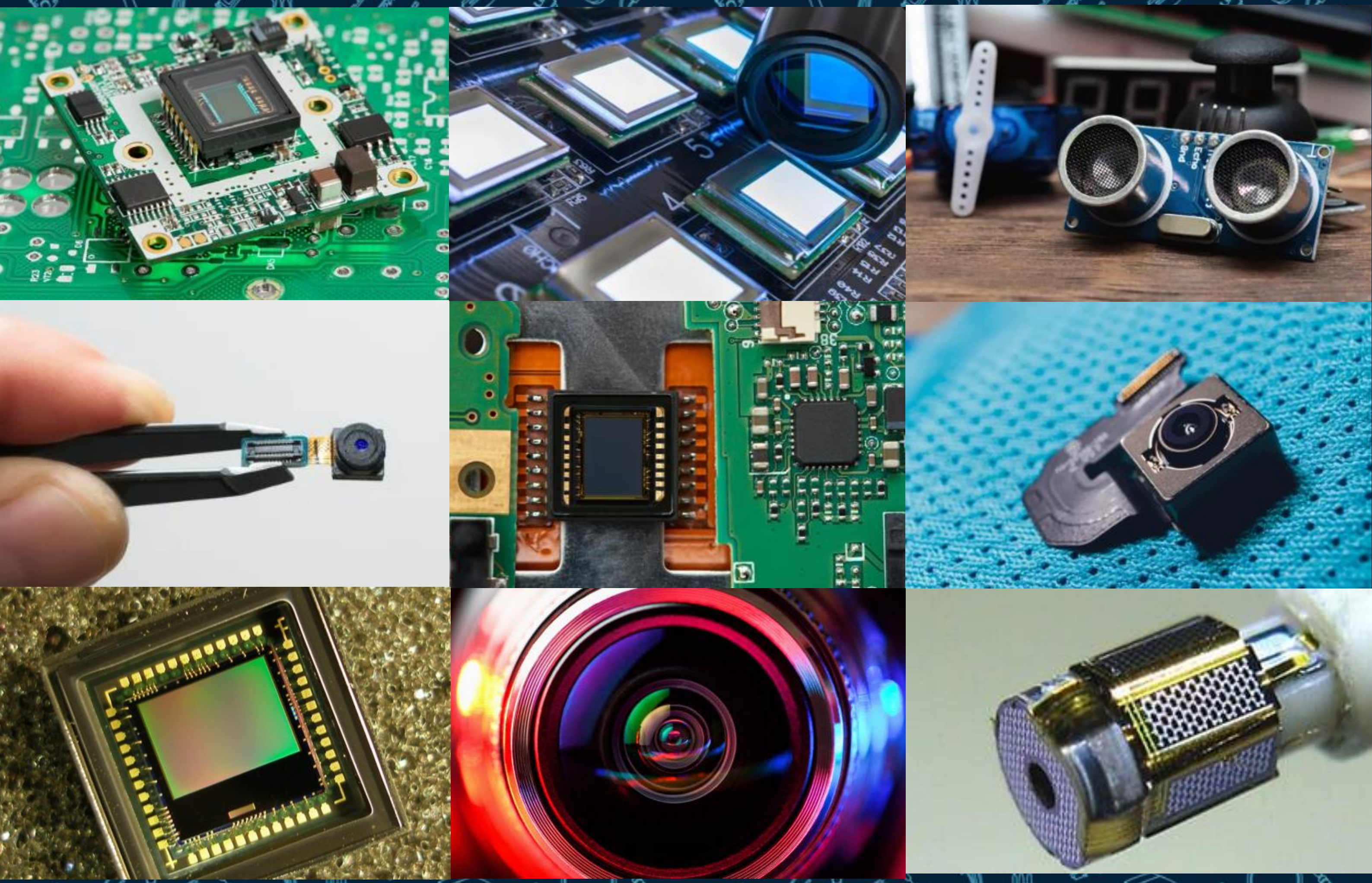
LENZI FRANCESCO – GIORGIO DONATO CARLO

MEMS AND MICROSENSORS



From Prof. G. Langfelder Lectures
Electronics Engineering
Politecnico di Milano

MEMS and Microsensors

Electronics Engineering

Lenzi Francesco, Giorgio Donato Carlo

Politecnico di Milano
Academic Year 2021-2022

Contents

1 TECHNOLOGIES FOR MICROFABRICATION OF MEMS

To understand what MEMS are, how they work, which principles they are based on, a knowledge of their fabrication process is very convenient.

Semiconductors and **dielectrics** constitute the two most heavily used material types for MEMS fabrication, owing in large measure to the ability (during the 90s) of leveraging the expertise, know-how, and physical infrastructure developed for Si and its dielectrics for the IC industry (MEMS technologies are somewhat sub-products of CMOS processes), and to their key properties (e.g. for Silicon: mechanical → brittle, electrical → good conductor, thermal → good conductor, chemical → $\mathrm{SiO_2}$. . .).

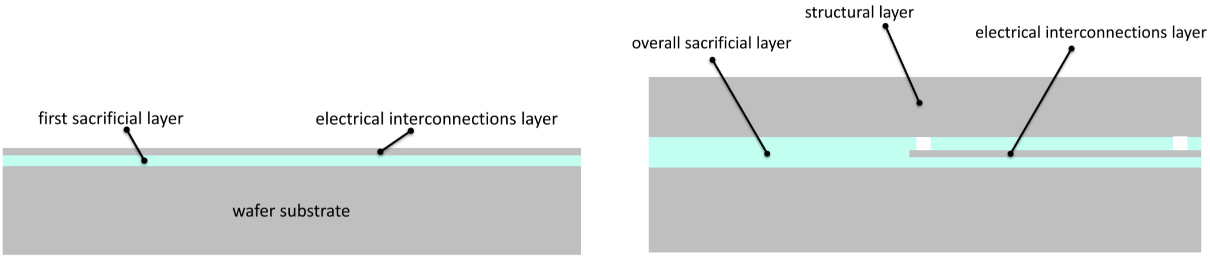
Performance of MEMS devices are often enabled/limited by their process (key parameters, repeatability, reliability. . .). To correctly design and optimize a MEMS sensor, it is thus mandatory to know its fabrication process steps. Moreover new ideas often sprout from the solution of process issues, or push for the solution of process issues.

For all these reasons, an overview of MEMS fabrication processes is useful at the beginning of the course.

1.1 A MEMS process at a glance

Several sensors work according to the following principle: the **measurement** is operated through a **displacement** caused by a **force** which is generated from a **physical quantity**. This itself is not new, but enabling the possibility to realize **moving parts at a micro-scale** makes MEMS a revolution in sensing!

MEMS are composed of a combination of a **Structural Layer** (micro-mechanical parts partially free to move) and some **Electrical Interconnections** (needed to apply and readout electric signals). This structure makes unavoidable to etch a **Sacrificial Layer** on the starting Si wafer.



(a) (b)

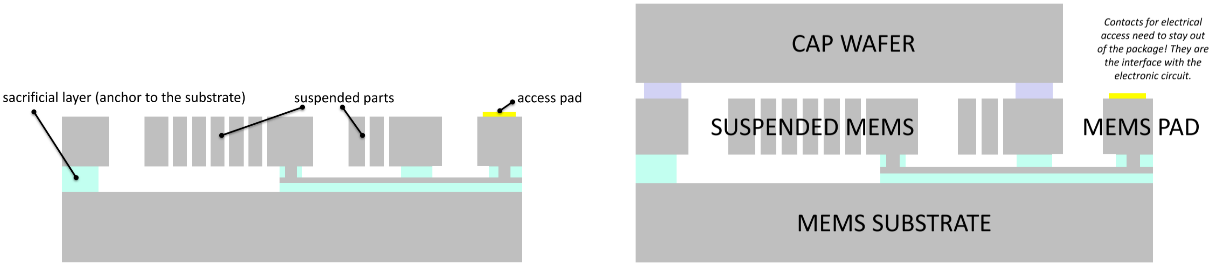
Figure 1: MEMS structure growth progression in a planar section

Thermal oxidation (high temperature and oxygenated environment) is operated on a Si substrate to create a first **"sacrificial"** SiO_2 **layer**. A thin layer of poly-silicon is deposited through Chemical Vapour Deposition[1]. It is then selectively etched to create paths, which will later form **buried electrical interconnections** to bring signals to the mechanical parts (fig. 1a).

