1 CMOS-MEMS Acoustic Devices

J.J. Neumann[1]*, K.J. Gabriel*

ECE Department
Carnegie Mellon University
USA

1.1 Introduction

Microelectromechanical Systems (MEMS) technology is an approach to fabrication that uses, as a basis, the materials and processes of microelectronics fabrication and conveys the advantages of *miniaturization, multiple components* and *on-chip signal processing* to the design and construction of integrated microstructures. Such systems are smarter, more reliable, and more capable, while also being less intrusive and less expensive than the traditional macroscopic components and systems they replace.

MEMS products are in applications ranging from acceleration sensors for automotive airbag safety systems to microoptical switches for telecommunications and micromirror arrays for data projectors and home theater systems. In this chapter, we explore the emerging application of MEMS technology to the design and fabrication of *acoustic transducers— microphone chips and speaker chips*. The advantages of MEMS— miniaturization, multiple components and on-chip signal processing— are all being employed to displace decades-old technology and bring the promise of new capabilities and features to audio devices and products.

Miniaturization and the relatively smaller size enabled by MEMS microphone chips open opportunities for broader and less-disruptive sampling of acoustic environments, ultrasonic devices where the wavelength of sound is on the same scale as geometries of the MEMS structures and repeatable sub-micron gaps between high-quality, reliable conductors and semiconductor materials bring the efficiencies, quality and scale of semiconductor fabrication to the manufacturing of audio devices.

Multiple components and the batch-fabricated, lithographic processes of MEMS fabrication bring the promise of arrays of acoustic elements, each element with precisely controlled geometries and material properties enabling extremely well-matched microphone pairs or clusters for the construction of multi-microphone features including: directional microphones, noise-suppression microphones and wind-immune microphones. Multiple and well-matched acoustic elements also lead to a completely new sound generation architecture, described at length later in this chapter, that relies on the direct and digital sound reconstruction enabled by the collective action of hundreds to thousands of individual and identical binary speakers.

On-chip signal processing is the final and integrating advantage that binds the miniature and multiple acoustic elements together to complete the integrated acoustic system-on-chip.

[1]Further author information: E-mail: jneumann@ece.cmu.edu, Telephone: 412 268 4404

Integrated analog, mixed-signal and digital circuitry provide not only the signal transduction and amplification typically needed for a sensor or actuator, but also open the door to on-chip analog-to-digital conversion and digital signal processing that will be needed to fully realize and deliver the capabilities and features of MEMS acoustic devices.

To date, most of the work reported on MEMS microphones and speakers have involved the use of piezoelectric materials and/or polymers deposited on silicon substrates[1, 2, 3]. Structures employed have included diaphragms, cantilever beams, and thermally actuated domes. Work at Carnegie Mellon University[4] has focused on building diaphragms out of the materials of standard CMOS, resulting in intrinsically integrated electronics with a simpler process flow.

A number of MEMS microphones have been built and tested using custom and captive fabrication processes usually using silicon or silicon nitride to form the diaphragm[5, 6, 7]. The approach employed by Carnegie Mellon University is to use the metal and oxide of CMOS fabrication to form a mesh diaphragm that acts as the skeleton for the deposition of a polymer that results in an airtight diaphragm[8]. A very different approach is taken with the "Microflown" which uses no diaphragm, but instead directly measures the particle velocity of air moving between two heating elements[9]. An even more radical design involves a micromachined mirror attached to a diaphragm. Modulations in the strength of the laser light reflected through a fiber optic cable are translated into an electrical signal[10].

In this chapter we discuss the design, fabrication and operation of audio-range microphones and microspeakers built using the CMOS-MEMS technology developed at Carnegie Mellon University. We will also discuss work done on a CMOS-MEMS ultrasonic sensor array designed for use in liquid using the same technology.

1.2 Microphones

A conventional microphone is made in one of several ways. One type is a dynamic microphone, which uses a coil of wire (usually embedded in the diaphragm) which moves relative to a permanent magnet when exposed to a sound wave. These are used mainly for low-budget handheld microphones, though there are notable exceptions (e.g. Shure SM series). A much broader class of microphone is the "condenser" or capacitive microphone. These are subdivided into two categories, electret, and externally polarized. In an electret microphone the electret material provides a permanent electric charge on either the diaphragm or backplate (which has holes to allow air flow). The change in capacitance C when sound impinges on the diaphragm creates a small voltage signal proportional to the displacement, which can be read with a high-impedance amplifier. In an externally polarized, or "phantom powered" microphone, the charge is provided by a voltage source through a very large resistance (such that $1/RC$ is smaller than the minimum frequency of interest). Electret microphones, which use a permanently charged electret material to polarize the capacitor, are typically used in lower-cost consumer products, and phantom powered mics are usually used in professional audio applications. Some microphones are also made with piezoelectric materials, and are sometimes used in low-cost applications, or when ruggedness is an overriding concern.

MEMS provides several advantages over conventional ways of building microphones. The most obvious are the economics of manufacturing mass quantities in the existing semicon-

ductor fabrication and packaging infrastructure, and the possibility of including integrated electronics. The smaller size also allows new applications, such as surveillance and multiple transducers in a small area. In theory, the thinner diaphragms also permit better noise, sensitivity, and vibration rejection for a given surface area than conventional microphones (See section 1.2.1). Several approaches to MEMS microphones have already been taken. In one similar to this work, additional layers (polyimide and metal) were deposited on a CMOS chip and micromachined[5]. Polysilicon[6] or silicon[7] may also be used to form the diaphragm. It is also possible, using thermal methods, to measure particle velocity (airflow) directly in three dimensions[9], rather than pressure, which greatly simplifies sound intensity measurements when used in conjunction with a single pressure microphone.

CMOS-MEMS microphones are usually made in the condenser microphone type. In the work presented here, the traditional "condenser" or capacitive microphone approach is used, but integrated with CMOS electronics in a way that minimizes custom processing steps. This results in a technology which can be mass produced commercially, while maintaining design flexibility and taking advantage of advances in semiconductor fabrication as they occur in the industry at large.

At Carnegie Mellon University (CMU), two audio-range microphone prototypes were constructed in CMOS-MEMS. In both cases, variation in the capacitance between a metal-oxide-polymer diaphragm and the silicon substrate was used to transduce sound into an electrical signal. (An ultrasonic "microphone" was also made with piezoresistors, discussed in section 1.4). In the first prototype a single diaphragm of size 700 by 700 μm acted as the transduction element. The capacitance between this diaphragm and the silicon substrate was part of an oscillator, and the frequency of the oscillator was modulated by the incoming sound. The carrier frequency of the oscillator was around 400 kHz, and the output was demodulated with a phase-locked loop. The second prototype was based on the same frequency-modulated concept, but the carrier frequency was designed to be around 100 MHz to facilitate demodulation by an FM radio receiver. This method was chosen to minimize the noise contribution from the demodulation section of the experimental setup, allowing us to better estimate the noise due to the transducer itself. The size of the entire CMOS chip was the same as the first prototype (2 mm square), but transduction was performed by six identical diaphragms, each 324 μm square. The smaller size was chosen to improve the high-end roll-off frequency over the first prototype. The performance of these microphones is shown in Section 1.2.4.

In this section, we will introduce the concepts necessary to understand the acoustics of MEMS microphones, and present the fabrication steps and results for our microphone prototypes.

1.2.1 Designing for Small Sizes

The small size of MEMS devices is attractive, yet brings up issues of physical limits and appropriate size scales for acoustic applications. Which scenarios are realistic? What physical limitations apply? Is smaller always appropriate and/or better?

MEMS microphone/speaker design involves many of the same issues as conventional microphones/speakers, but the scale difference changes their relative importance. For example, diffraction effects are still an issue for MEMS microphones, because the packaging is still macroscopic. Because the sizes of MEMS microphone diaphragms (usually 2 mm or less) are

much smaller than any audio wavelength of interest (> 17 mm), the shape of the diaphragm is not an issue; however, just like conventional microphones, diffraction effects (the increase in pressure at the face of a microphone relative to the free field) from packaging may still be the dominating effect on frequency response. Diffraction effects are ignored in this discussion. On the other hand, getting a smooth frequency response is usually easier in the case of the MEMS microphone because of the simple mechanical structure; system resonances can be designed well above the frequency range of interest, and damping can be introduced to tame the resonance of the fundamental mode of the microphone diaphragm.

Generally, one associates large-diaphragm microphones (typically 2.5 cm diameter for studio-grade mics) with low noise floors and overall better performance. What is interesting is that while a single MEMS-scale microphone element may have worse performance than a conventional-size element, for a given total area and rolloff frequency it can be shown that an array of very thin, low mass diaphragms will outperform a single large diaphragm in terms of noise floor, absolute sensitivity, and vibration rejection. We investigate each of these below.

Equivalent Input Noise

An important spec for microphones is equivalent input noise, usually given in terms of dB(A) SPL (The decibels is relative to 20 μPa sound pressure level, and the "A" modifier refers to A-weighting of the noise power spectrum). For conventional microphones, larger diaphragms correlate with lower noise floors. This is due partly to the reduced thermomechanical equivalent input noise (essentially the diaphragm interacting with the Brownian motion of individual air molecules), but also to a large extent on the increased electrical capacitance facilitating the job of the preamp electronics. In the case of the micromechanical microphone, thermomechanical noise may become greater than electronic noise if the designer is not careful about the noise of the acoustic circuit as well. At small size scales, the flow of individual air molecules impinging on the microphone diaphragm gives rise to an equivalent input pressure noise. This noise can be calculated in the same way as thermal (Johnson) noise in an electrical circuit.

In an electrical circuit, a voltage noise is generated across any resistance, as a result of the fluctuations in potential energy of electrons due to their interactions with other electrons and scatterers in the resistor. This is called the Johnson noise. From a circuit design standpoint, this can be modeled as a white noise source in series with an "ideal" (noiseless) resistor. The rms noise power density of the Johnson noise voltage v_n generated by a resistance R is $\sqrt{4k_BTR}$, where k_B is Boltzmann's constant and T is the absolute temperature[11]. Thermal noise at any point in the circuit can be calculated by substituting an ideal resistor plus a noise voltage source for every resistance in the circuit.

