Basic principles of MEMS microphones

John Widder, Alessandro Morcelli - May 14, 2014

Introduction

The application of MEMS (microelectro-mechanical systems) technology to microphones has led to the development of small microphones with very high performance. MEMS microphones offer high SNR, low power consumption, good sensitivity, and are available in very small packages that are fully compatible with surface mount assembly processes. MEMS microphones exhibit almost no change in performance after reflow soldering and have excellent temperature characteristics.

Figure 1 Top port and bottom port MEMS microphones

MEMS microphone acoustic sensors

MEMS microphones use acoustic sensors that are fabricated on semiconductor production lines using silicon wafers and highly automated processes. Layers of different materials are deposited on top of a silicon wafer and then the unwanted material is then etched away, creating a moveable membrane and a fixed backplate over a cavity in the base wafer. The sensor backplate is a stiff perforated structure that allows air to move easily through it, while the membrane is a thin solid structure that flexes in response to the change in air pressure caused by sound waves.
Figure 2 Cross-section diagram of a MEMS microphone sensor

![Cross-section diagram of a MEMS microphone sensor](image)

Figure 3 A typical MEMS microphone sensor viewed from above

Changes in air pressure created by sound waves cause the thin membrane to flex while the thicker backplate remains stationary as the air moves through its perforations. The movement of the membrane creates a change in the amount of capacitance between the membrane and the backplate, which is translated into an electrical signal by the ASIC.

**MEMS microphone ASICs**

The ASIC inside a MEMS microphone uses a charge pump to place a fixed charge on the microphone membrane. The ASIC then measures the voltage variations caused when the capacitance between the membrane and the fixed backplate changes due to the motion of the membrane in response to sound waves. Analog MEMS microphones produce an output voltage that is proportional to the instantaneous air pressure level. Analog mics usually only have 3 pins: the output, the power supply voltage (VDD), and ground. Although the interface for analog MEMS microphones is conceptually simple, the analog signal requires careful design of the PCB and cables to avoid picking up noise between the microphone output and the input of the IC receiving the signal. In most applications, a low noise audio ADC is also needed to convert the output of analog microphones into digital format for processing and/or transmission.

As their name implies, digital MEMS microphones have digital outputs that switch between low and high logic levels. Most digital microphones use pulse density modulation (PDM), which produces a highly oversampled single-bit data stream. The density of the pulses on the output of a microphone using pulse density modulation is proportional to the instantaneous air pressure level. Pulse density modulation is similar to the pulse width modulation (PWM) used in class D amplifiers. The difference is that pulse width modulation uses a constant time between pulses and encodes the signal in the pulse width, while pulse density modulation uses a constant pulse width and encodes the signal in the time between pulses.

In addition to the output, ground, and VDD pins found on analog mics, most digital mics also have
inputs for a clock and a L/R control. The clock input is used to control the delta-sigma modulator that converts the analog signal from the sensor into a digital PDM signal. Typical clock frequencies for digital microphones range from about 1 MHz to 3.5 MHz. The microphone’s output is driven to the proper level on the selected clock edge and then goes into a high impedance state for the other half of the clock cycle. This allows two digital mic outputs to share a single data line. The L/R input determines which clock edge the data is valid on.

The digital microphone outputs are relatively immune to noise, but signal integrity can still be a concern due to distortion created by parasitic capacitance, resistance, and inductance between the microphone output and the SoC. Impedance mismatches can also create reflections that can distort the signals in applications with longer distances between the digital mic and the SoC.

Although codecs are not required for digital MEMS microphones, in most cases the pulse density modulated output must be converted from single-bit PDM format into multibit pulse code modulation (PCM) format. Many codecs and SoCs have PDM inputs with filters that convert the PDM data into PCM format. Microcontrollers can also use a synchronous serial interface to capture the PDM data stream from a digital mic and convert it into PCM format using filters implemented in software.

**MEMS microphone packages**

MEMS microphone have hollow packages that consist of a substrate with pads that can be soldered to a circuit board or flex circuit, and a lid that creates a cavity where the acoustic sensor and the ASIC are located. Most MEMS microphones use separate die for the MEMS sensor and the interface ASIC, which allows the MEMS process to be optimized for creating moving structures while using a standard CMOS process to fabricate the ASIC. The ASIC is wire bonded to sensor and the substrate, and a lid is then placed over them and sealed to the substrate.

![Figure 4 A MEMS microphone with its lid removed](image)

MEMS microphones need to have a hole in their package to allow sound to reach the acoustic sensor. The sound inlet can be located either in the lid (top port) or on the bottom next to the solder...
pads (bottom port). Bottom port microphones also require a hole in the circuit board they are mounted on to allow sound to reach the sound inlet. The choice of whether to use a top port or bottom port microphone is usually determined by factors such as the location of the microphone in the product and manufacturing considerations, to mention a couple. Performance can also be a major factor in microphone port selection since top port microphones have traditionally had poorer performance than equivalent bottom port microphones. However, the introduction of high performance top-port mics such as ST’s MP34DT01 means this is no longer necessarily true.