---

[1] See paragraph 1.2.1

A **second thermal oxidation** is then operated, which increases the “sacrificial” layer thickness. It is later selectively etched to give access to electrical interconnections. On top of the sacrificial layer, the **poly-silicon “structural” layer** is grown with a certain thickness (nowadays 20-100 µm). This is the thickness that defines your structural layer height (i.e. **mass**) (fig 1b).



(a) Release of the suspended mass in a planar section

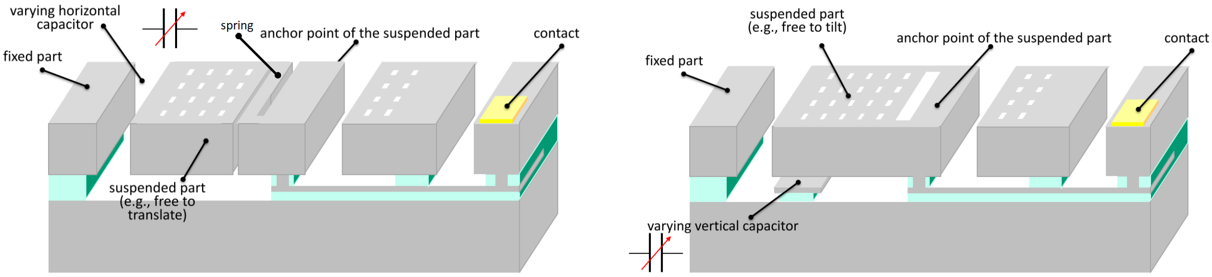
(b) Encapsulated MEMS sensor (note that the metal access pad is not under the cap).

Figure 2

As shown in Figure 2a the structural layer is **selectively removed** (in an anisotropic way) to define suspended and anchored parts. The release of the suspended structures is obtained by etching of the sacrificial layer. The etching advances also beneath anchored parts for a certain length: that is how you dimension maximum suspended widths. Lastly a metal (Al or Au) is deposited to define the **access pads**. The released MEMS device can thus move as a result of external forces (inertial, magnetic, pressure, acoustic...), or under applied electrostatic actuation. Typical (applied or readout) movements range from a few pm to tens of µm.

At the end of the fabrication process a “**cap**”, typically defined through another wafer, encapsulates the formed structures to protect them from environment, wear and debris. Note again (as shown in figure 2b) that electrical interconnections are obviously needed to apply and readout electrical signals.

In figure 3 is shown the 3D structure of the MEMS where it is visible how the suspended parts are anchored and their possible directions of motion. Note the square holes below, mandatory for the mass release.



(a) In plane capacitive MEMS sensor.

(b) Out of plane capacitive MEMS sensor.

Figure 3: 3D visualization of exemplar MEMS before packaging.

1.2 Structural layer growth and definition

The one shown so far is an overview of a so-called surface micromachining process, several other processes are derived from this one, with small differences. In general, aside from the process, an 8-inch[2] MEMS wafer can deliver up to few thousand MEMS sensors, which are then diced in small chips (dies).
Most performance of sensors fabricated in MEMS technologies are dictated by the **structural layer** and key performance parameters are represented by **thickness value and uniformity**

1.2.1 Structural layer growth

Chemical Vapour Deposition (CVD) Used for a long time in the IC industry, CVD is a process where a thin layer is formed by deposition of vapor-phase components onto a heated substrate. The vapor contains the constituent gases of the thin film. These precursor gases are introduced in the reactor in a regulated manner, to control gas mix and pressure. As shown in figure 4 the reactor consists of a horizontal tube, sized to accommodate a large number (÷50) of large-area (200 mm diameter) Si substrates. Vacuum seals mounted on each reactor enable operation at $0.1 \div 0.5mbar$. The chamber is kept at ÷500°C by a large resistive heater that envelops the reactor. **Growth rate is slow** $(5 \div 10\frac{nm}{min})$, and max obtainable thickness is limited to 1-2 µm only (this was not an issue for CVD as long as MEMS were not invented. . .).

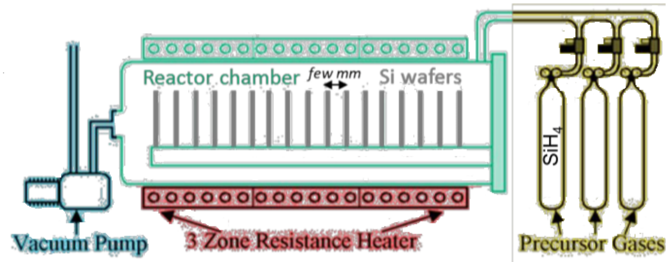


Figure 4: A sketch of a chemical vapour deposition reactor.

Epitaxial growth Some devices (e.g. inertial ones) benefit from thickness that cannot be achieved by "conventional" CVD. For such cases, epitaxial reactors were developed to grow **thick films**. Unlike CVD processes (typical deposition rates $< 10\frac{nm}{min}$), epitaxy has deposition rates of $\div 1\frac{\mu m}{min}$. This high rate results from the **higher substrate temperature** (> 1000°C) and **deposition pressures** $(> 60mbar)$. With this process thicknesses up to $> 30 - 60\ \mu m$ are nowadays commonly obtained. Nonuniformities of $\pm 1\mu m$ across the wafer are however common. Note that epitaxy is a "special case" of CVD-film growth where or a crystalline thin-film is grown upon a crystalline substrate such that the structure of the film is formed using the structure of the substrate as a "template" or, if surface migration[3] is inhibited (not enough time), then growth likely results in the formation of polycrystalline grains.

---

[2] Standard Si wafer size

[3] Silicon migration is a diffusive process that increases symmetry and smoothness of silicon surfaces as atoms flow into lower energy configurations.

Doping Epitaxial growth supports **"in situ" doping** using precursor gasses like PH_3 (phospine) and B_2H_6 (diborane) when used in combination with SiH_4 (silane). This technique allows to obtain conductive films with uniform doping profiles, avoiding the need of additional doping steps. Doping is important because, when exploiting electrostatic forces to drive or sense suspended masses, we need the electrostatic potential across the whole mass to be uniform. Hence, the structural layer needs to have high conductivity, whatever its type (P, N); therefore, it is convenient to uniformly dope it during the growth, at relatively high level ($> 10^{17} \div 10^{20} cm^{-3}$)

1.2.2 Structural layer definition - etching

The grown structural layer must be defined in shape. We need to define:

- A frame of the **suspended masses**;
- The shape of the **suspension springs**;
- Some **capacitive plates** for electrostatic actuation/sensing;
- A large amount of **holes** for the correct release.

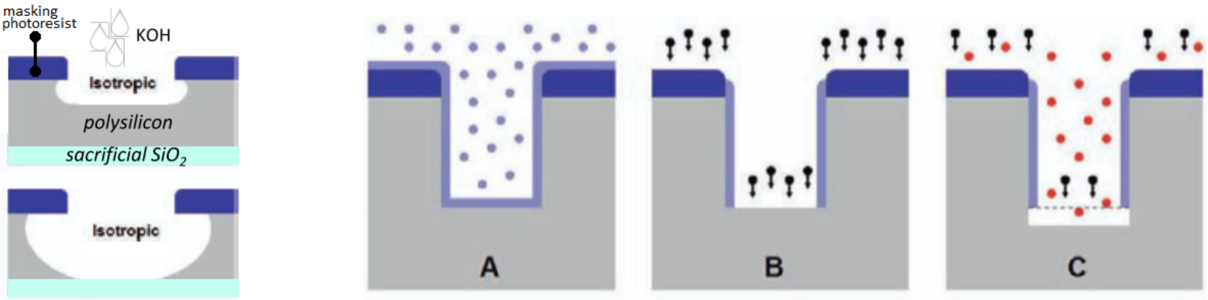
It is thus necessary to etch ("**dig & remove**") poly-silicon in a controlled manner, as the following parameters affect sensor performance:

- Sidewalls **orthogonality** in capacitor's gaps;
- The **gap size** between fixed and structural parts;
- The minimum feature (**dimension**) of structural parts (e.g. springs).