In analogy to electrical circuits, a "Johnson noise" is generated by the acoustic resistances. Acoustic resistances come from any dissipative mechanism associated with the propagation of sound into or through the device. Some examples relevant to microphone design are squeeze damping, vent holes, and radiation resistance. Squeeze damping is the damping due to air being squeezed outward from between two plates moving normal to their surfaces, such as a microphone diaphragm and backplate[12]. This effect is significant at small gap sizes (around 2 μm), and increases as the inverse of the cube of the gap. This is of particular concern in MEMS microphones, relative to conventional microphones. Vent holes are an important part of microphone designs, as they serve several functions: 1) they provide a path from the

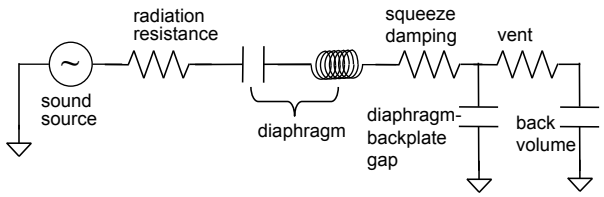

Figure 1.1: Equivalent electrical model of a typical microphone acoustic system.

backside of the diaphragm to a "back volume" which provides compliance, 2) they serve to damp the natural resonance of the diaphragm, providing a smoother frequency response, and 3) they provide a means to equilibrate the inside of the microphone with slow changes in ambient atmospheric pressure (to maintain constant gap and sensitivity, and prevent damage to the microphone). Their acoustic resistance is due to viscous (capillary) effects for very small radius tubes, and depends strongly on the radius. Radiation resistance is the dissipative force that a microphone diaphragm feels as it vibrates in a medium. It should be noted that not only does the microphone diaphragm vibrate under the influence of an incoming sound wave, but the resulting vibrations radiate new sound waves into the free field of the medium, and thus take energy away from the system. The radiation resistance is proportional to frequency squared at low frequencies (where the diaphragm size is much smaller than the wavelength), but approaches a constant (the product of the air density and sound speed, divided by area) at higher frequencies. A large radiation resistance is desirable for speakers, as it determines the amount of sound energy radiated for a given displacement of the diaphragm or speaker cone. The radiation resistance generates some noise in the MEMS microphone, but by far the largest sources are the resistances of the vent(s) and squeeze damping.

An equivalent electrical circuit is shown in Figure 1.1 for the acoustic circuit of a typical microphone. This electro-mechanical analogy is one in where sound pressure corresponds to voltage, and volume velocity (volume of air moving past a given plane during unit time) corresponds to electrical current. Other analogies are possible, and are described in detail, along with formulas for acoustic components, in other articles[13]. In any reasonable microphone design, the resistances will be kept small enough to prevent significant damping, meaning the high frequency rolloff due to damping will be located at or above the frequency where the mass

of the diaphragm (modeled by the inductor) causes a high-frequency rolloff. This allows us to make some simplifying assumptions in order to estimate the noise density in the middle of the audio range, where noise specifications are usually reported. In this limit, the total noise power is distributed uniformly across the bandwidth of the microphone[12], which we will take to be the natural resonant frequency of the diaphragm, $\omega_0 = 1/\sqrt{LC}$. This can be seen by analyzing the electrical equivalent circuit, and calculating the noise voltage (pressure) across the capacitor-inductor (compliance-mass) combination, which represents the diaphragm by itself.

There is another way to analyze the noise which provides an interesting insight: in the regime described above, it turns out the total noise (and thus the noise density, since the distribution is nearly uniform) is dependent only on the compliance of the diaphragm. What this means, is that we can to a very good approximation ignore the magnitudes of the resistances and focus on the compliance to determine the overall noise. The equipartition theorem of statistical mechanics[14] states that each quadratic term in the expression for the energy (potential and kinetic) of a system has associated with it a thermal energy $k_B T/2$ when the system is in thermal equilibrium. The displacement of the diaphragm is what is transduced into the microphone signal, so the relevant expression is

$$\frac{Cp^2}{2} = \frac{k_B T}{2}, \tag{1.1}$$

where C is the diaphragm acoustic compliance, and p is the noise pressure, in analogy to the energy stored in a capacitor, $CV^2/2$.

Consider a square diaphragm[15] with thickness t, Young's modulus E, and side length a. We can show that the thermomechanical equivalent input noise is

$$p^2{}_{diaphragm} = \frac{k_B T E t^3}{(0.61)\alpha a^6}, \tag{1.2}$$

where $\alpha = 0.0138$. This shows clearly that larger areas are better, which seems at first to be bad news for MEMS microphones. However, we can instead consider an array of diaphragms which fill an area L^2, and choose the size of the individual diaphragms to achieve a given cutoff frequency ω_0. When the signals from the individual diaphragms are averaged together, the noise power drops by a factor $N = (L/a)^2$. Then

$$p^2_{array} = \frac{p^2_{diaphragm}}{N} = k_B T \frac{E t^3}{(0.61)\alpha a^6} \frac{a^2}{L^2} = k_B T \omega_0^2 \frac{t\rho}{L^2}. \tag{1.3}$$

(This is the total noise energy; to get the noise density in the flat part of the passband, one should divide by $\omega_0/4$).

Now we see that the thermomechnical noise performance is determined by the areal density of the diaphragm, $t\rho$, that is the mass per unit area. This is a quantity usually associated with the particular technology being used, rather than the particular design (especially in MEMS). Here CMOS has a great advantage: Typical CMOS diaphragm thicknesses are on the order of 1 or 2 μm, rather than about 25 μm for a conventional microphone (the densities of the materials, usually metals, are similar).

Sensitivity and Vibration Rejection

A similar analysis can be performed that shows the sensitivity, i.e. the change in capacitance with sound pressure is:

$$\frac{dC}{p} = \frac{0.61L^2\epsilon_0\alpha}{g^2\omega_0^2 t\rho}. \tag{1.4}$$

Again, we see that it is better if the quantity $t\rho$ is kept small. It should be noted that although the ratio of the capacitance change to nominal capacitance is an obvious way to define sensitivity, making dC larger facilitates making a lower equivalent input noise for the electronics as well. In MEMS technologies, the gaps can be made much smaller than in conventionally assembled microphones, so this can make up for the smaller electrode area.

The vibration rejection, which compares the diaphragm displacement due to sound pressure to displacement due to inertial effects (e.g. bumping the microphone), also improves by using diaphragms of small areal density. The following relation expresses the ratio of diaphragm displacement due to sound pressure p to the displacement due to an acceleration $\ddot{x}$:

$$\frac{x_p}{x_{accel}} = \frac{p}{\ddot{x}}\frac{1}{t\rho}. \tag{1.5}$$

1.2.2 Microphone Design and Fabrication

The microphone diaphragms are formed from the existing CMOS layers using a variant of the CMOS-based micromachining technique developed at Carnegie Mellon University[16, 4, 8]. A serpentine metal and oxide mesh pattern (0.9 μm-wide beams and gaps) is repeated to form the diaphragm skeleton, and the underlying silicon is etched out to form a suspended mesh skeleton. A Teflon®-like conformal polymer (0.5-1.0 μm) is then deposited onto the chip, covering the mesh and creating an airtight seal over a cavity. Depending on the diaphragm geometry and gap between the diaphragm and substrate, the capacitance will typically range around 0.1 pF to 1 pF.

Vent holes are needed to allow air to flow from the sub-diaphragm cavity to a back volume (which reduces the total impedance of the system), and to provide a mechanism for controlled damping of resonant oscillations. Before the release of the CMOS structures on the front side, the backside vent holes are deep reactive ion etched (DRIE) using either a timed etch, or the oxide layer as an etch-stop. Vent hole sizes are chosen large enough to decrease the acoustic resistance, and positioned close enough together to reduce squeeze damping in the diaphragm-substrate gap. After the vent holes are etched, processing of the CMOS (front) side of the chip follows the steps shown in Figure 1.2:

1. The chip comes from a CMOS fabrication foundry covered with a layer of protective glass (silicon dioxide).

2. The glass is etched anisotropically down to the silicon substrate, the metal layers acting as a mask to define the mesh structure.

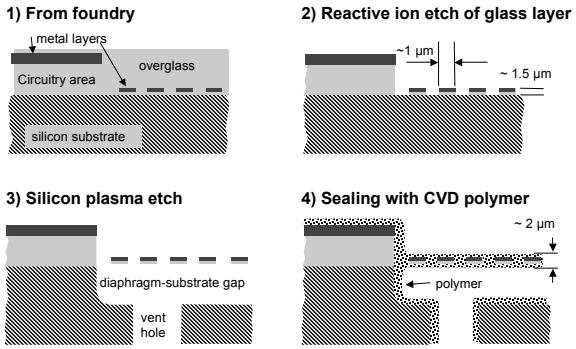

Figure 1.2: Sequence of steps to form CMOS-MEMS diaphragms.

3. The underlying silicon substrate is etched with a 12 minute DRIE anisotropic deep etch followed by a 7 minute isotropic etch. After this, the mesh structure is released from the underlying silicon and there is a gap between the diaphragm and the substrate. In the figure, we see a CMOS MEMS beam and the metal layers inside which can be used as electrodes for sensing and actuation, or wires for connecting to the on-chip circuitry.

4. In the final step, the released CMOS MEMS structure is coated with polymer in a chemical vapor deposition process. The polymer conforms to all sides of the beams, until the gaps are sealed, creating an airtight diaphragm suspended over the gap.

