-G.; et al. Graphene-on-paper sound source devices. *ACS Nano* **2011**, *5*, 4878–4885. [CrossRef] [PubMed]

95. Wang, D.; He, X.; Zhao, J.; Jin, L.; Ji, X. Research on the electrical-thermal-acoustic conversion behavior of thermoacoustic speakers based on multilayer graphene film. *IEEE Sens. J.* **2020**, *20*, 14646–14654. [CrossRef]