In this class we will learn how to design a MEMS microphone. We will choose the diameter of the membrane and its bias voltage to comply with noise specifications; we will dimension the front-end according to full-scale range specs.

**PROBLEM**

We have to design a MEMS microphone. The schematic of the device is shown in Fig. 1. The parameters of the system are reported in Table 1.

We are asked to…

1. Choose the diameter of the membrane in order to comply with noise requirements, aiming at a well-balanced sensor in terms of noise performance.

2. Calculate the required bias voltage, $V_{DC}$, in order to have a well-balanced sensor in terms of noise performance.

3. Compare the required bias voltage with the pull-in voltage. Does the electrostatic softening impact on the previous calculation?

4. Suitably dimension the feedback network of the front-end.
A microphone is a dual-die device consisting of two components, the integrated circuit and the sensor. The sensor uses MEMS technology and it is basically a poly-silicon capacitor. The capacitor consists of two poly-silicon plates/surfaces. One plate is fixed while the other one is movable (respectively, the green plate and the grey one shown in Fig. 1). The fixed surface is covered by an electrode to make it conductive and is full of acoustic holes which allow sound to pass through. The movable plate is able to move. A ventilation hole, allows the air compressed in the back chamber to flow out and consequently allows the membrane to move back. A microphone MEMS sensor is basically a variable capacitor where the transduction principle is the coupled capacitance change between a fixed plate (back plate) and a movable plate.
(membrane) caused by the incoming wave of the sound. A simplified model is reported in Fig. 2. The integrated circuit converts the change of the polarized MEMS capacitance into an analog output.

A microphone is a sound-to-electricity transducer: any output signal corresponds to a specific input signal. The input signal of a microphone is sound pressure, $p$. Sound pressure (or acoustic pressure) is the local pressure deviation from the ambient (average, or equilibrium) atmospheric pressure, caused by a sound wave. A sound pressure wave can be described as

$$p = p_a \sin(2\pi f_a t),$$

where $f_a$ is the acoustic frequency. The acoustic frequency range is defined from 20 Hz to 20 kHz.

Usually, sound pressure is expressed in dBSPL:

$$p_{SPL} = 20 \log_{10} \frac{p_a}{p_{ref}},$$

where $p_{ref}$ is the reference pressure,

$$p_{ref} = 20 \mu Pa,$$

i.e.

$$p_{ref,SPL} = 0 \text{dBSPL},$$

which is commonly considered as the threshold of human hearing (roughly the sound of a mosquito flying 3 m away).

Fig. 3 reports examples of sound levels and the corresponding maximum exposure time required to prevent ear damage. We can thus easily translate our requirements in pascals:

$$p_{a,min} = p_{ref} \cdot 10^{\left(\frac{p_{a,min,SPL}}{20}\right) / \left(\frac{p_{a,ref,SPL}}{20}\right)} = 893 \mu Pa,$$
that corresponds to a quiet whisper, and

$$p_{a,\text{max}} = p_{\text{ref}} \cdot 10^{\frac{P_{a,\text{max},\text{SPL}}}{20}} = 28.3 \text{ Pa},$$

that corresponds to a loud rock concert.

Given the frequency range of interest, we can determine the input-referred pressure noise density, by simply dividing the minimum detectable signal for the bandwidth of our sensor, i.e. the whole audio bandwidth:

$$s_{n,p} = \frac{p_{a,\text{min}}}{\sqrt{f_{\text{max}} - f_{\text{min}}}} \approx \frac{p_{a,\text{min}}}{\sqrt{f_{\text{max}}}} = 6.31 \mu \text{Pa}/\sqrt{\text{Hz}}.$$  

Given a certain pressure, $p$, applied to the membrane, the corresponding applied force can be simply evaluated as

$$F = pA = p\pi r^2,$$
where

\[ A = \pi r^2 \]

is the area of the membrane, and \( r \) is its radius. The force \( F \) is a uniformly distributed force that causes a certain displacement of the membrane.