Wet etching In wet etching, after a masking layer is coated and patterned on the wafer, the wafer is immersed into a solution specific to the material to be chemically etched (e.g. potassium hydroxide, KOH). As shown in figure 5a, wet etching of poly-silicon is isotropic: the etch rate is the same in all directions. This is not helpful when we need to define gaps of known (possibly vertical) geometry. The **aspect ratio** (i.e. vertical etching over horizontal etching) is poor.
This etching technique was quite used at the beginning of the MEMS era. However, now it is much less used because it cannot achieve gap orthogonality.

Dry etching: Deep Reactive Ion Etching (DRIE) Dry etching exploits a combination of chemical etching (but now in vapor or plasma phase) and physical etching. Modern DRIE tools use three distinct phases which are iterated until desired etch depth is achieved:

- **Passivation cycle**: C_4F_8 plasma in the chamber deposits a uniform protective polymer on all surfaces; (fig. 5b A)
- **Intermediate cycle**: the accelerated ions directionally remove the passivation layer from the base of the trenches, but not from the sidewalls; (fig. 5b B)
- **Etch cycle**: SF_6 plasma in the chamber attacks the exposed silicon areas at the base of the trenches; (fig. 5b C)



(a) Wet etching is **isotropic**.

(b) Deep reactive ion etching can achieve **orthogonal gaps**.

Figure 5: Etching processes.

Among dry etching techniques, DRIE revolutionized the fabrication of micromachined inertial devices, allowing to etch deep and high (e.g. 20) aspect ratio trenches with vertical sidewalls. With deep reactive ion etching, typically achievable gaps are in the order of $1\mu m$, with nonuniformities in the order of $\pm 0.1\mu m$, mantaining vertical and almost straight sidewalls: as shown in figure 6 some unavoidable, but overall negligible, scallops are produced.

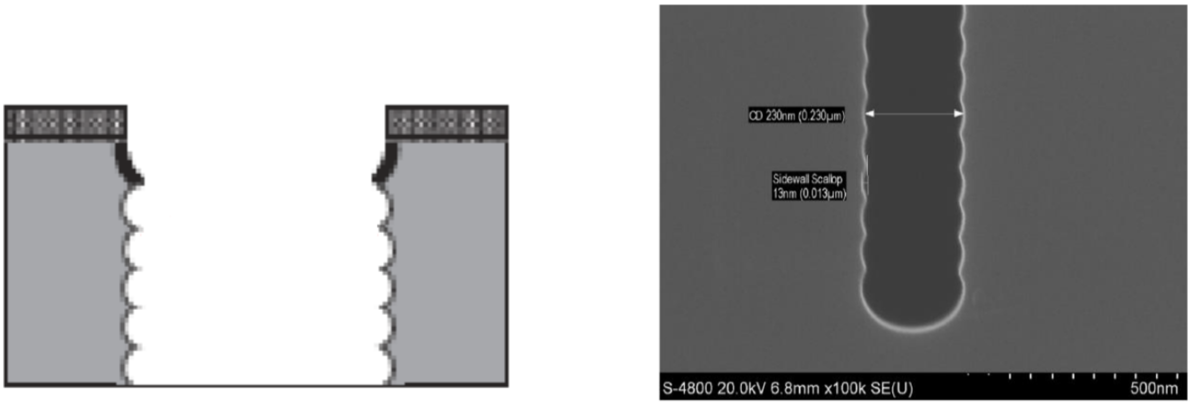


Figure 6: Residual scallops from deep reactive ion etching process.

Structural parts release Deep reactive ion etching stops when the SiO_2 layer is reached. At this point the wafer is put in a gaseous environment (typically hydrofluoric acid, HF, or XeF_2), which begins removing SiO_2. The gas reaches SiO_2 surface through the trenches formed by DRIE (including holes of perforated masses) and starts removing the SiO_2 advances from exposed surfaces towards regions beneath polySi. For a **release distance** d_{rel}, all structural parts having a width lower than $2d_{rel}$ (usually about 10 µm) are completely released, suspended over the substrate. Where the width is larger than $2d_{rel}$, the layer remains anchored (electrodes, anchor points) as shown in figure 7.

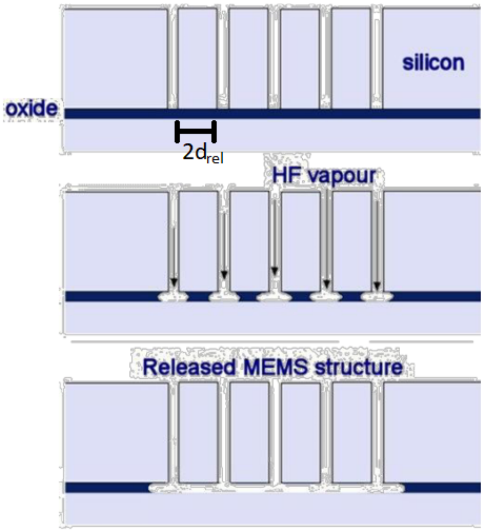


Figure 7: MEMS structure release.

1.3 MEMS packaging

The performances of a MEMS system depend as much on the packaging as they do on the sensor itself. Packaging is needed for:

- **Protection** against debris, dust...;
- Creation of the correct **operative pressure** (typically $100mbar \div 50\mu bar$);
- Creation of **non-reactive environment** (no oxygen, ok nitrogen or noble gases only).

Packaging also influence the **Quality factor Q**[4] (and its repeatability from part to part) and **noise** (Brownian noise is a function of pressure) values.
We discuss first the "zero-level" packaging, i.e. of the MEMS only without its coupling to the integrated circuit (IC).

1.3.1 Wafer bonding techniques

Glassfrit bonding It consists of a patterning of low-melting-point glass (a powder) on the wafer which melt and fuses at $<$ 450°C (below melting point of Al connections). If cooling profiles are used during the process, the hermetic sealing between wafers is optimal. The major drawback of this technique is area and volume occupations: typical height is $20 \div 30$ μm, and widths are $250\mu m$[5] which makes also arise the cost.

---

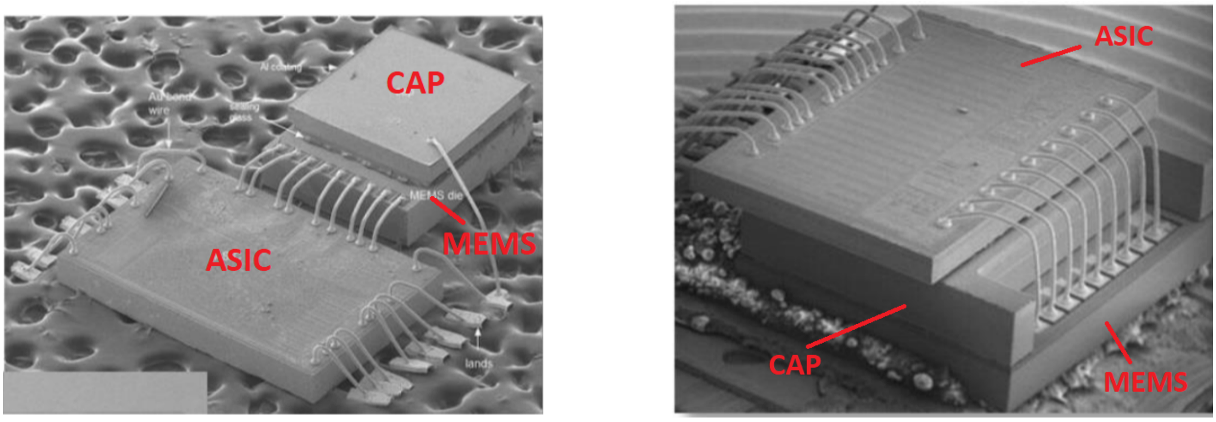
[4] We will largely discurss this parameter later in the course.