One difficulty in creating a large diaphragm with CMOS MEMS is the buckling caused by stresses inherent in the oxide and metal, as well as from bimorph temperature-dependent stress differences[17]. For a 50 μm-long cantilever beam made of metal and oxide, the out-of-plane curl can be as much as 1 μm, increasing with the square of the length. To solve this problem, we developed the serpentine mesh design shown in Figure 1.3. This design works as follows. Serpentine springs run in both the x and y directions, and are arranged in such a way as to cover the entire area. The x and y springs connect and cross over in the center of each unit cell. Small tabs are added near the end of the long beams (running along the boundaries of the unit cells) in order to partially close the remaining gaps, making the gaps uniform for conformal coating in a later step. The springs provide a great deal of stress relief for either compressional or tensile stress. Treating the mesh as a lumped-parameter equivalent plate, the "effective Young's modulus" for flexural and torsional stress is about nL/b and $nL/12b$ times less, respectively, than it would be for a solid plate of the same material and thickness, where n is the number of legs in each spring section, b is the lateral width of the beams, and L is the length of the spring members. The individual members which make up the springs are kept

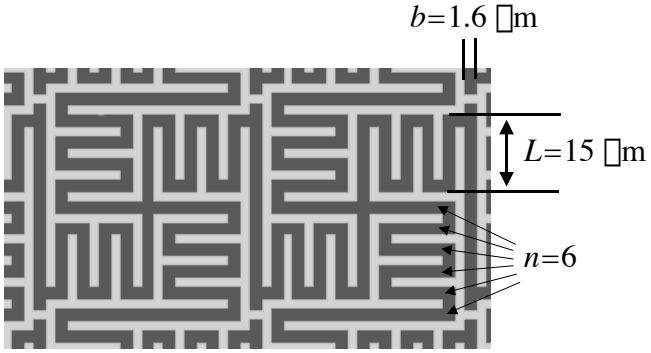

Figure 1.3: Drawing of a typical serpentine mesh pattern. Dark areas are beams, light areas are gaps.

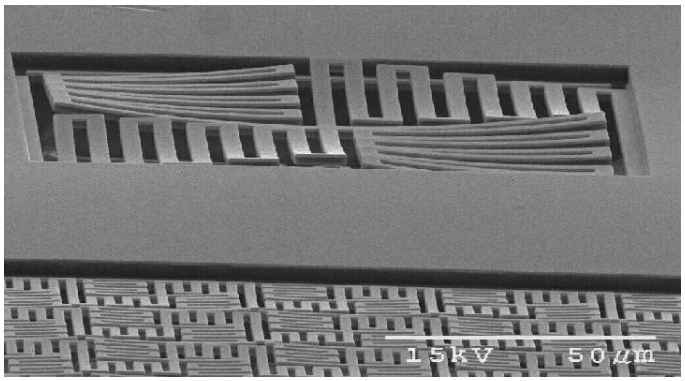

Figure 1.4: SEM of serpentine mesh unit cell.

short ($<$50 μm) so as to limit curl. It can be seen in the scanning electron micrograph (SEM) (Figure 1.4) that the greatest contribution to the curl in the mesh comes from the long beams, and that the mesh designs with shorter beams curl much less (bottom part of picture).

1.2.3 Acoustic Model of the CMOS MEMS Microphone

The transduction behavior may be predicted by considering the acoustic and the electrical behavior of the chip and package. Figure 1.5 shows a cross section of the chip mounted on a dual in-line package, along with the corresponding equivalent electrical circuit. A hole is drilled in

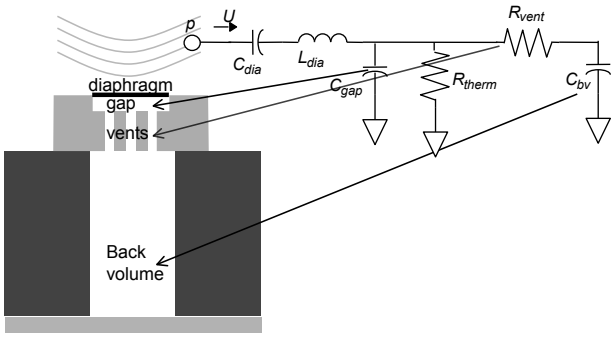

Figure 1.5: Equivalent electrical model of microphone acoustic system.

the DIP and a metal cover is glue over the bottom to the hole to create an isolated back volume, to which the vent holes in the chip connect. It is permissible to treat the acoustic behavior with a lumped-parameter equivalent electrical circuit because the acoustic wavelengths of interest ($>$17 mm) are much greater than the dimensions of the microphone and the back volume. In this analogy[13, 18], the sound pressure p (deviation from ambient atmospheric pressure) is analogous to voltage in the electrical equivalent circuit, and the volume velocity U (volume of air moved per unit time) corresponds to electrical current. Thus, acoustic impedance is defined to be $Z = p/U$. Acoustic compliances are modeled as capacitors, usually with one end tied to ground (the exception is modeling the diaphragm, as there is air flow on both sides). The compliance of the air in the diaphragm-substrate gap is given by $C_{gap} = (7/5)V/\rho c^2$, where V is the volume of the cavity, ρ is the density of air, and c is the sound speed of the air. For macroscopic air volumes, the $7/5$ factor is omitted, as the compression at audio frequencies is nearly adiabatic. At the small size scales of MEMS microphones, the area in contact with the thermally conductive silicon is high compared to the air volume, and it can be shown that the air undergoes an almost perfectly isothermal compression and expansion, which increases the compliance by the $7/5$ factor (and also introduces a negligible amount of damping). Note that one end of the capacitor in the electrical equivalent circuit represents the pressure in the cavity, and the other (grounded) terminal represents the wall of the cavity. The acoustic compliance of the diaphragm can be calculated based on formulas for deflection of a plate, and the definition of acoustic impedance.

The mass of the air and the diaphragm are modeled as inductances (“inertance” in acoustic jargon). For a diaphragm of uniform thickness, the inductance $L = \rho' t/A$ where ρ' is the material density and A is the diaphragm area. For the vent holes, we can calculate the

inductance as $L = \rho l/A$, where l is the length of the vent hole and A is the cross-sectional area. However, at MEMS scales and audio frequencies the impedance due to this inertance is negligible compared to the acoustic resistance.

There are several sources of damping (acoustic resistance). First, squeeze damping occurs between the diaphragm and the substrate. This is due to the viscosity of the air creating friction as the air is pushed sideways towards the vent holes, and is sensitive to the gap (inverse cube dependence) between the diaphragm and the cavity bottom. In the particular microphone design we investigated, this effect was dominated by the resistance of the vent holes, down to a gap of about 2 μm. The resistance of the vent holes also stems from viscosity, and can be calculated from formulas for capillary resistance. Typical dimensions of the vent holes in our microphone prototypes were diameters between 32 and 75 μm, and lengths from 200 to 600 μm. As mentioned earlier in this section, there is a very slight damping effect from heat loss through the silicon walls of the sub-diaphragm cavity. Another small source of damping occurs due to radiation resistance of the air in front of the microphone. This occurs because vibrations of the microphone diaphragm, due to incoming sound, radiate sound back into the air.

Because the device and packaging are both much smaller than most acoustic wavelengths of interest, we may neglect the pressure-doubling effect caused by diffraction. Material properties of the diaphragm derived in previous work[4] were used. We used an effective Young's modulus $E = 800$ MPa, and a density $\rho' = 1400$ kg/m^3 for the mesh/polymer combination. It can be shown that the dissipative part of the compression is negligible in the limits of pure adiabatic and pure isothermal compression, so we set R_{therm} to infinity. C_{bv} is the compliance of the back volume, in our case several mm^3 drilled in the DIP package. This compression is approximately adiabatic because of the smaller surface to volume ratio. For a given sound pressure p, the change in (electrical) capacitance may be calculated by finding the volume velocity U and deriving the diaphragm displacement. The predicted response, assuming the above material properties, is shown together with the measured response in the results section.

1.2.4 Experimental Results

The frequency response of the MEMS microphone was measured by placing it into an anechoic box (Bruel and Kjaer (B&K) 4232) and measuring its output as a loudspeaker in the box was driven over the frequency range of interest by a function generator and power amplifier. The box also contained a reference microphone (B&K 4939) placed at a symmetrical position with respect to the speakers, to measure the sound pressure level (SPL) at the microphone position. It should be noted that at higher frequencies (> 5 kHz) the sound pressure field becomes complicated due to the wavelength being smaller than the size of the speaker, but comparable to the distance between the microphones. Thus the difference in position between the reference microphone and the device under test (DUT) may cause significant variations in the measured response. The reference microphone signal was fed through a B&K 2669 preamp and B&K 2690 Nexus conditioning amp. The electrical output of the MEMS device was connected, in the case of the prototype with 400 kHz output ("P1"), to a phase-locked loop frequency-to-voltage demodulator. In the case of the prototype with 100 MHz output ("P2"), the MEMS device was connected to the antenna input of a stereo receiver (Pioneer SX-303R) through a resistor and capacitor in series, chosen to protect the output of the MEMS device.

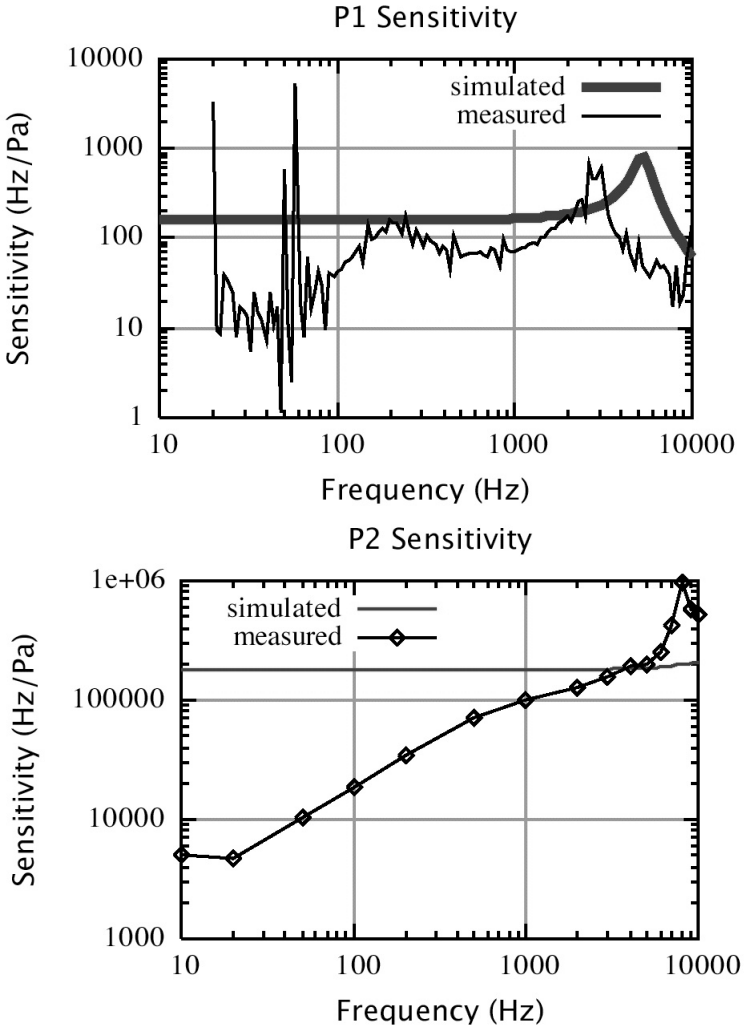

Figure 1.6: Predicted and measured frequency response of the two prototype microphones.