To be more precise, the membrane deflects if there's a pressure difference between its two sides. As one can observe from Fig. 1, both the chambers faced by the membrane are kept at the atmospheric pressure, so that the DC displacements induced by the atmospheric pressure are canceled; the pressure difference applied to the membrane, \( p_m \), is

\[ p_m = (p_{atm} + p_{sound}) - p_{atm} = p_{sound}. \]

Figure 4: 2D vs 1D model of the membrane.

It can be demonstrated (e.g. in I. O. Wygant, M. Kupnik and B. T. Khuri-Yakub, *Analytically calculating membrane displacement and the equivalent circuit model of a circular CMUT cell*, in 2008 IEEE Ultrasonics Symposium, Beijing, 2008, pp. 2111-2114) that a vibrating membrane can be well-approximated as a 1D mass-spring-damper system (Fig. 4). This 1D model is defined in this way: given a uniformly distributed force applied onto the membrane, the dynamic behavior of the membrane can be equivalently modeled as a 1D piston that uniformly moves in \( y \)-direction, whose mass is \( m \), whose spring constant is \( k_m \), and whose damping coefficient is \( b \). For a circular membrane of radius \( r \), the 1D parameters are

\[ k_m = \frac{192\pi D}{r^2}, \]

\[ m = 1.84\pi r^2 h\rho, \]
\[ b = b_{\text{area}} \pi r^2. \]

\( D \) is the so-called flexural rigidity of the membrane,

\[ D \approx \frac{E h^3}{12}. \]

The resonance frequency of such a membrane can be evaluated as

\[ f_r = \frac{1}{2\pi} \sqrt{\frac{k_m}{m}} = \frac{1}{2\pi} \sqrt{\frac{192\pi D}{r^2}} = \frac{1}{2\pi} \frac{10.22}{r^2} \sqrt{\frac{h \rho}{D}}. \]

The model is defined in such a way that, given a certain force \( F \) applied to the membrane, the corresponding displacement

\[ y = \frac{1}{k_m} F, \]

represents the average displacement of the membrane (see Fig. 4).

We can immediately calculate the flexural rigidity of the membrane of our microphone

\[ D = \frac{E h^3}{12} = 12.5 \times 10^{-9} \text{Nm}. \]

**QUESTION 1**

The diameter of the membrane can be chosen from noise specifications. Aiming at a well-balanced sensor in terms of noise performance, i.e.

\[ s_{n, \text{MEMS}, p} = s_{n, \text{ELN}, p}, \]

we get

\[ s_{n, \text{TOT}, p} = \sqrt{2} s_{n, \text{MEMS}, p} = \sqrt{2} s_{n, \text{ELN}, p}, \]

i.e.

\[ s_{n, \text{MEMS}, p} = \frac{s_{n, \text{TOT}, p}}{\sqrt{2}}, \]

As \( F = Ap \), input-referring MEMS thermo-mechanical noise, we get

\[ s_{n, \text{MEMS}, p} = \frac{s_{n, \text{MEMS}, F}}{A}, \]

where

\[ s_{n, \text{MEMS}, F} = \sqrt{4k_b T b} = \sqrt{4k_b T b_{\text{area}}A}. \]

Hence,

\[ s_{n, \text{MEMS}, p} = \frac{s_{n, \text{MEMS}, F}}{A} = \frac{\sqrt{4k_b T b_{\text{area}}A}}{A} = \frac{\sqrt{4k_b T b_{\text{area}}\pi r^2}}{\pi r^2}. \]
From
\[ s_{n,\text{TOT}} = \sqrt{2}s_{n,\text{MEMS}} = \sqrt{2}s_{n,\text{ELN}}, \]
\[ \sqrt{\frac{4k_B T \text{b}_{\text{area}} \pi r^2}{\pi r^2}} = s_{n,\text{TOT},p} \]

hence,
\[ r = \frac{\sqrt{4k_B T \text{b}_{\text{area}}}}{s_{n,\text{TOT},p}} = 304 \text{ nm}. \]

We can thus choose a radius of 304 \( \mu \text{m} \), i.e. a membrane diameter of 608 \( \mu \text{m} \).