[5] compare it to typical accelerometer sizes of 300x400 μm^2!

Eutethic bonding It is done by creating an alloy between two materials (e.g. Au-Si or Al-Ge) at high temperature (however lower than the melting T of each of the components). Melting is obtained through mechanical pressure: e.g. a layer of Au is placed onto a Si surfaces, and a suitable force is applied and in the meanwhile, the system is heated. This process has the advantages of having small bonding ring widths ($\div 50\mu m$), an optimal bonding contacts conductivity and an easy distribution of interconnections on the cap.
Bonding rings can be also used to create separate cavities for devices operating at different pressures on the same dice.

1.3.2 MEMS integration

The integration of the MEMS sensors on the chip is important to achieve an optimal design. Indeed, each process step dedicated to MEMS (or ASIC) is limited also by the constraints given by the ASIC (or MEMS) this in the end limits the optimization capabilities. First MEMS accelerometers had structural parts and readout electronics on the same chip (examples from Analog Devices), but currently there are also different integration techniques. As shown in figure 8, MEMS sensor are often integrated as a **System in Package** (SiP). This kind of integration grants two major advantages: firstly each chip (MEMS and ASIC) can be individually optimized, with less process constraints; and secondly an evolution of the ASIC can be directly applied to existing MEMS without the need of a complete re-design (and vice-versa).



(a) Lateral integration.

(b) Stacked integration.

Figure 8: ASIC+MEMS integrations. Those can be also combined when using more MEMS sensors.

2 SPRING-MASS-DAMPER SYSTEM

In inertial MEMS sensors, there is a need to measure the motion of objects (translations and rotation) induced by external actions. Though obvious, it is worthwhile to give a complete description of the reference equations describing the system and to fix, once and for all, a formalism (the formalism will be here initially simplified to the case of translations only).
We will then derive the linear **Spring-Mass-Damper Model** (Mathematical expression at the basis of MEMS studies), both in the time and frequency domains: it is the physical model at the basis of the solution of any problems related to MEMS. The model will be also quickly extended to the torsional case.
Finally, it will be completed with electrostatic forces, which are unavoidable during MEMS actuation or sensing.

2.1 Kinematics of relative motion

An **Inertial Reference Frame** (system) is a frame where the first Newton's law applies: an object moves at a constant velocity, if not perturbed by an external force. All inertial reference frames are in a state of constant, rectilinear motion (they are not accelerating). Therefore, physical laws take the same form in all inertial frames. In a non-inertial reference frame, the laws of physics depend upon the particular frame of reference, and the usual physical forces must be supplemented by "fictitious" or "apparent" forces.
Considering two reference frames and a point like mass, we could define the following quantities:

- **Oxyz** : Inertial (absolute) reference frame;
- **O'x'y'z'** : non-inertial (relative) reference frame;
- $\mathbf{r_{O'a}}$: vector describing the position of O' w.r.t. the absolute reference;
- $\mathbf{r_{Pr}}$: vector describing the point P w.r.t. the relative reference;
- $\mathbf{r_{Pa}}$: vector describing the point P w.r.t. the absolute reference;
- $\mathbf{a_{O'a}}$: vector describing the acceleration of the relative reference;
- $\mathbf{\Omega_{O'a}}$: vector describing the angular velocity of the relative reference.
- $\mathbf{v_{Pr}}$ $\mathbf{a_{Pr}}$ $\mathbf{v_{Pa}}$ $\mathbf{a_{Pa}}$: vector describing the velocity and the acceleration related respectively to the inertial frame and the absolute frame.

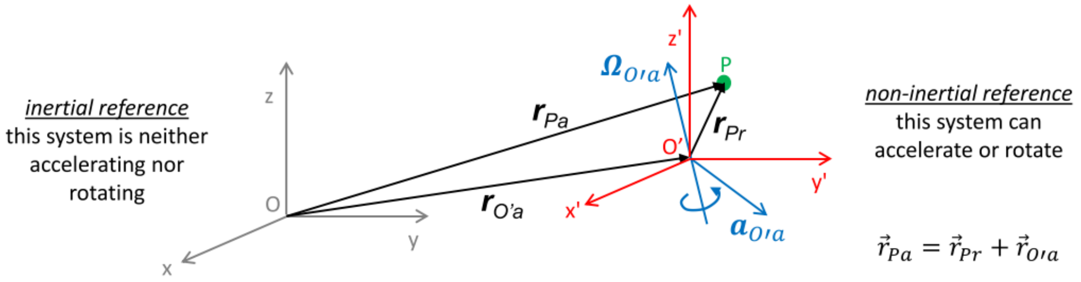


Figure 9: Relation between inertial and non-inertial system

We can now write the expressions of **position** $\vec{r}_{Pa}$, **velocity** $\vec{v}_{Pa}$ and **acceleration** $\vec{a}_{Pa}$ the point P with respect to the absolute reference system. In the most general situation, we need to take into account the translational velocity, the translational acceleration, the angular velocity, the angular acceleration and the Coriolis acceleration.

$$\begin{cases} \vec{r}_{Pa} = \vec{r}_{Pr} + \vec{r}_{O'a} \\ \vec{v}_{Pa} = \vec{v}_{Pr} + \vec{v}_{O'a} + (\vec{\Omega}_{O'a} \times \vec{r}_{Pr}) \\ \vec{a}_{Pa} = \vec{a}_{Pr} + \vec{a}_{O'a} + \cancel{(\overrightarrow{\dot{\Omega}_{O'a}} \times \vec{r}_{Pr})} + \cancel{\vec{\Omega}_{O'a} \times (\vec{\Omega}_{O'a} \times \vec{r}_{Pr})} + 2(\vec{\Omega}_{O'a} \times \vec{v}_{Pr}) \end{cases}$$

The last equation shows us that the "true" Newton acceleration is obtained by the sum of the acceleration seen in O', the acceleration of O', the angular acceleration (<u>negligible</u>), centripetal acceleration (<u>negligible</u>), *Coriolis* acceleration.
The goal is to measure the motion of O'x'y'z' (i.e. of our non-inertial system) with respect to Oxyz (i.e. the reference frame). To do it, we exploit the motion $r_{Pr}(t)$ of the MEMS mass P relative to O'x'y'z', described through fictitious forces. Obviously, the MEMS mass is not floating and disconnected from the package, it is suspended through elastic means and encapsulated in a gaseous package.

This, give us the true forces to consider for the point-mass P:[6]

- the restoring **elastic force** : $F_{elastic} = -k \cdot r_{Pr}$
- the restoring **damping force** : $F_{damping} = -b \cdot v_{Pr}$

2.2 Linear spring-mass-damper system (1-D case)

We could pass to the 1-D case substituting $\vec{r}$ with $\vec{x}$. This is a real case, as several MEMS sensor are designed to be sensitive to a single direction.[7]

$$m\vec{a}_{Pa} = m\vec{a}_{Pr} + m\vec{a}_{O'a} + 2m(\vec{\Omega}_{O'a} \times \vec{v}_{Pr})$$
$$m\vec{\ddot{x}}_{Pa} = m\vec{\ddot{x}}_{Pr} + m\vec{a}_{O'a} + 2m(\vec{\Omega}_{O'a} \times \vec{\dot{y}}_{Pr})$$
$$-k\vec{x}_{Pr} - b\vec{\dot{x}}_{Pr} = m\vec{\ddot{x}}_{Pr} + m\vec{a}_{O'a} + 2m(\vec{\Omega}_{O'a} \times \vec{\dot{y}}_{Pr})$$

---

[6] Note that those forces do happen in the relative frame.

[7] Note that for the Coriolis force to be on the x direction, the motion has to be on the y axis.