The frequency to volts factor of the radio was measured to be 10 μV/Hz, and a standard FM radio pre-emphasis curve was applied[19]. The signals from the MEMS device under test and the reference microphone were measured simultaneously with a signal analyzer (HP 3562A).

Figure 1.6 shows the output of the microphones in response to the loudspeaker being driven from 20 Hz to 10 kHz, along with the predicted sensitivties based on the above-discussed acoustic models. The measured response of P1 matched the predicted response well, considering the imprecision with which the material properties were known. The material properties were estimated in previous work[4], in which the resonant peak was somewhat ambiguous because of damping and resonances from the packaging. We used other approxi-

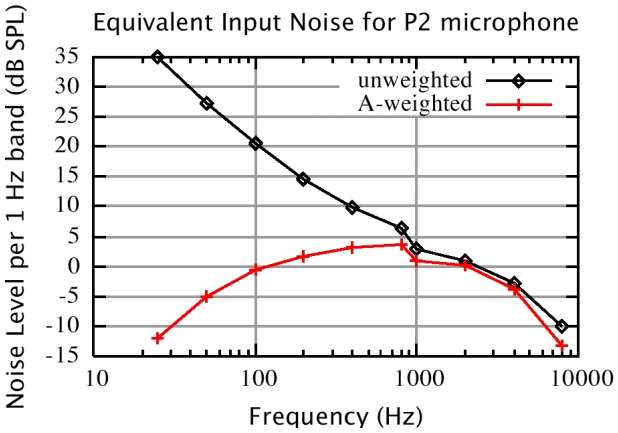

Figure 1.7: Noise power spectrum of microphone.

mations which may have an effect on the acoustic parameters, such as treating the diaphragm as a flat plate (there is actually several microns of buckling), and approximating the capacitance as between two flat plates. Another feature of the response curves is the drop-off at low frequencies, which is not predicted by the acoustic model, unless a leakage path is added. We believe this may be due to cracks in the sealing polymer, or holes which may be etched through the silicon but which were not sealed. A blanket mask with a grid of holes was used to pattern the vents, and in some cases the timed etch reached the upper surface of the CMOS chip, creating leaks. An attempt was made to seal these holes, but we were unable (in P2) to cure this problem completely.

The noise level of the P2 microphone was measured by repeating the above experiment without driving the loudspeaker. Figure 1.7 shows the equivalent input noise power density of the P2 MEMS microphone, unweighted, and also A-weighted (A-weighting is a standard function that takes into account the frequency response of the human ear[20]). The noise power spectrum takes into account the frequency response of the microphone, however the uncorrected (output) noise spectrum still has a $1/f$ shape, indicating the source of the noise is electronic rather than thermomechanical. In any case the measured noise (46 dB SPL, A-weighted) was far above the thermomechanical limit, which can be calculated from Equation (1.2.1) to be -15 dB SPL in a 1 Hz band around 1000 Hz (or 30 dB SPL over the whole bandwidth). Possible sources of electronic noise in our device include $1/f$ noise from the transistors, Johnson noise, and sensitivity to the power supply voltage (the oscillator frequency was observed to be proportional to the power supply voltage). An Agilent E3631A power supply with heavy RC filtering was used to power the device at 5 volts.

The frequency response of the microphone was measured before and after exposure of the P2 microphone to various elevated temperatures. (Figure 1.8). This is an important issue for commercially viable microphones, as the commonly used electret microphones lose their charge and ability to function after exposure to high temperatures, e.g. sitting on a hot dashboard. Because the CMOS-MEMS microphone does not use electret to polarize the capacitor, this is not an issue. However, we wanted to make sure that the mechanical properties of the

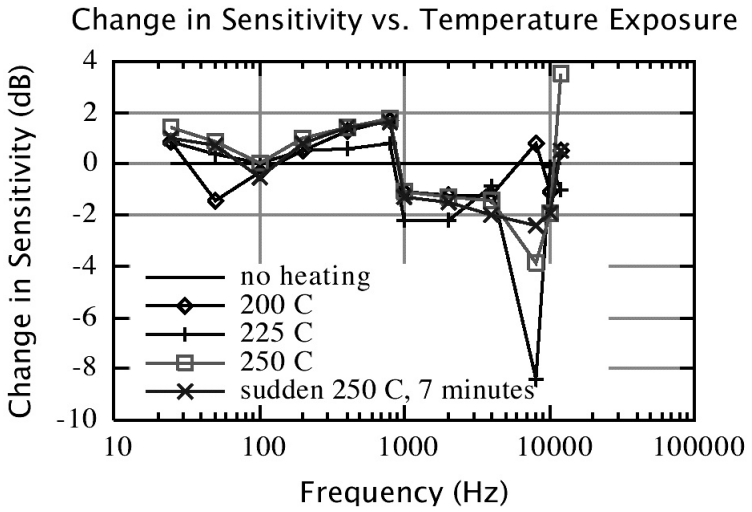

Figure 1.8: Change in sensitivity as a function of frequency, for several temperatures.

sealing polymer are not significantly affected, so we exposed the microphones to high temperatures and measured their sensitivities afterward. For all but one measurement, the packaged chip was placed in a room temperature (20 ^{o}C) oven, and the temperature was gradually increased over 30 minutes until the temperature noted on the graph was reached. Then the heating element was turned off and the oven was allowed to cool slowly (another 30 minutes) down to room temperature. The chip was taken out of the oven and the frequency response was measured. The chip was raised to a number of peak temperatures in this manner. Finally, the oven was pre-heated to 250 ^{o}C, and the packaged chip was placed in for 7 minutes, and then cooled by setting it out in the room temperature air. The measured change in frequency response relative to the unheated microphone is shown in Figure 1.8. The behavior of the frequency response with respect to temperature changes appears small and random, after the initial heating cycle to 200 ^{o}C. This suggests an initial "heat curing" process would guarantee repeatable performance through future heating cycles.

It appears that with some work the performance of a CMOS-MEMS microphone could reach a level sufficient for cell phones, hearing aids and other applications. Obvious improvements include covering a larger area with diaphragms (to increase capacitance) and develop low-noise circuitry that does not depend on frequency modulation. The limits of performance appear to be related to the quality of the circuit design at least as much as physical limits of the technology.

1.3 Speakers

At CMU, microspeakers for the audible frequency range were fabricated in CMOS-MEMS. MEMS speakers face greater challenges than microphones because of their small size. As will be explained shortly, the strategy for producing a flat frequency response is different for

Figure 1.9: Photographs of earphone construction.

MEMS speakers versus conventional speakers. We will see that despite the challenges of building a MEMS speaker, there are possibilities that will motivate us.

The first CMU microspeaker device, shown in Figure 1.9, was packaged as an earphone. This allowed measurements in a standard size "ear simulator" measurement setup (Brüel and Kjaer 4157), and also made a demo that could be used with a portable CD player. In this section, we will discuss the acoustics of the CMOS-MEMS earphone, and show the fabrication steps. The experimental and music demonstration setups will be presented. We will discuss the performance of small speakers in free space. Finally, we will consider the possibilities of arrays of microspeakers, for example, digital sound reconstruction.

1.3.1 Traditional Speaker versus MEMS Speaker

Traditional speakers work as they do due to a fortuitous combination of speaker mechanics and acoustics. The pressure generated in the air near the cone is proportional to the velocity of the cone, and the proportionality constant is the acoustic impedance of the air, more exactly the radiation impedance, which is a function of the ratio of the wavelength to the cone size. For calculating the sound energy radiated at a distance from the speaker, we use the real, or resistive part of the radiation impedance, the radiation resistance. The radiation resistance approaches the characteristic impedance of the air multiplied by the area of the cone as the wavelength gets small compared to the speaker cone. For much of the audio range, however, the wavelengths are about the size of the cone or longer and the radiation resistance is roughly proportional to the frequency[21]. Therefore, in order to have a flat response, it is necessary that the speaker be designed so that its range of operation is above the resonant frequency of

the cone, so that the velocity of the cone for a given voltage is inversely proportional to the frequency (mass-controlled rather than stiffness controlled).

The MEMS speaker is much smaller than the wavelengths it is trying to excite, and in this regime the radiation resistance is proportional to the square of the frequency[21]. In addition, because its resonance frequency is designed above the audio range, the displacement is proportional to the voltage, i.e. stiffness controlled (assuming electrostatic deflection with signal plus DC bias). This results in a free-field response that goes as frequency cubed, which is not acceptable for most applications. Therefore, MEMS speakers may be useful mainly for in-ear applications such as hearing aids and portable music devices. Out-of-ear applications may still be possible using very large arrays of microspeakers, and signal processing to achieve the desired frequency response.

1.3.2 Fabrication

Most of the processing steps for fabricating the CMOS-MEMS microspeaker are the same as for the CMOS-MEMS microphone. The main difference is in the packaging, and the choice of geometric design parameters. The microspeaker chip contained a central diaphragm, 1442 μm square, surrounded by 27 smaller variations, structures to test the effect of varying the serpentine mesh design parameters (gaps, beam widths, and number of turns). Only the large central diaphragm was electrically connected to the bond pads, and was actuated[4]. Vent holes about 150 μm across were etched through the silicon substrate from the back, in a 3 by 3 grid behind the diaphragm, using the bottom glass layer as an etch stop. The chips were carefully cleaned to remove photoresist, and then the usual CMOS-MEMS processing[16] was performed to release the mechanical structures. Finally, a chemical vapor deposition of polymer was applied to seal the diaphragm. The gap between the microspeaker diaphragm and the silicon substrate was typically 60-80 μm deep, which is significantly greater than the gap in the microphones. This is so the diaphragm had room enough for the large deflections necessary for generating large sound pressures. A DC bias voltage of 67 volts was necessary to pull the diaphragm down to its operating position. This was chosen to be where the deflection vs. voltage curve has the greatest slope, in order to maximize the sensitivity to the signal voltage (see Figure 1.11).