Given the radius of the membrane, we can infer the 1D mass-spring-damper equivalent coefficients:
\[ k_m = \frac{192\pi D}{r^2} = 16\pi E h^3 = 81.5 \text{ N/m}, \]
\[ m = 1.84\pi r^2 h \rho = 1.27 \text{ nkg}, \]
\[ b = b_{\text{area}} \pi r^2 = 101 \cdot 10^{-6} \text{ N/(m/s)}. \]

The resonance frequency is
\[ f_r = \frac{1}{2\pi} \sqrt{\frac{k_m}{m}} = 40.2 \text{ kHz}, \]

while the quality-factor is
\[ Q = \frac{2\pi f_r m}{b} = 3.2. \]

Given the 1D model of the mechanical structure, we can easily report its mechanical transfer function, as
\[ G(j\omega) = \frac{Y(j\omega)}{F(j\omega)} = \frac{1}{k_m + j\omega b - \omega^2 m}, \]
where \( F \) is the applied force. As we are dealing with a microphone, i.e. a (sound) pressure sensor in the acoustic frequency range (20 Hz - 20 kHz), we can more conveniently evaluate the mechanical transfer function of our microphone as the displacement divided for the input pressure. As \( F = \pi r^2 p, \)
\[ T(j\omega) = \frac{Y(j\omega)}{P(j\omega)} = \pi r^2 \frac{1}{k_m + j\omega b - \omega^2 m}, \]
whose magnitude is reported in Fig. 5.

This is a good dynamic behavior! In fact, as we are working with a very large signal bandwidth, the best choice for our structure is to have the resonance frequency outside the frequency range of interest (as in accelerometers), possibly with a low quality-factor, in order to reject under-damped-related phenomena, such as long time constant, overshoots, ... The 40 kHz value of the resonance frequency is thus perfect for bandwidth constraints, as it automatically filters out any force/pressure at frequencies higher than the audio range (e.g. ultrasound).
As, for microphone purposes, we are working at frequencies lower than the resonance frequency, our microphone can be equivalently modeled as the spring constant only. As our input is a pressure, we can more conveniently define the mechanical sensitivity of our sensor, $\frac{\partial y}{\partial p}$, defined as the ratio between the average membrane displacement, $y$, and the pressure, $p$:

$$\frac{\partial y}{\partial p} = |T(\omega)|_{\omega < \omega_r} = \frac{\pi r^2}{k_m} = \frac{\pi r^2}{16\pi Eh^3} \frac{r^4}{r^2} = 3.56 \text{ nm/Pa.}$$

How can we read out the pressure-induced displacement? The variable-gap structure is basically a variable gap capacitor, whose capacitance variation can be easily read-out with a TCA-based front-end.

We can calculate the rest capacitance of the structure,

$$C_0 = \frac{\varepsilon_0 A}{g} = \frac{\varepsilon_0 \pi r^2}{g} = 2.57 \text{ pF},$$

and the capacitance variation per unit $y$-axis displacement,

$$\frac{\partial C}{\partial y} = \frac{C_0}{g} = \frac{\varepsilon_0 \pi r^2}{g^2} = 2.57 \cdot 10^{-6} \text{ F/m.}$$
The previously calculated thermo-mechanical noise can be re-calculated at the output of the mechanical domain of our sensor, i.e. as capacitance noise:

\[ s_{n,\text{MEMS,C}} = s_{n,\text{MEMS,P}} \frac{\partial y \partial C}{\partial p \partial y} = s_{n,\text{MEMS,F}} \frac{1}{k_m} \frac{\partial C}{\partial y} = 40.9 \text{ zF/\sqrt{Hz}}, \]

flat within the whole frequency range of interest. Our hypothesis of a well-balanced sensor is satisfied if the input-referred noise of the front-end is equal to 40.9 zF/\sqrt{Hz}.