Rearranging the terms we get:

$$m\ddot{\vec{x}}_{Pr} + b\dot{\vec{x}}_{Pr} + k\vec{x}_{Pr} = -m\vec{a}_{O'a} - 2m(\vec{\Omega}_{O'a} \times \dot{\vec{y}}_{Pr})$$

$$\mathbf{m\ddot{x} + b\dot{x} + kx = F_{inertial}}$$

2.3 Frequency description

For the description of the dynamic behavior of the MEMS, it is interesting and relevant to apply the Laplace transform to the found equation, in order to evaluate the transfer function between a generic external force and the displacement, in terms of both modulus and phase, as a function of the frequency of that force.

$$m\ddot{x} + b\dot{x} + kx = -ma_{ext} = F_{ext} \quad \longrightarrow \quad mX(s)s^2 + bX(s)s + kX(s) = F_{ext}(s)$$

Where:

- $\mathbf{X(s)}$ is the Laplace transform of the relative position between the point-like mass and the non-inertial frame;
- $\mathbf{F_{ext}}$ is the Laplace transform of the external force applied to the non-inertial frame.

Let's find the **Transfer Function $\mathbf{T_{XF}}$**:

$$X(s)(ms^2 + bs + k) = F_{ext}(s) \quad \to \quad X(s) = \frac{1/m}{(s^2 + \frac{b}{m}s + \frac{k}{m})}F_{ext}(s)$$

$$\implies T_{XF} = \frac{1/m}{s^2 + \frac{b}{m}s + \frac{k}{m}}$$

We highlight the dependence on two characteristic parameters of the system, which are the resonance (angular) frequency ω_0 and the quality factor Q.

$$\omega_0 = \sqrt{\frac{k}{m}} \qquad\qquad Q = \omega_0\frac{m}{b}$$

$$\implies T_{XF} = \frac{1/m}{s^2 + \frac{\omega_0}{Q}s + \omega_0^2}$$

The system has a low-pass 2^{nd} order transfer function with two singularities which can be either complex conjugate, or coincident real, or split real poles.

$$\Delta = \left(\frac{\omega_0}{Q}\right)^2 - 4\omega_0^2$$

- Split real poles: $\Delta > 0 \implies Q < 0.5$
- Coincidence real poles: $\Delta = 0 \implies Q = 0.5$
- Complex conjugate poles: $\Delta < 0 \implies Q > 0.5$

Below is a plot of the transfer function modulus for various Q:

$$|T_{XF}(j\omega)| = \frac{1/m}{\sqrt{(\omega_0^2 - \omega^2)^2 + \left(\frac{\omega\omega_0}{Q}\right)^2}}$$

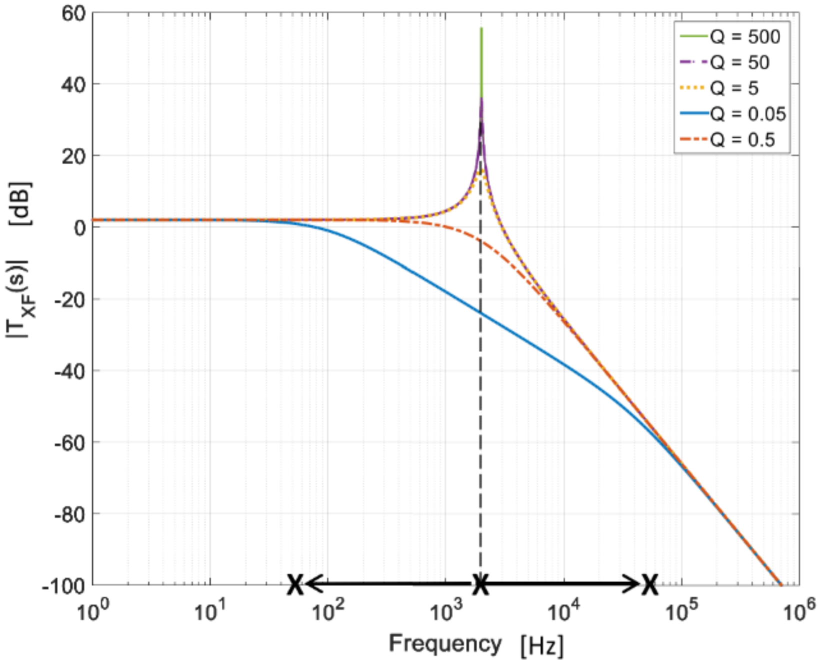


Figure 10: Bode Diagram of transfer function modulus

In particular:

- $\omega \ll \omega_0 \implies |T_{XF}(j\omega)| = \frac{1/m}{\omega_0^2} = \frac{1}{k}$
- $\omega = \omega_0 \implies |T_{XF}(j\omega)| = \frac{Q/m}{\omega_0^2} = \frac{Q}{k}$
- $\omega \gg \omega_0 \implies |T_{XF}(j\omega)| = \frac{1/m}{\omega^2}$

The phase decreases by 180°, as the system is characterized by two poles. The phase shift at the resonance frequency is exactly 90°.

$$\phi[T_{XF}(j\omega)] = arctan\left|\frac{Im\,(T_{XF}(j\omega))}{Re\,(T_{XF}(j\omega))}\right|$$

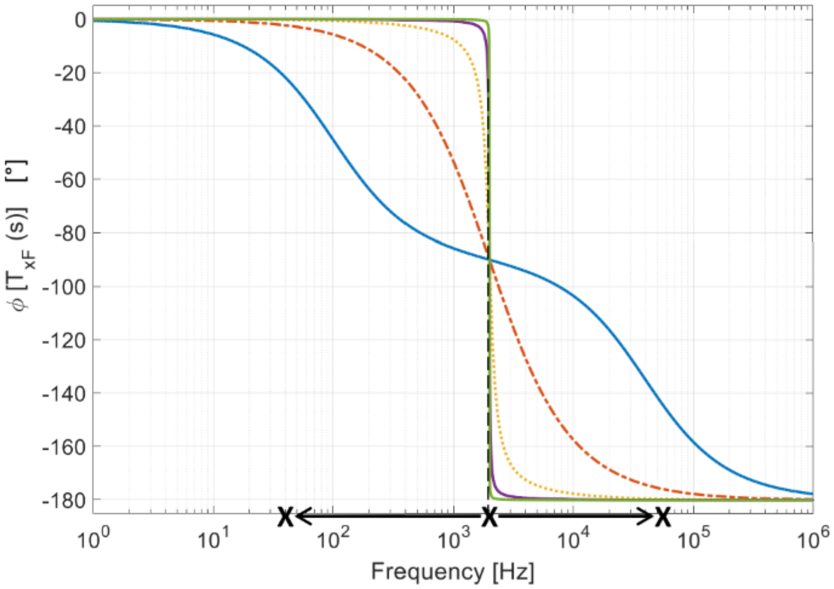


Figure 11: Bode Diagram of transfer function phase

Please note that the larger is the quality factor, the steeper is the phase transition.

2.3.1 Operating Regions

MEMS devices operate in different regions of the transfer function:

- **Accelerometers**, microphones and pressure sensors typically operate under forces occurring far before resonance; as we will see, they usually have relatively low Q factors (typically < 10, or even < 1);
- **Resonators** (including the gyroscopes drive) operate at resonance (few tens kHz to few MHz) and require high quality factors (typically few thousand to ten thousand);
- Other sensors (**gyroscopes**, magnetometers. . .) operate slightly before the resonance frequency (off-resonance or mode-split operation), due to a modulation in frequency of the applied forces;
- No devices operate beyond the resonance frequency.

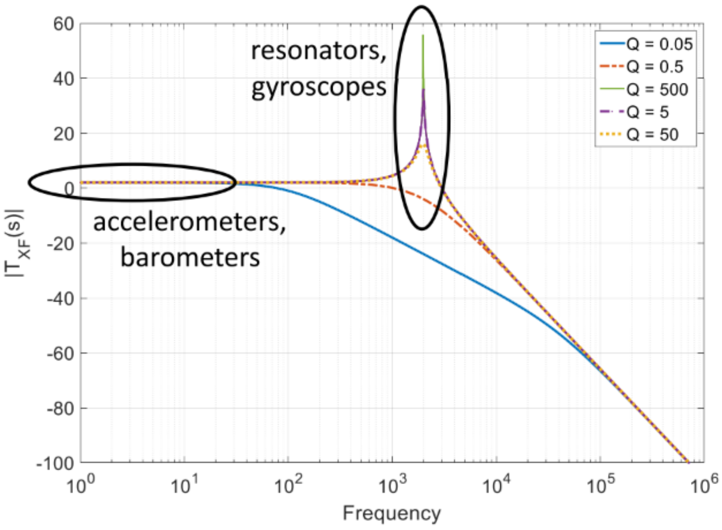


Figure 12: Operating regions of MEMS Sensors

2.3.2 Step responses at different quality factor

The different **quality factor** implies not only a different response amplification at resonance, but also a different response to pulses or steps (both waveform types ideally include all the frequencies).