The MEMS speaker device, in its housing, is placed in the ear with the diaphragm facing into the ear canal. The vent holes between the diaphragm-substrate gap and the back volume have important effects on the behavior of the device, such as reducing the acoustic impedance on the backside of the diaphragm (allowing greater displacements for a given electric force), and adding a resistive component that damps out the unwanted resonances. The backside of the chip needs to make contact to the outside world (or at least a significant volume of air), without creating an acoustic “short circuit” between the back side of the diaphragm and the side facing into the ear canal. In equivalent electrical terms, the backside of the diaphragm (capacitor-inductor combination) should have a path to ground, either directly or via a large capacitance. This can be accomplished by mounting the CMOS die on a substrate with a hole in it, and sealing around the edge of the chip with epoxy, as shown in Figure 1.10. The backside of the substrate is then facing an enclosed, but relatively (compared to the dimensions of the CMOS-MEMS acoustic system) large air volume.

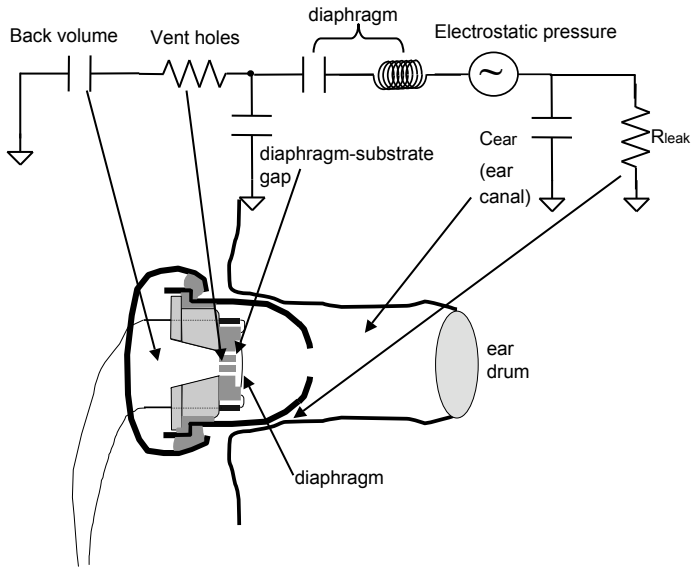

Figure 1.10: Electrical equivalent model of earphone acoustic system.

1.3.3 Acoustics in a closed coupler

A simplied acoustic model was developed to investigate the effects of the design parameters on the behavior of the MEMS earphone (Figure 1.10). Like the situation with the CMOS-MEMS microphone (Section 1.2.3), most of the wavelengths of interest (audio frequencies) are much greater ($\gg$17 mm) than the size of both the device and the ear canal, so we may ignore the effects of wave propagation and model the system as discrete components (the sound pressure is assumed uniform throughout each of the air volumes). The signal voltage source in series with the diaphragm models the AC signal part of the electrostatic pressure created by the driving voltage. The voltages in the equivalent acoustic model correspond to sound pressures, and currents correspond to volume velocity of air. The voltage on C_{ear} represents the sound pressure created in the ear canal. It is clear from the equivalent electrical circuit that as the volume of air that the earphone drives gets smaller, the pressure will increase. It is also clear that the acoustic impedance of the back volume should be kept smaller than or comparable to that of the diaphragm to get good performance. In our prototype CMOS-MEMS earphone, the acoustic compliance of the diaphragm was estimated to be 5.5×10^{-6} s^2·cm^4/g, which corresponds to an air volume 7.7 cm^3. This volume is several times larger than what

is available in the earphone shell, so there is some attenuation. The acoustic resistances and compliances are calculated from the device geometry and the air properties:

$$R = \frac{8l\mu}{\pi r^4}, \qquad C = \frac{V}{\rho c^2}, \tag{1.6}$$

where μ is the viscosity of air, l is the length of the vent hole, and r is the radius of the vent hole. V is the volume of air trapped between the diaphragm and substrate, ρ is the density of air, and c is the sound speed of air at standard temperature and pressure. It was assumed that when the device is placed in the ear canal, there would be roughly 1 cm^3 of volume, with some leakage (modeled by R_{leak}) around the device into the outside world.

The serpentine mesh structure is too complicated for which to derive an exact analytical model, but with some simplifying assumptions (e.g. considering the flexure, but not the torsion of the individual beams[22]), it was possible to show that the effective stiffness is approximately 5-60 times less than a solid sheet of the same thickness and material, depending on whether flexural or torsional stresses dominate. Thus, for purposes of acoustic modeling we used as a starting point the properties of polymer, as it has a density similar to the aluminum-glass beams, and is the compliant material filling the gaps between beams (in other words, we expect the polymer to set the upper limit on the stiffness). Using the results from our earphone and microphone experiments[4, 8], we were able to estimate values for the effective density and Young's modulus of the coated diaphragms.

1.3.4 Results

Optical measurements were made of static deflection as a function of voltage (Figure 1.11) in order to determine an operating voltage, in this case 67 volts. The deflection shown in the graph is given relative to the zero-volt position of the diaphragm, which may be several microns above the chip surface because of the buckling or bulging mentioned above.

The CMOS-MEMS speaker chip was packaged in a conventional earphone shell for testing. First the chip was mounted on a small TO package and wire bonded. The TO package was then epoxied inside the housing from a Radio Shack 33-175B earphone in such a way as to seal off the front of the TO package from the back (Figure 1.10). To present an acoustic load to the device similar to a human ear, we measured the acoustic behavior in a Brüel and Kjaer (B&K) 4157 Ear Simulator rather than at an arbitrary distance from a free-field reference microphone. The earphone and measurement microphone were put inside a B&K 4232 anechoic test chamber that provided about 40-50 dB isolation.

The response of the MEMS device mounted in the earphone housing was measured, driven with a 14.3 V-peak signal on the top of a 67 VDC bias (Fig. 1.12). We were able to determine, by comparing results with a conventional earphone, that the resonant peaks near 4 kHz were from the earphone housing and ear simulator, and not from the earphone diaphragm itself.[4]. The peaks near 12 kHz are due to the frequency response of the microphone (ear simulator) itself, which mimic the resonance of the ear canal. This peak's center frequency is sensitive to the trapped air volume between the earphone diaphragm and the ear simulator's microphone, so it may not coincide with the peak in the calibration data supplied with the microphone. For this reason, we have chosen to present the data without correction for the microphone response to avoid misinterpretation.

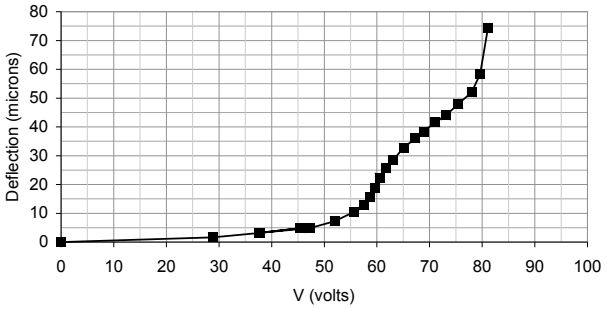

Figure 1.11: Diaphragm deflection as a function of voltage, for a 1442 μm square diaphragm and 60 μm diaphragm-substrate gap.

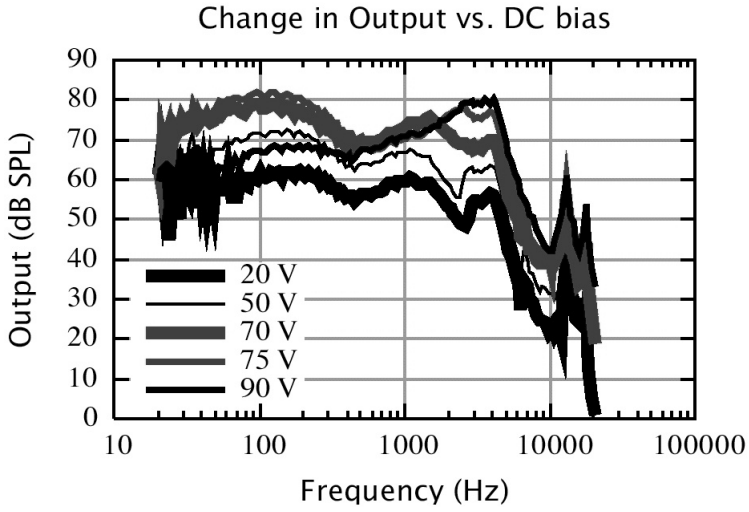

Figure 1.12: Effect of DC bias voltage on frequency response of CMOS-MEMS earphone.

Fig. 1.12 illustrates the effect of varying bias on the frequency response. The overall magnitude of the response increases rapidly as the bias voltage is increased from 20 V to about 70 V. At around 75 V bias, a qualitative change in behavior occurs, with a broad peak

forming between 2 and 4 kHz. Further increase of the bias to 80 V completes the transition, with a sudden drop in the low-frequency (20-200 Hz) response of about 15 dB, while the 3 kHz peak remains. Though we did not measure distortion, we heard a clear increase in distortion at the higher input signal levels in this region of operation, as well as the drop-off in bass response. We believe that this change in behavior corresponds to the snapping down of the diaphragm between its two bistable states (slightly convex and slightly concave), which occurs around 75-80 V in most of the chips we studied.