**QUESTION 2**

![Figure 6: MEMS microphone readout system.](image)

The readout electronics system is reported in Fig. 6. The membrane is biased at \( V_{DC} \) with a dedicated voltage supply, while the other electrode is connected with the virtual ground of the front-end. The transfer function of the front end is

\[ \frac{V_{out}(j\omega)}{I(j\omega)} = -\frac{R_F}{1 + j\omega R_F C_F}. \]

Remembering that

\[ i = \frac{\partial C}{\partial t} V_{DC}, \]

hence

\[ I(s) = sV_{DC}C(s), \]

i.e.

\[ I(j\omega) = j\omega V_{DC}C(j\omega), \]

we can re-evaluate the transfer function of the TCA, now defined as the ratio between the output voltage, \( V_{out} \), and the capacitance variation \( C \), as

\[ T_{TCA}(j\omega) = \frac{V_{out}(j\omega)}{C(j\omega)} = -j\omega V_{DC} \frac{R_F}{1 + j\omega R_F C_F}. \]
which is an high-pass filter, whose pole frequency is

\[ f_p = \frac{1}{2\pi R_F C_F}. \]

If the pole frequency is designed to be equal to (or lower than) the minimum signal frequency, 20 Hz, the gain of the TCA (in the audio range) is flat and equal to

\[ \frac{\partial V_{out}}{\partial C} = |T_{TCA}(\omega)|_{\omega \approx \omega_p} = \frac{V_{DC}}{C_F}. \]

The input-referred noise of the front-end, assuming that the dominant noise source is the voltage noise of the op-amp, can be expressed as

\[ s_{n,ELN,C} = \frac{1 + \frac{C_p}{C_F}}{s_{n,OA,v} \frac{V_{DC}}{C_F}} \approx s_{n,OA,v} \frac{C_p}{V_{DC}}. \]

In the previous expression, we assumed \( C_p \gg C_F \). Note that, as shown in Fig. 6, the parasitic capacitance is given by the sum of two contributions: the interconnections capacitance (wire bonding + chip metal wires), modeled as \( C_I \), and the DC (rest) capacitance of the membrane, \( C_0 \), that contributes, as well, to the calculation of the overall capacitance from the virtual ground to ground:

\[ C_p = C_I + C_0 = 7.57 \text{ pF}. \]

The DC bias voltage, \( V_{DC} \), that forces electronics noise to be equal to MEMS thermo-mechanical noise is thus:

\[ V_{DC} = s_{n,OA,v} C_p \frac{V_{DC}}{s_{n,ELN,C}} = s_{n,OA,v} C_p \frac{V_{DC}}{s_{n,MEMS,C}} = 3.70 \text{ V}. \]

**QUESTION 3**

We can check whether the required DC bias voltage, \( V_{DC} \), is lower than the pull-in voltage, \( V_{PI} \). Let’s derive the expression. The electrostatic force applied to the 1D-modeled membrane is:

\[ F_e = \frac{1}{2} \frac{\partial C}{\partial y} V_{DC}^2. \]

As the complete expression of the capacitance is

\[ C(y) = \frac{\varepsilon_0 A}{g - y}, \]

\[ \frac{\partial C}{\partial y} = \frac{\varepsilon_0 A}{(g - y)^2}. \]

For small displacements, i.e. for \( y \ll g \),

\[ \frac{\partial C}{\partial y} = \frac{\varepsilon_0 A}{(g - y)^2} \approx \frac{\varepsilon_0 A}{g^2} + \frac{2\varepsilon_0 A}{g^3} \cdot y. \]
Hence,

$$F_e = \frac{1}{2} \frac{\partial C}{\partial y} V_{DC}^2 = \frac{1}{2} \frac{\varepsilon_0 A}{g^2} V_{DC}^2 + \frac{\varepsilon_0 A}{g^3} V_{DC}^2 y.$$  

The first term, constant, is responsible of a DC displacement of the membrane towards the fixed electrode. The second term, which is linearly proportional with the displacement $y$ is responsible of the electrostatic spring-softening effect, as

$$F_e \propto \frac{\varepsilon_0 A}{g^3} V_{DC}^2 y = k_e y,$$

where

$$k_e = \frac{\varepsilon_0 A}{g^3} V_{DC},$$

is the electrostatic-spring constant. The pull-in voltage can be evaluated as the DC voltage for which the electrostatic spring constant is equal to the mechanical one:

$$k_m = k_e = \frac{\varepsilon_0 A}{g^3} V_{DC},$$

from which,

$$V_{PI} = \sqrt{\frac{g^3 k_m}{\varepsilon_0 A}} = 5.63 \text{ V},$$

higher than the desired $V_{DC}$.