For **under-damped** systems, the time constant is given by $\tau = \frac{Q}{\pi f_0}$, so it increases linearly with Q.
For **over-damped** systems, the time constant is dominated by the first pole of the T_{FX}.

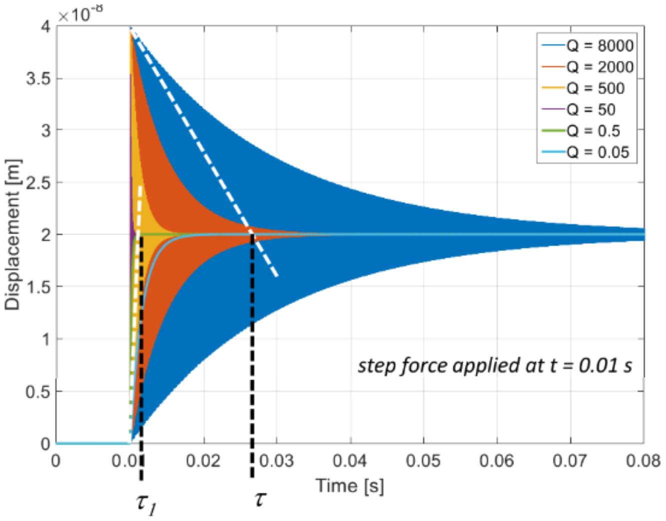


Figure 13: Step response at different Q

2.4 Torsional Spring-Mass-Damper-System

As we will study in detail during the description of the individual sensors, some MEMS operate through torsion of structural elements, rather than through deflection.
Therefore, the torsional spring-mass-damper system should be considered in those cases:

- displacement x → angle θ;
- mass m → moment of inertia I;
- force F → torque M (obtained as the force applied in the gravity center times the arm R between the rotation hinge and the gravity center).

$$\omega_0 = \sqrt{\frac{k_\theta}{I}} \qquad Q = \omega_0 \frac{I}{b_\theta}$$

The stiffness and damping coefficient become their torsional counterparts (and note, consequently, the change in their unit).
All the equations are just extensions of the linear to the torsional case:

$$I\ddot{\theta} + b_\theta\dot{\theta} + k_\theta\theta = M_{ext}$$

The moment of inertia, also known as the angular mass or rotational inertia, of a rigid body determines the torque needed for a desired angular acceleration about a rotational axis.
It depends on the body's mass distribution around the chosen rotation axis, with larger moments requiring more torque to change the body's rotational state.

2.5 Electrostatics forces

The force balance defined so far includes the elastic and damping forces, as well as the external action to be measured.

$$m\ddot{x} + b\dot{x} + kx = ma_{ext} = F_{ext}$$

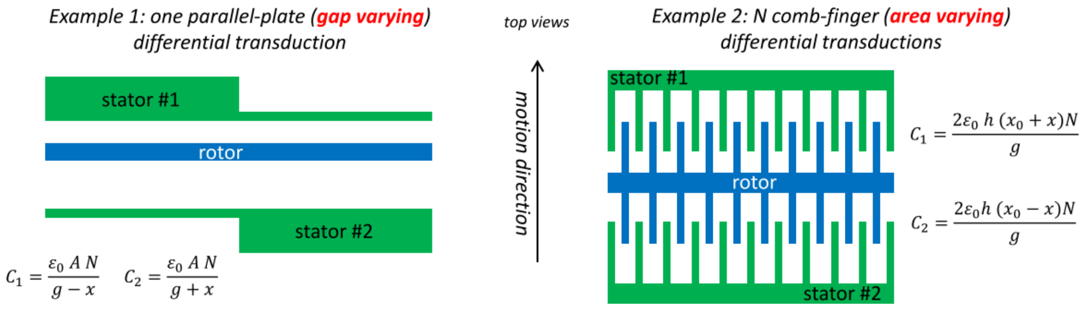
However, in the most general and common situation, there are also electrostatic forces that take part in the force balance:

- when the MEMS is used as a **capacitive actuator**, a static/dynamic voltage is applied to generate a force (not the case in accelerometers);
- when the MEMS is used as a **capacitive sensor**, a voltage is again applied to readout the value of the capacitance (see e.g. the next class).

The used *variable capacitance* can be of **area-varying**(comb-finger) type or **gap-varying**(parallel-plate).

$$C = \frac{\varepsilon_0\, A\, N}{g}$$

In analogy with the "macro" world of electrical motors, moving parts of micro-systems are called "**rotors**" even if they do not rotate. Fixed electrodes are, again from the analogy above, called **stators**.



2.5.1 Parallel-plate configuration

Let us consider a single-ended parallel-plate capacitor and let us assume a DC voltage applied to its plates. Using the virtual works principle, we can compute the value of the electrostatic force: in equilibrium conditions, electrostatic forces balance the mechanical force.
Any variation of the energy E_c stored in the capacitor should be given by the work W done by the mechanical or electrical forces.

$$dE_C = dW_{mech} + dW_{elec}$$

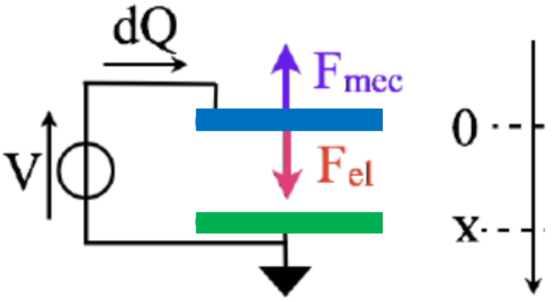


Figure 14: Single-ended parallel-plate capacitor

We know that the electrostatic energy stored in a capacitor is:

$$dE_C = d\left(\frac{1}{2}CV^2\right) = \frac{V^2}{2}dC$$

The work done for a displacement dx is:

$$dW_{mech} = -F_{mech}dx$$

The electrostatic work done to change the voltage over a capacitor is:

$$dW_{elec} = V\; dQ = V\; d(CV) = V(C\; dV + V\; dC) = V^2 dC$$

Therefore we could write:

$$\frac{V^2}{2}dC = -F_{mech}dx + V^2dC \longrightarrow F_{mech}dx = \frac{V^2}{2}dC$$

Knowing that the derivative of the capacitance is:

$$dC = d\left(\frac{\varepsilon_0\, A\, N}{g-x}\right) = \frac{\varepsilon_0\, A\, N}{(g-x)^2}dx$$

Putting together all the equations developed so far we obtain the equilibrium solution:

$$F_{mech}dx = \frac{V^2}{2}dC = \frac{V^2}{2}\frac{\varepsilon_0\, A\, N}{(g-x)^2}dx \implies F_{elec} = F_{mech} = \frac{V^2}{2}\frac{\varepsilon_0\, A\, N}{(g-x)^2}$$

The situation can be easily extended by considering the **differential** configuration with opposite forces:

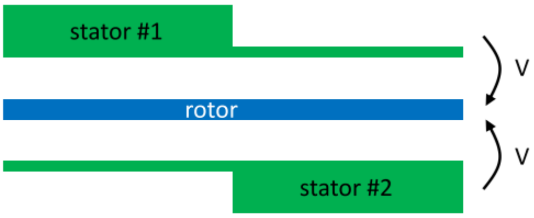


Figure 15: Differential parallel-plate capacitor

$$F_{elec} = \frac{V^2}{2}\frac{\varepsilon_0\, A\, N}{(g-x)^2} - \frac{V^2}{2}\frac{\varepsilon_0\, A\, N}{(g+x)^2}$$

Note that the applied force is a function of the displacement itself, both in single-ended and in differential configurations.