Simulation results (Figure 1.13) were compared with the experimental data to help us to estimate the polymer material properties. The set of known quantities for the device include

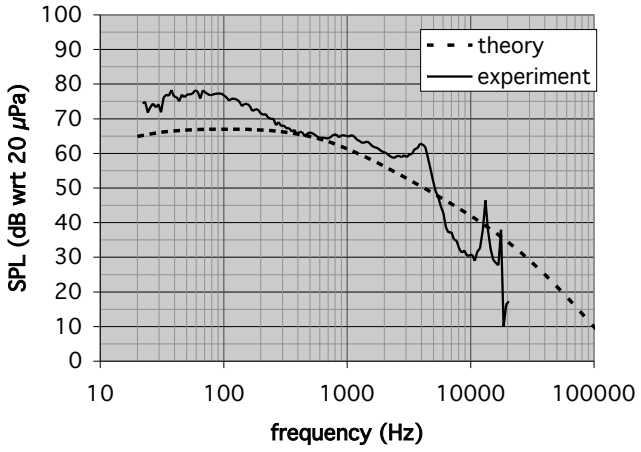

Figure 1.13: Predicted earphone response (best fit of material property parameters).

the etch depth (70 μm), diaphragm size (1442 μm^2), applied voltage (67 V bias + 9.75 Vrms signal), approximate diaphragm thickness (2 μm), and size (150 μm across) and length (about 500 μm) of the nine vent holes. Other knowns include the density of air and the speed of sound. It was estimated that the earphone housing fills about half of the volume of a typical ear canal or our ear simulator, so we use 1 cm^3 for the ear canal volume. Unknown quantities include the material properties of the polymer, and the leak size around the perimeter of the device. As initial trial values, we used a density of 1.4×10^3 ng/μm^3 for the polymer density and 3000 MPa for the Young's modulus. We assumed that the observed peaks near 4 and

12 kHz are artifacts of the earphone housing and microphone response (as explained earlier), so we did not try to match these. For the leak area, we started with no leak and increased the leak area. The density and Young's modulus of the diaphragm were adjusted in order to bring the simulation result in line with the experimental data.

Figure 1.13 illustrates the results of the model calculations. Starting with the parameters described above, we adjusted the unknowns to make a reasonable fit to the data. First, the leak area around the device appears to be very small, or zero. This is to be expected since the earphone shell fits tightly into the rubber gasket supplied with the ear simulator. Second, the corner frequency of the measured earphones is ambiguous, especially given the fact that the housing itself has a significant effect on the frequency response. A range of values for both the density and the Young's modulus were simulated[4], and it was found that the earphone design was relatively insensitive to the specific material property values, probably in part because the acoustic impedance of the diaphragm was small compared to other parts of the system (back volume, ear canal, and vent holes). Later work with CMOS-MEMS microphones refined our estimates to 800 MPa for the Young's modulus, and 1.4×10^{-3} ng/μm^3 for the density.

1.3.5 Digital Sound Reconstruction

In Section 1.3.1 we described the strengths and weaknesses of the CMOS-MEMS speakers in relation to conventional loudspeakers. With a conventional loudspeaker, great care is taken with materials and design to create speakers with both power and linearity. We have established that any single-element MEMS-scale transducer will be relatively ineffective as a radiator at audio frequencies because of their small size (and hence acoustic coupling to air) and small displacements. At CMU, we have demonstrated a different approach to creating sound, which we call "Digital Sound Reconstruction" (DSR)[23]. In this approach, an array of transducers, each of limited power and linearity, together overcome both of these problems. Before MEMS, this was not a reasonable approach, because of the high assembly costs and the lack of uniformity between individual transducers. However, CMOS-MEMS micromachining provides a method for easily creating large arrays, and the uniformity is a result of using state of the art CMOS foundries to perform the most critical steps.

Consider a single CMOS-MEMS speaker as described in the earlier sections. The speaker size can be scaled down until the frequency response of the speaker covers the entire audio band (20 kHz). If we take an array of such speakers (call them "speaklets") we can then control the amplitude of the volume velocity by controlling the number of speaklets that snap up or down. Note that regardless of the linearity of an individual speaklet, as long as the speaklets have identical responses a linear overall response is achieved as the pressures add (ignoring the effect of phase cancellation due to different path lengths). Note also that frequency response of the individual speaklets becomes irrelevant below their resonant frequency. If one conceives of each speaklet having sequence of "up" and "down" states at a given sampling rate, then we have a digital speaker system.

In the past, there have been ideas put forth for building digital speakers[24] or digital speaker arrays[25], but there have been limitations with conventional technology. If one takes a single speaker, and subdivides its speaker coil[24], it is difficult to achieve the exact power-of-two ratios that are necessary for linearity. Similarly, if one builds an array with conventional transducers, mismatch is also a problem, and the large overall size causes phase cancellations

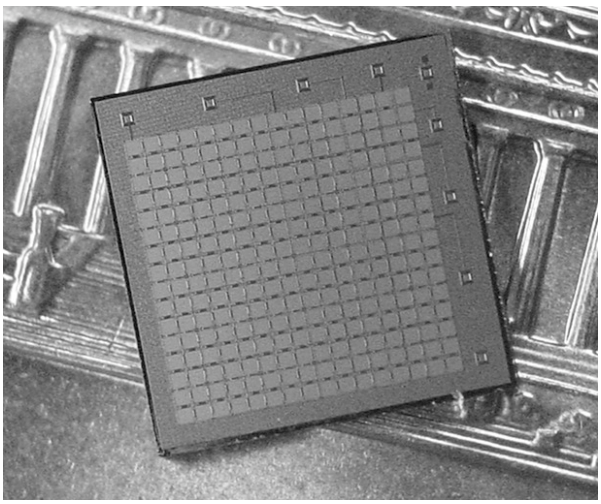

Figure 1.14: Photograph of DSR chip fabricated in CMOS.

due to the path differences being comparable to the wavelengths of interest (not to mention the nightmare of interconnecting such a system). MEMS can solve the size and uniformity problem; CMOS-MEMS in particular facilitates interconnect and any control electronics that are necessary (for example converting an N-bit digital word into 2^N control signals).

As proof of concept of this technology, we built[23] an array of 16×16 speaklets, each 216 μm square, on a CMOS-MEMS chip, using the AMS (Austria Micro Systems) process offered by MOSIS (Metal-Oxide-Semiconductor Implementation System, Marina del Rey, CA). A photograph of the chip is shown in Figure 1.14. Experiments were performed to verify the additive nature of the pressure pulses[26]. The impulses added as expected, however the response was slower than desired. Faster responses can be obtained with stiffer diaphragms and higher voltages and/or pulse shaping. Still, the technology works well enough that one can demonstrate a successive approximation to a low-frequency (500 Hz) sine wave, by adding in successively less significant bits (Figure 1.15).

As mentioned earlier, the frequency response from the system driven in a digital manner should be independent (up to the cut off frequency) of the frequency response of the individual speaklets. Figure 1.16 shows the results of driving the array digitally (speaklets snap up and down individually in numbers proportional to targeted sound pressure) and the result of driving all the speaklets in the array together in an analog manner (voltage signal proportional to targeted sound pressure). While there is some deviation from flatness for the digital mode of driving the array, this is mainly due to the response of the ear simulator. As a final note, it should be mentioned that this array was easily audible from a distance of a meter or two in a noisy lab environment.

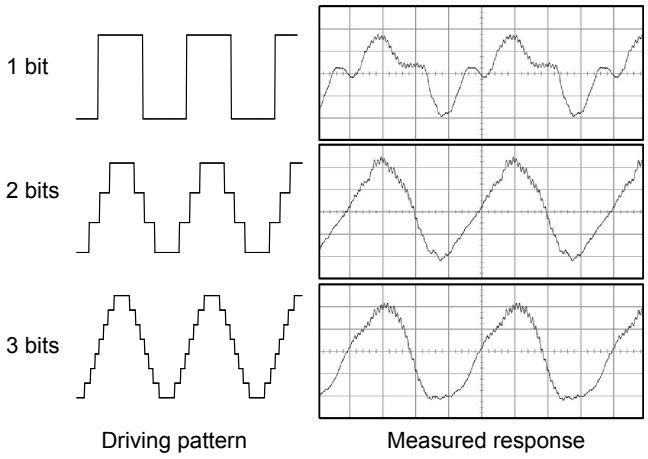

Figure 1.15: Successive approximation to a 500 Hz sine wave, by increasing the number of bits from 1 to 3.

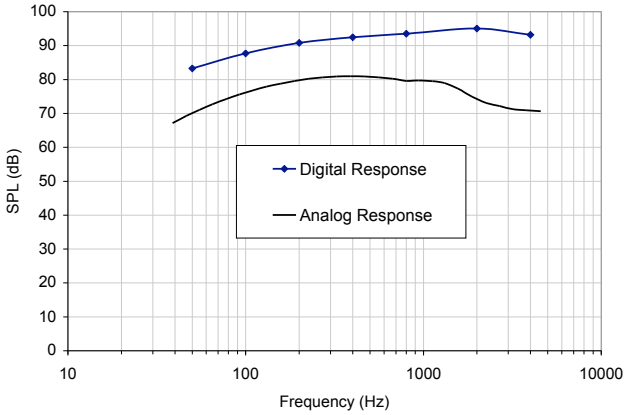

Figure 1.16: Frequency response of DSR chip, driven digitally vs. analog driven.

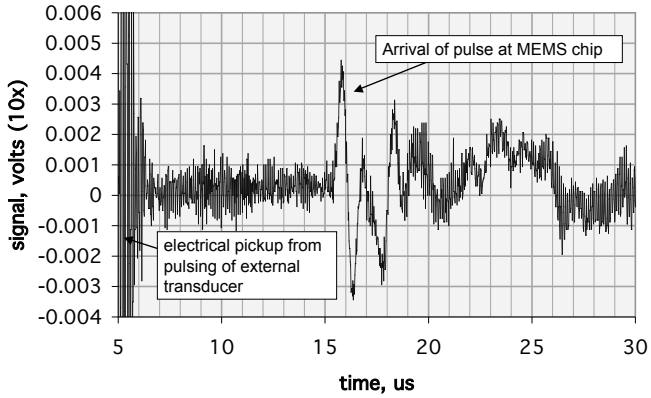

Figure 1.17: Waveform measured with capacitive sensor on CMOS-MEMS chip with on-chip amplifier.

1.4 Ultrasonics

An important class of acoustic sensors is ultrasonics sensors. Ultrasonic sensors are regularly used for flaw detection and medical imaging. Both of these applications typically use piezoelectric transducers because of the large acoustic energies generated from a modest voltages. Many MEMS sensors use capacitive sensing, and are called "Capacitive MEMS Ultrasonic Transducers" (cMUTs). Capacitive transducers suffer from parasitics in the cabling leading to read out electronics. Integrating electronics with an on-chip amplifier, as is possible with CMOS MEMS, is an effective way to solve this problem. Figure 1.17 shows a waveform from a capacitive sensor on a CMOS (TSMC 0.35 μm process) sensor chip packaged for use in water, and intended for gravimetric mass detection. A single diaphragm on this chip was only 130 μm square, and connected to a high-impedance amplifier with a gain of about 10, whose output was digitized. Wiring multiple diaphragms in parallel would improve the noise level, or multiple diaphragms could be placed around the chip area to perform phased array operations. The packaging of this chip was the same as the piezoresistive chip discussed here, and is discussed later.