The membrane of the acoustic sensor divides the interior of a MEMS microphone into two sections. The area between the sound inlet and the sensor membrane is generally referred to as the front chamber, and the section on the other side of the membrane is known as the back chamber (Figure 5). The sensor in bottom port microphones is usually placed directly over the sound inlet which provides several benefits.

![Figure 5 Cross-section diagram of a typical bottom port MEMS microphone](image)

The sensitivity of most MEMS microphones increases at higher frequencies. This increase in sensitivity is caused by the interaction between the air in the sound inlet and the air in the front chamber of the microphone. This interaction creates a Helmholtz resonance, which is the same phenomenon that creates sound when blowing into a bottle. As with bottles, smaller air volumes create higher resonant frequencies and larger air volumes create lower resonant frequencies. The microphone sensor is mounted directly over the sound inlet in most bottom port microphones, which results in a relatively small front chamber and a high center frequency for the Helmholtz resonance. Because the Helmholtz resonance is normally located in the upper part of the audio band, increasing the resonant frequency leads to a flatter frequency response.

Placing the sensor directly over the sound inlet also creates a relatively large back chamber. A larger volume of air in the back chamber makes it easier for the membrane to move in response to sound waves, which improves the sensitivity of the microphone and leads to higher SNR. A large back chamber also improves the microphone’s low frequency response. The construction of top port microphones has traditionally been very similar to bottom port microphones, with the sensor and the interface IC mounted on the substrate with a hollow lid enclosing them. Traditionally, the only real difference between top port and bottom port microphones is that the sound inlet is located in the microphone lid instead of in the substrate. For these microphones, moving the sound inlet to the lid turns what was previously the front chamber into the back chamber and vice versa.
The smaller air volume in the back chamber of traditional top port MEMS microphones makes it more difficult for the membrane to move, which hurts the sensitivity of the sensor and leads to a lower SNR. In addition, the larger air volume in the front chamber between the sound inlet and the membrane lowers the resonant frequency, hurting the microphone’s high frequency response. This combination of lower SNR and poorer frequency response at both high and low frequencies is the reason that most top port microphones have poorer performance than an equivalent bottom port microphone.

An exception to this rule is STMicroelectronics’ MP34DT01 top-port digital MEMS microphone. ST’s proprietary packaging technology makes it possible to mount the MEMS sensor and the interface IC on the bottom side of the lid of the MP34DT01, directly beneath the sound inlet (Figures 7 and 8). This results in a small front chamber and a large back chamber and allows the MP34DT01 to achieve the same level of performance as the bottom port version of this mic, the MP34DB01.
Measuring MEMS microphone performance

The SI unit for pressure is the pascal (Pa), which is a linear measure of force per unit area (1 Pa = 1 N/m²). However, logarithmic scales are more convenient when discussing sound pressure levels due to the large dynamic range of the human ear, which can detect sounds from as low as 20 micropascals to more than 20 pascals. Because of this, the key measures of microphone performance are normally expressed in decibels (dB). 0dB SPL is equal to 20 µPa and 1 Pa is equal to 94dB SPL. The following parameters are normally the most important indicators of microphone performance:

Signal-to-noise ratio (SNR)

The signal-to-noise ratio (SNR) is the most important measure of microphone performance in most applications. The signal-to-noise ratio is the difference between a microphone’s sensitivity and its noise floor and is expressed in dB. The SNR of current MEMS microphones ranges from about 56 dB to about 66 dB.

Sensitivity

The sensitivity of a microphone is a measure of its response to a given sound pressure level. Sensitivity is normally specified at a frequency of 1 kHz and at 94 dB SPL (1 Pa). The sensitivity of analog microphones is usually expressed in decibels relative to 1 volt RMS (dBV) while the sensitivity of digital microphones is normally expressed in decibels relative to the microphone’s full scale output (dB FS).

Noise floor

The noise floor of a microphone is the amount of noise on its output in a perfectly quiet environment. Both the sensor and the interface ASIC contribute noise to the output of a microphone. The noise contributed by the sensor is created by the random Brownian motion of air molecules while the noise from the ASIC is created by the preamplifier and, in the case of digital microphones, the delta-sigma modulator. The noise floor is measured across the full audio band and A-weighting filters are used to provide a better measure of the noise level as perceived by human ears.

The noise floor is not always specified in microphone datasheets but it can be calculated by subtracting the microphone’s SNR from its sensitivity, providing a result in dBV or dB FS. The noise floor can be expressed as an equivalent input noise in dB SPL by subtracting the SNR from the sound pressure level that the sensitivity is measured at (usually 94 dB SPL).