2.5.2 Comb-finger configuration

Let us consider a single-ended comb-finger capacitor. Let us assume a DC voltage applied to its plates. Using again the same approach, we evaluate the electrostatic force: in equilibrium conditions, electrostatic forces balance the mechanical force.
Comb-finger capacitors are of the area-varying type:

- the facing area is $A = h \cdot (x_0 + x)$, h being the process height and x_0 the facing length at rest;
- for N fingers on the rotor, a factor 2 should be added as the facing is on two sides.

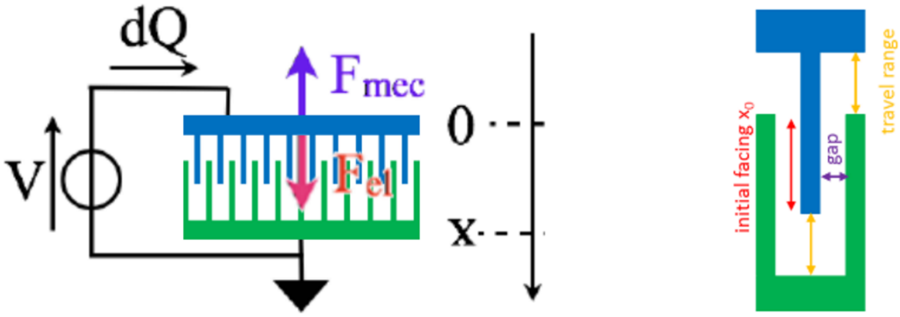


Figure 16: Comb-finger configuration capacitance

$$C = \frac{2\,\varepsilon_0\;A\,N}{g} = \frac{2\,\varepsilon_0\;h\,(x+x_0)\;N}{g}$$

$$\implies\;F_{elec} = \frac{V^2\,\varepsilon_0\,h\,N}{g}$$

Note that, in both types of capacitive sensing (area or gap varying), the differential configuration is advantageous confronted to the single ended one because:

- The effect of electrostatic forces is minimized;
- The linearity of the response in improved;
- Common mode disturbances are better rejected.

3 ACCELEROMETER

Accelerometers are the "most intuitive" MEMS devices and were the first MEMS sensors to be commercialized in automotive (airbag activation, 1993), electrical appliances (washing machines, 2001), and consumer markets (gaming joysticks, 2006).
In several MEMS, the displacement induced by the force to quantify (i.e. in this case the inertial acceleration force) is measured via parallel-plate (gap-varying) capacitive variations.

3.1 General Architecture of MEMS acceleration

Basic structural elements of a MEMS accelerometer (axel) are:

- suspended **proof mass** (or seismic mass, frame, rotor or what else...);
- suspending **springs** (anchored to the substrate at one end);
- **rotor** plates or **electrodes**;
- **stators** plates or contacts or electrodes (often – if not always – differential).

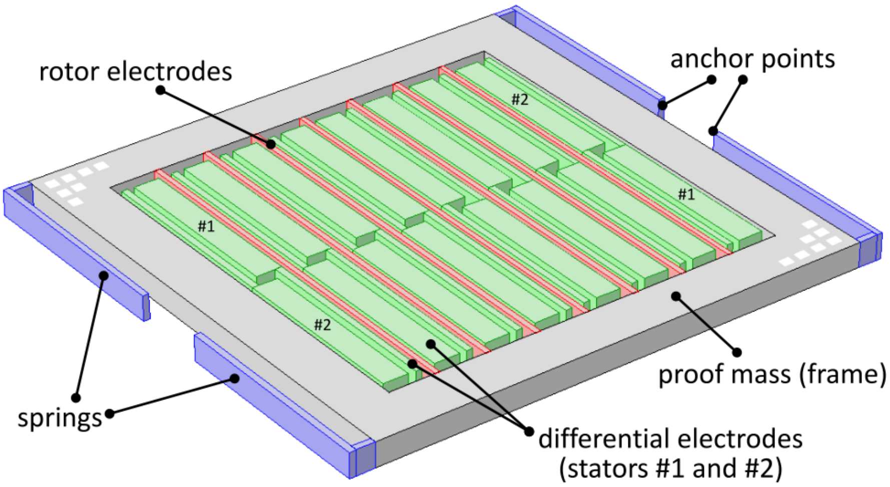


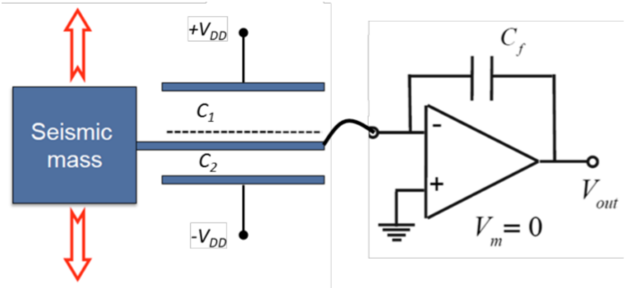
Figure 17: Architecture of MEMS Accelerometer

In order to readout a signal from an accelerometer, we need to capacitively measure its displacement to recover the information on accelerations.
Note that the capacitance is **variable**! The general expression of the current through such a capacitor, biased by a generic voltage V, is:

$$i_c = \frac{dQ}{dt} = \frac{d(CV)}{dt} = C\,\frac{dV}{dt} + V\,\frac{dC}{dt}$$

3.2 Parallel Plate (PP) readout configuration

- **Application of a constant DC voltage ±VDD between each stator and the rotor.**

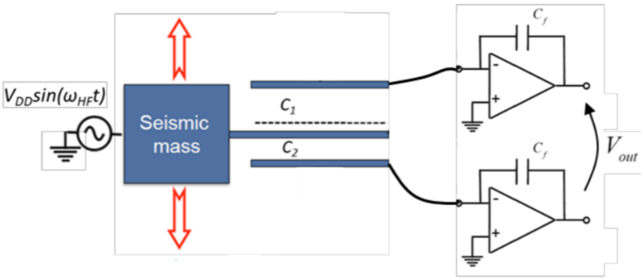


The current that flows through each capacitor is:

$$i_{C_i} = \frac{dQ_i}{dt} = \cancel{C_i \frac{dV_i}{dt}} + V_i \frac{dC_i}{dt} = V_i \frac{dC_i}{dt}$$

This term is proportional to the capacitance derivative: not to lose the stationary (DC) value of the acceleration, we need to integrate the signal. The feedback impedance is capacitive.

- **Application of a modulated AC voltage between each stator and the rotor.**



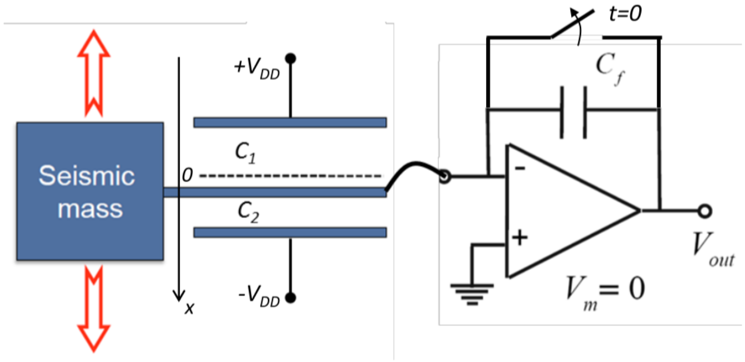
The current that flows through each capacitor is:

$$i_{C_i} = \frac{dQ_i}{dt} = C_i \frac{dV_i}{dt} + \cancel{V_i \frac{dC_i}{dt}} = C_i \frac{dV_i}{dt}$$

This term dominates for very-high frequency AC signals (the derivative is indeed proportional to the AC frequency). Any feedback impedance is okay.