We can evaluate the electrostatic spring constant,

$$k_e = \frac{\varepsilon_0 A}{g^3} V_{DC} = 35.2 \text{ N/m},$$

which is not negligible if compared with the mechanical one (81.5 N/m)! Hence, we should re-write the mechanical sensitivity expression, taking into account this effect, i.e. modeling the spring constant of the system as

$$k = k_m - k_e.$$

The new expression for MEMS noise, evaluated as capacitance noise, is

$$s_{n,\text{MEMS},C} = s_{n,\text{MEMS},F} \frac{1}{k} \frac{\partial C}{\partial y} = s_{n,\text{MEMS},F} \frac{1}{k_m - k_e} \frac{\partial C}{\partial y} = \sqrt{4k_b T b} \frac{1}{k_m - \frac{\varepsilon_0 A}{g^3} V_{DC}^2} \frac{\partial C}{\partial y},$$

which should be equal to

$$s_{n,\text{ELN},C} = s_{n,\text{OA},C} \frac{C_p}{V_{DC}},$$

i.e.

$$\sqrt{4k_b T b} \frac{1}{k_m - \frac{\varepsilon_0 A}{g^3} V_{DC}^2} \frac{\partial C}{\partial y} = s_{n,\text{OA},C} \frac{C_p}{V_{DC}}.$$
from which,
\[
V_{DC} = \sqrt{\frac{\sqrt{4k_b T b} \frac{\partial C}{\partial y}}{s_{n,OA,\nu} C_p}} + \frac{4 \varepsilon_0 A}{g^3} k_m \frac{\partial C}{\partial y} - \frac{\sqrt{4k_b T b} \frac{\partial C}{\partial y}}{s_{n,OA,\nu} C_p} = 2.79 \text{ V}.
\]

We can re-calculate some useful parameters, as
\[
k_e = 20 \text{ N/m},
\]
\[
k = k_m - k_e = 61.5 \text{ N/m},
\]
\[
\frac{\partial y}{\partial p} = \frac{\pi r^2}{k} = \frac{\pi r^2}{k_m - k_e} = \frac{16 \pi E h^3}{r^2} - \frac{\varepsilon_0 A}{g^3} \frac{V_{DC}}{2} = 4.72 \text{ nm/Pa},
\]
\[
\frac{\partial C}{\partial p} = \frac{\partial y}{\partial p} \frac{\partial C}{\partial y} = 12.1 \text{ fF/Pa},
\]
\[
y_{min} = \frac{\partial y}{\partial p} p_{min} = 4.22 \text{ pm},
\]
\[
y_{max} = \frac{\partial y}{\partial p} p_{max} = 133 \text{ nm},
\]
\[
C_{min} = \frac{\partial C}{\partial p} p_{min} = 10.8 \text{ aF},
\]
\[
C_{max} = \frac{\partial C}{\partial p} p_{max} = 343 \text{ fF}.
\]

Regarding noise,
\[
s_{n,MEMS,C} = \sqrt{4k_b T b} \frac{1}{16 \pi E h^3} \frac{\varepsilon_0 \pi r^2}{r^2} - \frac{\varepsilon_0 A}{g^3} \frac{V_{DC}^2}{2} = 54.2 \text{ zF/Hz},
\]
\[
s_{n,ELN,C} = s_{n,OA,\nu} \frac{C_p}{V_{DC}} = 54.2 \text{ zF/Hz}.
\]

As expected, the two contributions are equal. The total input referred noise is thus,
\[
s_{n,TOT,p} = \sqrt{2} s_{n,MEMS,C} \frac{\partial C}{\partial p} = 6.32 \mu \text{ Pa/Hz},
\]
that, as expected, was not affected from \( V_{DC} \) re-design.