Distortion (THD)

Distortion is a measure of how accurately a microphone can capture sound. Distortion is usually specified at about 94 dB – 100 dB SPL in order to provide a good indication of a microphone’s audio quality at normal sound levels.
Acoustic Overload Point (AOP)
Distortion normally does not increase much as the sound pressure level increases until the sound pressure level starts to approach the acoustic overload point of the microphone. When this happens the distortion starts rising rapidly. The acoustic overload point of a microphone is normally defined as the sound pressure level where the distortion reaches 10%.

Frequency response
The frequency response of a MEMS microphone refers to the change in its sensitivity at various frequencies. The frequency response of a microphone is usually set to 0 dB at 1 kHz to normalize the results. The sensitivity of most MEMS microphones falls off below 100 Hz and starts rising between about 4 kHz – 6 kHz due to the Helmholtz resonance. This is the reason that many MEMS mics only specify their frequency response between 100 Hz and 10 kHz. However, high performance MEMS microphones are available with a relatively flat frequency response over the full audio band from 20 Hz to 20 kHz.

Power supply rejection (PSR)
The power supply rejection of a microphone is a measure of its ability to prevent noise on the microphone’s power supply input from appearing on its output. PSR is usually specified with a 217 Hz square wave to simulate the TDMA noise generated by GSM cellular radios and/or a swept sine wave across the audio band.

Future trends

Higher SNR
MEMS microphone performance continues to improve. SNRs have increased from 55 – 58 dB a few years ago to 63 – 66 dB today, resulting in cleaner audio capture and allowing microphones to be used at greater distances with the same level of clarity. High SNR levels are needed by automatic speech recognition algorithms to achieve good word accuracy rates.

Higher sound pressure levels
Many microphone users are also requesting higher acoustic overload points to prevent distortion in loud environments. Distortion caused by clipping at sound pressure levels above the acoustic overload point can make recordings made in loud environments such as rock concerts unusable.

Smaller package sizes
MEMS microphone package sizes are also shrinking as consumer demand for thinner, lighter products continues to increase. Early MEMS microphones had package sizes of 3.76mm x 4.72mm x 1.25mm while today 3mm x 4mm x 1mm and 2.95mm x 3.76mm x 1mm packages are
common. Newer MEMS microphones are available in 2.5mm x 3.35mm x 0.98mm and 2.65mm x 3.5mm x 0.98mm packages. This trend is likely to continue, although smaller microphone packages make it more difficult to maintain or improve audio quality due to the shrinking size of the microphone’s back chamber.

**Ambient noise reduction**

Many smartphones and tablets are starting to use more than one microphone to enable features such as video recording. Another common way in which multiple microphones are used is for ambient noise reduction. Many smartphones use a microphone located on the top or the back of the phone to detect noise in the surrounding environment and subtract it from the output from the voice microphone(s) to help improve intelligibility. Microphones whose primary purpose is video recording are frequently also used for ambient noise reduction.

**Beamforming**

Arrays of two or more microphones are also being used to perform beamforming, which processes the outputs from a microphone array to increase the sensitivity in a particular direction while at the same time rejecting sounds from other directions. Most microphones have omnidirectional outputs, i.e., the sensitivity is the same in all directions, but in many cases, it is desirable to focus the sensitivity in a particular direction and reduce the sensitivity in other directions in order to improve intelligibility. Beamforming uses the phase differences of sounds arriving from different directions to focus the sensitivity of the microphones in a particular direction. Beamforming can also be used to locate the direction a sound is coming from. Beamforming is particularly useful in applications where the microphone is not close to the person speaking such as living rooms, conference rooms, automobiles, etc. It can also be very useful when using speakerphones or videoconferencing in noisy environments.

**Tighter control of sensitivity**

The performance algorithms used to perform functions such as noise cancellation and beamforming usually assume that the sensitivity of the microphones being used is the same, so variations in sensitivity between the microphones in an array hurt the performance of the algorithms.

This has created a demand for tighter sensitivity matching. MEMS microphones typically have a ±3 dB tolerance on sensitivity, but this can be tightened to ±1 dB by screening microphones to a tighter tolerance (binning) and/or trimming of the microphone ASICs to compensate for normal variation in microphone parameters.

The use of MEMS microphones is increasing rapidly. IHS predicts that the usage of MEMS microphones will almost double over the next four years, rising from about 2.7 billion units in 2013 to about 5 billion units in 2017. The growth of applications such as tablets and smartphones is a major contributor to the growth in the use of MEMS microphones. The growth in the usage of MEMS microphones is also being fueled by the increasing use of two or more microphones in products to enable features such as video recording, active noise cancellation, and beamforming.

Also see:
• Acoustic Design for MEMS Microphones
• Akustica’s MEMS microphones for smartphones and wearables
• Microphones: A sound technology choice for communication and control