Another approach is to use piezoresistors on chip, and provide a low-impedance output which is relatively unaffected by cable capacitance (it may also be amplified on chip if desired). It should be noted that the piezoresistors are not added to the CMOS MEMS device through micromachining, but rather are made from the polysilicon layer normally used for transistor gates. Thus the resistors are aligned to the mechanical structures as accurately as the CMOS process would make a transistor.

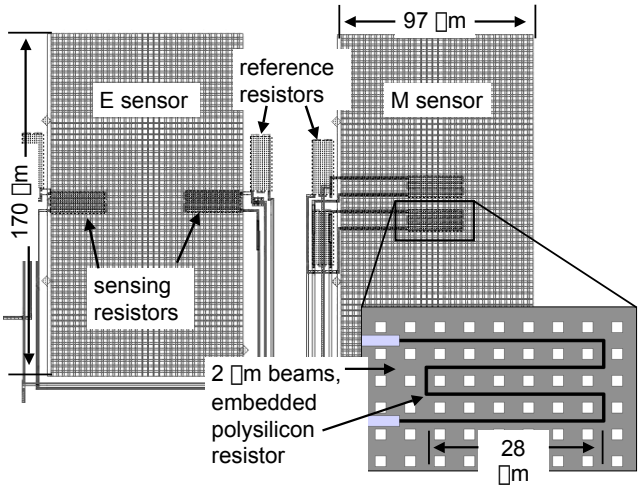

Figure 1.18: Layout of piezoresistors on diaphragms.

In this section, we will discuss the fabrication of the sensors, and how they are packaged for use in water. We will discuss the acoustic environment in which these ultrasonic sensors are used, and how this differs from the environment in which audio MEMS devices are used. A description of the experimental setup will be presented, followed by results for the CMOS MEMS sensors.

1.4.1 Fabrication

Piezoresistive sensors were fabricated using the Agilent 0.5 μm process at MOSIS (Marina del Rey, CA), and the CMOS-MEMS process developed at Carnegie Mellon Univerisity (CMU)[16], described earlier (Section 1.2.2). Rather than building the diaphragms out of a serpentine mesh structure, as we did for the audio-range microphone and speaker, we used a solid plate with 0.9 μm holes 2.0 μm apart in a rectangular grid (Figure 1.18). The holes serve to allow the underetch of the silicon during the CMOS-MEMS processing. Piezoresistors were formed by running a serpentine line of polysilicon through part of the diaphragm, and performing the silicide block only on the longer sections of the polysilicon in order to maximize $\Delta R/R$ when the diaphragm flexes. The resistors were placed midway between the far ends of the diaphragm in the long direction, and so we can approximate the mechanics in the region of the resistors as being like fixed-fixed beams running in the short direction, parallel to the long members of the polysilicon resistors. For uniform loading of the beams, we expect the greatest strain at the edges (assuming a true clamped scenario), and also significant strain at the middle of the beams. Though the peak strain in the middle is only half that at the edges, there is the advantage that the strain remains significant over a longer length, so the average

strain on the resistor remains large. For this reason, we made two sensor designs, one with the resistors at the edge, and one with the resistors at the middle.

Our prototype chip had 16 individual sensors, each with a resistive bridge connected to an amplifier. Half of the sensors had the sensing resistors near the middle of the diaphragm ("**M**" sensors) and half had the resistors near the edge ("**E**" sensors). There were also two sensors on each chip that were wired directly to bond pads to be measured directly with external circuitry. The individual sensor diaphragms have a pitch of 150 μm.

1.4.2 Acoustics in Water and Experimental Setup

Many of the envisioned applications of ultrasonic sensing arrays take place in wet mediums, such as in the body, or aided by gels. After the release of the mechanical structures using the standard CMU CMOS-MEMS processing[16], we packaged the chip in a dual in-line package (DIP) and then slowly poured PDMS (polydimethylsiloxane, DOW Corning Sylgard 184) over the the chip and bond wires. The thickness of the cured gel above the piezoresistive sensor was measured optically to be between 500 and 600 microns. This served to electrically insulate the wires, and protect the structures on the chip. As the cured PDMS also has an acoustic impedance similar to that of water, it provides a means to acoustically couple the pressure waves in the water to the mechanical structures. The MEMS structures are designed to have an acoustic impedance significantly less than the surrounding medium (the gel), so to a good approximation their movement follows the particle velocity of the medium, and provide a much wider bandwidth than conventional ultrasonic transducers such as PZT.

To hold water above the chip, plastic weighing dishes were glued to the top surface of the DIP packages with cyanoacrylate glue and then the seam was sealed with epoxy. The DIP package was inserted into a protoboard for electrical testing. Figure 1.19 shows an equivalent electrical circuit for the sensor in our experimental setup. Designing sensors for use in liquid uses a very different set of considerations than designing for sensing in air. The acoustic impedance of a typical CMOS-MEMS structure is much greater than that of air, but less than that of water. In the schematic, the water and silicone are modeled as transmission lines. Their length and characteristic impedance should be chosen based on the distance the waves travel through the medium, and the characteristic acoustic impedance of the medium. If one compares this system with the acoustic system for the microphone (Section 1.2.3) or the speaker (Section 1.3.3), one notes that we can ignore the impedance of any air trapped in the diaphragm-substrate gap. This is reasonable as the overall scale of acoustic impedances in the system is much higher than that of the microphones and speakers. No vent holes are necessary for this reason, which simplifies the fabrication process. It is also not necessary to coat the structures with chemical vapor deposition (CVD) polymer to seal the gaps, as this is taken care of by the PDMS gel.

1.4.3 Measurement and Results

The piezoresistive transducers were characterized by exciting them with an external source, a Krautkramer PZT (Lead Zirconate Titanate) transducer partially submerged in the water approximately 16 mm above the sensor surface. As a means to assess their sensitivity, the same measurements were taken with a MEMS capacitive transducer, for which the incoming

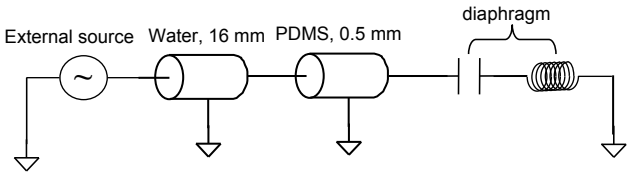

Figure 1.19: Equivalent electrical circuit for acoustic system of ultrasonic experiments

medium displacement could be calculated from the output signal and known geometry[27]. A DC voltage of 10 volts was applied across opposite corners of the bridge formed by the piezoresistors, and the other two corners of the bridge, which gave signals of opposite polarity, were measured separately and algebraically subtracted later. Both types of sensors, **E** and **M** were measured and compared. The signals from the **E** and **M** sensor elements have opposite polarities because of the opposite strains experienced in the edge versus the middle of the diaphragm. The results of the piezoresistive sensor measurements are shown in Figure 1.20. The waveform shown is the algebraic difference of the signals from opposite corners of the resistive bridge, to reduce the effect of the electromagnetic coupling to the source transducer. It is still possible to discern the time of the initial excitation pulse applied to the external transducer. The first arrival pulse appears as a series of pulses that die out, which is consistent with a model including reflections between the water-silicone interface and the sensor chip. The acoustic impedance of water is $1.5\text{x}10^5$ g/(cm^2·s), and we have previously estimated the acoustic impedance of the silicone gel to be $2\text{x}10^5$ g/(cm^2·s)[28]. The acoustic impedance of the CMOS diaphragm was calculated (based on a solid plate of the same thickness as our beams), to be $9\text{x}10^3$ g/(cm^2·s), about an order of magnitude less than the surrounding medium.

Because the acoustic impedance of the medium (water or PDMS) dominates the system, the deflection produced in the transducer diaphragm is roughly the same over a large range of transducer acoustic impedances. Using this fact, we were able to calculate the sound pressure at the sensor due to the external source, and estimate the gauge factor of the polysilicon. The deflection of the capacitive transducer, based on its geometry and voltage measurements, was approximately 7 nm at its center. Assuming the same central deflection for the piezoresistive sensor, and averaging the strain along the resistors, we calculate from the measured voltage

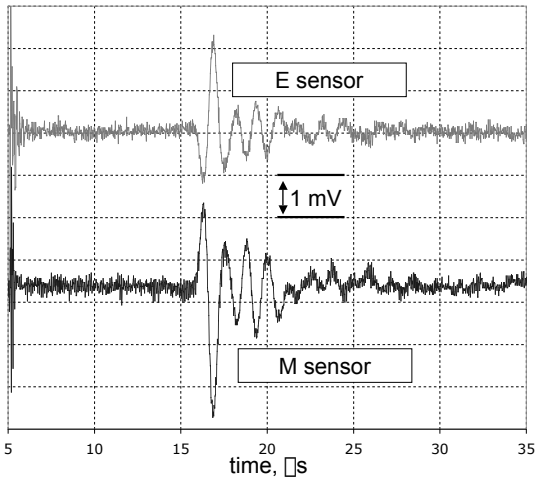

Figure 1.20: Waveform measured from piezoresistive sensor, with source 16 mm from surface of sensor.

a gauge factor of -44 and -51 for the **E** and **M** resistors respectively. This is higher than the range for n-type polysilicon reported elsewhere[29, 30] of -15 to -22, but the difference may come from a different acoustic reflection pattern (pressure doubling or diffraction) in the capacitive vs. piezoresistive sensors. It is possible that there is more effect of pressure doubling in the case of the piezoresistive sensor because of the greater fraction of device area that is not covered by sensors, i.e. the average acoustic impedance is closer to a hard wall.

1.4.4 Phased Array Behavior

The **E** and **M** sensors have a center-to-center distance of 150 μm. To test the phased array behavior of the sensor chip, we captured waveforms from the two sensor elements while the source transducer was positioned about a centimeter away from the sensors along the line passing through the sensor elements.