The new resonance frequency and \( Q \)-factor are 34.9 kHz and 2.7, respectively.
QUESTION 4

The feedback capacitance of the TCA determines the gain of the front-end,

\[
\frac{\partial V_{\text{out}}}{\partial C} = |T_{\text{TCA}}(\omega)|_{\omega = \omega_p} = \frac{V_{\text{DC}}}{C_F}.
\]

It can be dimensioned in such a way that, given the maximum capacitance variation (that corresponds to the maximum displacement, i.e. on the maximum sound pressure), the output voltage variation is equal to \(V_{\text{max}}\):

\[
V_{\text{max}} = \frac{V_{\text{DC}} C_{\text{max}}}{C_F},
\]

where \(C_{\text{max}} = 343 \text{ fF}\). Hence,

\[
C_F = \frac{V_{\text{DC}} C_{\text{max}}}{V_{\text{max}}} = 383 \text{ fF}.
\]

A short comment about non-linearity. The linearity error, \(\epsilon_{\text{lin}}\), of a single-ended parallel-plate-based capacitance measurement can be evaluated as

\[
\epsilon_{\text{lin}} = \frac{\Delta C_{\text{real}} - \Delta C_{\text{lin}}}{\Delta C_{\text{real}}} = \frac{(\varepsilon_0 A - \varepsilon_0 A y - \varepsilon_0 A)}{g} \left( \frac{g}{y} \right) - \varepsilon_0 A \left( \frac{g}{y} \right) = \cdots = \frac{y}{g}.
\]

This means that, with the previously calculated maximum displacement, 133 nm, the linearity error is greater than 10%! Remember that, if the capacitance measurement is differential, the linearity error can be expressed as

\[
\epsilon_{\text{lin,diff}} = \frac{\Delta C_{\text{diff,real}} - \Delta C_{\text{diff,lin}}}{\Delta C_{\text{diff,real}}} = \cdots = \left( \frac{y}{g} \right)^2.
\]

Hence, if the readout was differential, with the same maximum displacement, the linearity error would be lower, slightly higher than 1%. This is why a differential readout would be beneficial.

Given a feedback capacitance of 383 fF, the required feedback resistance, in order to have a pole frequency lower at 20 Hz is

\[
R_{\text{max}} = \frac{1}{2\pi C_F \omega_p} = 20 \text{ G}\Omega.
\]

The overall transfer function of the sensor, defined as the ratio between the output voltage and the input pressure, is

\[
T_{\text{TOT}}(j\omega) = \frac{V_{\text{out}}(j\omega)}{P(j\omega)} = T(j\omega) \frac{\partial C}{\partial y} T_{\text{TCA}}(j\omega) = -\pi r^2 \frac{1}{k + j \omega b - \omega^2 m} \frac{\partial C}{\partial y} j \omega V_{\text{DC}} \frac{R_F}{1 + j \omega R_F C_F}.
\]
Figure 7: Magnitude of the mechanical transfer function of the sensor, from sound pressure to displacement.

Figure 8: Measured magnitude of the frequency response of a MEMS microphone.
whose magnitude, reported in Fig. 7 resembles a pass-band filter where the bass-band frequency range is, correctly, the frequency range of interest! Note that this theoretical frequency response is very similar to an experimental one: see the measured magnitude of the frequency response of a MEMS microphone shown in Fig. 8.

The total sensitivity of the sensor, defined as output voltage variation per unit input pressure in the audio range is

\[
S_{TOT} = \frac{\partial V_{out}}{\partial p} = [T_{TOT}(\omega)]_{\omega\in[20-20k]} = \frac{\partial y}{\partial p} \frac{\partial V_{out}}{\partial y} \frac{\partial y}{\partial C} = \frac{\pi r^2}{16\pi Eh^3} \frac{\varepsilon_0 \pi r^2 V_{DC}}{g^3 C_F} \frac{\varepsilon_0 A V_{DC}}{g^3}
\]

\[
S_{TOT} = 88.5 \text{ mV/Pa}.
\]

Fig. 9 reports a photograph of the membrane of a commercial MEMS microphone.
Figure 9: SEM photograph of the membrane of a MEMS microphone.