Waveforms were captured from both sensor elements of the piezoresistive chip, while positioning the source transducer at opposite ends of the water reservoir. Close up views of the detected signal are shown in Figure 1.21, corrected for the polarity difference between the sensor elements. In the left hand graph, the source was one the side of the sensor chip closest to the E sensor element; we see the pulse arrival is about 0.1 μs earlier for the E sensor element than the M element, which is consistent with a 1550 m/s sound speed in water and 150 μm element pitch. The opposite is true when the source is placed on the side of the chip closer to the M sensor element, demonstrating that an array of individual sensors may be used to determine direction. Increasing the spacing of the transducers would improve the sensitivity

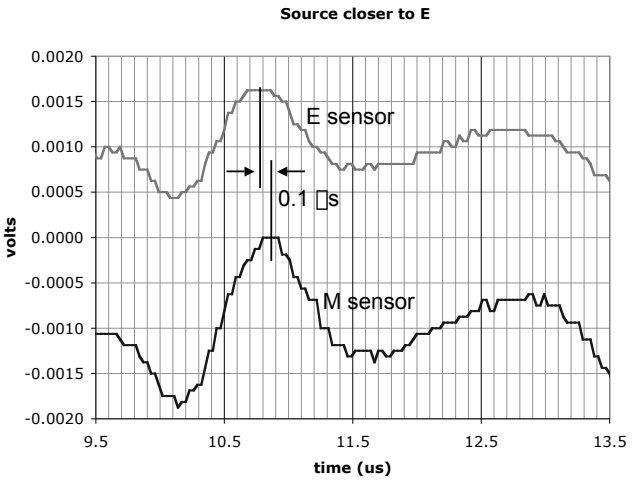

Figure 1.21: Pulse arriving at slightly different times for the two sensor elements.

to direction.

1.5 Conclusions

We have explored the emerging application of CMOS-MEMS technology to the design and fabrication of both microphone chips and speaker chips: audio to ultrasonic frequencies, for use in air and in liquid. The advantages of MEMS— miniaturization, multiple components on a single chip, and on-chip signal processing have been highlighted in use for different devices in a variety of applications. MEMS technology allied with the signal processing power of standard CMOS is the key element and unique power of CMOS-MEMS, delivering new capabilities to audio products today and enabling revolutionary audio products of the future.

1.6 Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. CMS-0329880. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. Support was also provided by the Pennsylvania Infrastructure Technology Alliance, a partnership of Carnegie Mellon, Lehigh University, and the Commonwealth of Pennsylvania's Department of Economic and Community Development. Other support was provided by Adtranz, and DARPA through the MARCO, ICESCATE, and ASIMPS projects.

The authors would like to acknowledge gifts from Krautkramer Inc., and thank Dr. Lars Erdmann for advice and assistance in diaphragm processing, and Prof. Irving Oppenheim for his analysis of the serpentine mesh mechanics.

Bibliography

[1] C.H. Han and E.S. Kim. Parylene-diaphragm piezoelectric acoustic transducers. In *Proceedings of the 13th IEEE International Conference on Micro Electro Mechanical Systems (MEMS 2000), Miyazaki, Japan, Jan. 23-27, 2000*, pages 148–152, 2000.

[2] S.S. Lee, R.P. Ried, and R.M. White. Piezoelectric cantilever microphone and microspeaker. *J. MEMS*, 5 (4):238–242, 1996.

[3] C.H. Han and E.S. Kim. Fabrication of piezoelectric acoustic transducers built on cantilever-like diaphragm. In *Proceedings of the 14th IEEE International Conference on Micro Electro Mechanical Systems (MEMS 2001), Interlaken, Switzerland, 21 - 25 January 2001*, pages 110–113, 2001.

[4] J.J. Neumann and K.J. Gabriel. CMOS-MEMS membrane for audio-frequency acoustic actuation. *Sensors and Actuators A*, 95:175–182, 2002.

[5] M. Pedersen, W. Olthuis, and P. Bergveld. An integrated silicon capacitive microphone with frequency-modulated digital output. *Sensors and Actuators A*, 69:267–275, 1998.

[6] M. Brauer, A. Dehe, T. Bever, S. Barzen, S. Schmitt, M. Fuldner, and R. Aigner. Silicon microphone based on surface and bulk micromachining. *J. Micromech. Microeng.*, 11:319–322, 2001.

[7] Y.B. Ning, A.W. Mitchell, and R.N. Tait. Fabrication of a silicon micromachined capacitive microphone using a dry-etch process. *Sensors and Actuators A*, 53:237–242, 1996.

[8] J.J. Neumann and K.J. Gabriel. A fully integrated CMOS-MEMS audio microphone. In *proceedings of IEEE Transducers'03: The 12th International Conference on Solid-State Sensors, Actuators and Microsystems, Boston, Massachusetts, USA, June 8 - 12, 2003 (www.transducers03.org).*, pages 230–233, 2003.

[9] F.J.M. van der Eerden, H-E.de Bree, and H. Tijdeman. Experiments with a new acoustic particle velocity sensor in an impedance tube. *Sensors and Actuators A*, 69:126–133, 1998.

[10] W.W. Gibbs. Micromicrophones- new sensors detect sound using light and heat. *Scientific American*, May 1999.

[11] P. Horowitz and W. Hill. *The Art of Electronics*. Cambridge University Press, Cambridge, 1980.

[12] T.B. Gabrielson. Fundamental noise limits for miniature acoustic and vibration sensors. *Journal of Vibration and Acoustics*, 117:405–410, 1995.

[13] B.B. Bauer. Equivalent circuit analysis of mechano-acoustic structures. In *Microphones (Anthology)*. Audio Engineering Society, 1979.

[14] H.B. Callen. *Thermodynamics and an Introduction to Thermostatistics (Second Edition)*. John Wiley and Sons, New York, 1985.

[15] R.J. Roark, R.G. Budynas, and W.C. Young. *Roark's Formulas for Stress and Strain*. McGraw-Hill, 2001.

[16] G.K. Fedder, S. Santhanam, M.L. Reed, S.C. Eagle, D.F. Guillou, M.S.-C. Lu, and L.R. Carley. Laminated high-aspect-ratio microstructures in a conventional CMOS process. *Sensors and Actuators A*, 57:103–110, 1996.

[17] M.S.-C. Lu, X. Zhu, and G.K. Fedder. Mechanical property measurement of 0.5-mm CMOS microstructures. *Mat. Res. Soc. Symp. Proc.*, 518:27, 1998.

[18] H.T. Souther. An adventure in microphone design. In *Microphones (Anthology)*. Audio Engineering Society, 1979.

[19] http://www.fmsystems-inc.com/eng_desp.htm (discussion and table of pre-emphasis curve).

[20] F.A. Everest. *Master Handbook of Acoustics, Fourth edition.* McGraw Hill, 2001.

[21] A.D. Pierce. *Acoustics- An Introduction to Its Physical Principles and Applications*. The Acoustical Society of America, Woodbury, NY, 1991.

[22] I.J. Oppenheim. private conversation.

[23] B.M. Diamond. *Digital Sound Reconstruction Using Arrays of CMOS-MEMS Microspeakers (M.S. thesis)*. Carnegie Mellon Univeristy, Pittsburgh, Fall 2002.

[24] U.S. patents 4555797, 4360707, 4566120.

[25] www.1limited.com.

[26] B.M. Diamond, J.J. Neumann, and K.J. Gabriel. Digital sound reconstruction using arrays of CMOS-MEMS microspeakers. In *proceedings of IEEE Transducers'03: The 12th International Conference on Solid-State Sensors, Actuators and Microsystems, Boston, Massachusetts, USA, June 8 - 12, 2003 (www.transducers03.org).*, pages 238–241, 2003.

[27] J.J. Neumann, D.W. Greve, and I.J.Oppenheim. Comparison of piezoresistive and capacitive ultrasonic transducers. In *proceedings of SPIE Symposium on Smart Structures & Materials/NDE 2004, San Diego, California, March 14-18, 2004*, 2004.

[28] D.W. Greve, J.J. Neumann, I.J. Oppenheim, S.P. Pessiki, and D. Ozevin. Robust capacitive MEMS ultrasonics transducers for liquid immersion. In *2003 IEEE International Ultrasonics Symposium Oct 5-8, 2003 Honolulu, Hawaii conference of the Ultrasonics Ferroelectrics, and Frequency Control Society*, pages 3F–4, 2003.

[29] L. Cao, T. S. Kim, J. Zhou, S. C. Mantell, and D. L. Polla. Calibration technique for MEMS membrane type strain sensors. In *Proc. Thirteenth Biennial University/Government/Industry IEEE Microelectronics Symposium*, pages 204–210, 1999.

[30] P.J. French and A.G.R. Evans. Polycrystalline silicon strain sensors. *Sensors and Actuators*, 8:219–225, 1985.

List of Figures

1.1 Equivalent electrical model of a typical microphone acoustic system. 5
1.2 Sequence of steps to form CMOS-MEMS diaphragms. 8
1.3 Drawing of a typical serpentine mesh pattern. Dark areas are beams, light areas are gaps. 9
1.4 SEM of serpentine mesh unit cell. 9
1.5 Equivalent electrical model of microphone acoustic system. 10
1.6 Predicted and measured frequency response of the two prototype microphones. 12
1.7 Noise power spectrum of microphone. 13
1.8 Change in sensitivity as a function of frequency, for several temperatures. . . 14
1.9 Photographs of earphone construction. 15
1.10 Electrical equivalent model of earphone acoustic system. 17
1.11 Diaphragm deflection as a function of voltage, for a 1442 μm square diaphragm and 60 μm diaphragm-substrate gap. 19
1.12 Effect of DC bias voltage on frequency response of CMOS-MEMS earphone. 19
1.13 Predicted earphone response (best fit of material property parameters). . . . 20
1.14 Photograph of DSR chip fabricated in CMOS. 22
1.15 Successive approximation to a 500 Hz sine wave, by increasing the number of bits from 1 to 3. 23
1.16 Frequency response of DSR chip, driven digitally vs. analog driven. 23
1.17 Waveform measured with capacitive sensor on CMOS-MEMS chip with on-chip amplifier. 24
1.18 Layout of piezoresistors on diaphragms. 25
1.19 Equivalent electrical circuit for acoustic system of ultrasonic experiments . . 27
1.20 Waveform measured from piezoresistive sensor, with source 16 mm from surface of sensor. 28
1.21 Pulse arriving at slightly different times for the two sensor elements. 29