3.2.1 Example of a differential capacitance sensing

Assume that the accelerometer is initially in the rest position and that C_f is initially discharged (e.g. via a switch, as shown):



$$V_m = 0 \qquad C_1 = \frac{\varepsilon_0\ A\ N}{g+x} \qquad C_2 = \frac{\varepsilon_0\ A\ N}{g-x} \qquad Q_F = 0$$

If the position changes by a quantity x, we have opposite changes in the value of the differential capacitances:

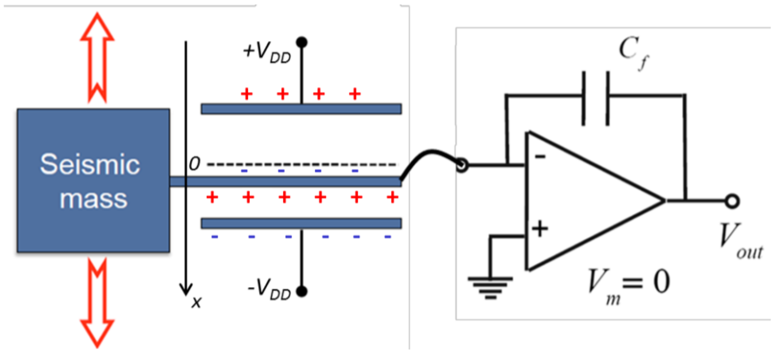
- for x > 0 (as in the figure) → C_1 decreases and C_2 increases;
- for x < 0 → C_1 increases and C_2 decreases.

The charge on the rotor has the sign in the figure and is given by:

$$Q_1 = -C_1\ V_{DD} \qquad\qquad Q_2 = C_2\ V_{DD}$$

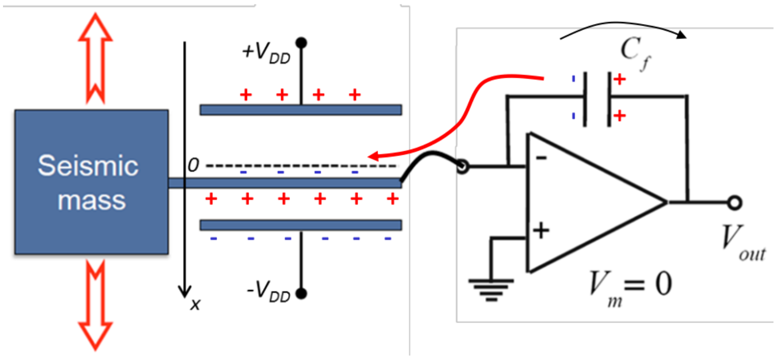
While the values for $x = 0$ are equal and opposite (**charge neutrality**), for positive displacements a net charge on the rotor appears:

$$\Delta Q_m = -C_1\ V_{DD} + C_2\ V_{DD} > 0$$



As the OpAmp input is an ideal high impedance, the charge (i.e. the current flow) is delivered through the feedback capacitance. As a consequence, the output voltage of the operational amplifier changes by a quantity:

$$\Delta V_{out} = \frac{\Delta Q_m}{C_f} = \frac{-C_1\, V_{DD} + C_2\, V_{DD}}{C_f} = \frac{V_{DD}}{C_f}\Delta C_{diff}$$



Let's calculate ΔQ_m as a function of x:

$$\Delta Q_m = -C_1\, V_{DD} + C_2\, V_{DD} = V_{DD}\, \Delta C_{diff} = V_{DD}\left(-\frac{\varepsilon_0\, A\, N}{g+x} + \frac{\varepsilon_0\, A\, N}{g-x}\right) =$$

$$V_{DD}\frac{\varepsilon A N}{g}\left(-\frac{1}{1+x/g} + \frac{1}{1-x/g}\right) = V_{DD}C_0\left(-\frac{1}{1+x/g} + \frac{1}{1-x/g}\right) =$$

$$= V_{DD}C_0\left[\frac{-1+x/g+1+x/g}{1-(x/g)^2}\right] = V_{DD}C_0\left[\frac{2\,x/g}{1-(x/g)^2}\right]$$

For small displacement x of the suspended mass with respect to the air gap g between parallel plates:

$$\text{for } x \ll g \implies \frac{x}{g} \ll 1 \implies \Delta Q_m = 2V_{DD}C_0\frac{x}{g} \implies \Delta C_{diff} = 2C_0\frac{x}{g}$$

The output voltage can be now calculated as a function of the displacement x generated by the external acceleration:

$$\Delta V_{out} = \frac{\Delta Q_m}{C_f} = \frac{\Delta C_{diff} V_{DD}}{C_f} = 2\frac{V_{DD}}{C_f}\frac{C_0}{g}x$$

The charge amplifier output voltage is, for small displacement x from the equilibrium position, a linear function of the displacement.

3.2.2 Parasitic capacitances affecting rotors (and stators)

Several are the sources of parasitic capacitance between the rotor (or stators) and ground:

- direct facing towards the **substrate**;
- facing of the **path** to the pad towards the substrate;

- facing of the **pad** itself towards the substrate;
- bonding **wires** to the ASIC pad;
- **ASIC pad parasitic** to ground;
- other parasitic towards the PCB ground. . .

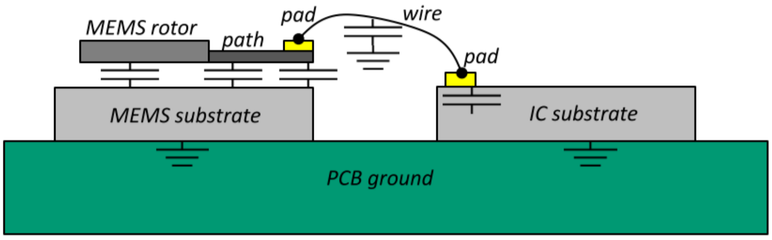


Figure 18: Parasitic capacitances affecting rotors

Even if C_{par} is very large, there is no charge flow through this capacitance (and thus no loss of signal), because it is continuously kept between ground and another fixed potential (virtual ground). In other words, the charge on C_{par} never changes.

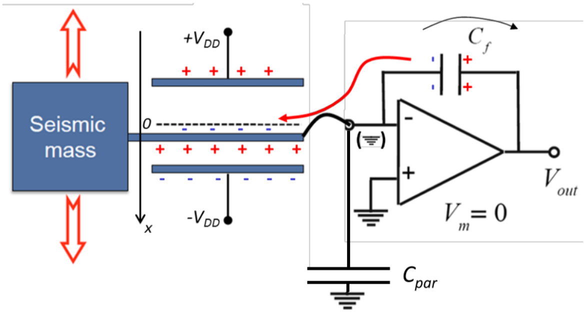


Figure 19: Schematic of Parasitic capacitances

However we will see that large parasitic capacitances have a bad impact on noise performance.

3.2.3 Effects of electrostatic forces

Once $\Delta V_{out}(x)$ is found, we need to relate x to the external acceleration (through a motion equation) in order to to find the sensitivity. We saw in the last chapter (section 2.5) that the application of a voltage for the readout generates unavoidable electrostatic forces, which should be considered in the force balance of the motion equation.

$$F_{elec,1} = -\frac{V_{DD}^2}{2}\frac{\varepsilon_0 AN}{(g+x)^2} \qquad F_{elec,2} = +\frac{V_{DD}^2}{2}\frac{\varepsilon_0 AN}{(g-x)^2}$$

$$\implies F_{elec} = \frac{V_{DD}^2}{2}\frac{\varepsilon_0 AN}{(g-x)^2} - \frac{V_{DD}^2}{2}\frac{\varepsilon_0 AN}{(g+x)^2}$$

We can now put the net electrostatic force inside the dynamic equation of the suspended mass. For sake of simplicity, we start from the case of <u>no external acceleration</u> and we look only at the quasi-stationary behavior:

$$\cancel{m\ddot{x}} + \cancel{b\dot{x}} + kx = \cancel{ma_{ext}} + F_{elec}$$

$$kx = F_{elec} = \frac{V_{DD}^2}{2}\frac{\varepsilon_0 AN}{(g-x)^2} - \frac{V_{DD}^2}{2}\frac{\varepsilon_0 AN}{(g+x)^2}$$

Let's find now the stable and unstable equilibrium points. It is quite instructive to solve this equation in stationary conditions (equilibrium points) with a **graphical approach**.

$$\frac{V_{DD}^2}{2}\frac{C_0}{g}\left[\frac{1}{(1-x/g)^2} - \frac{1}{(1+x/g)^2}\right] - kx = 0$$

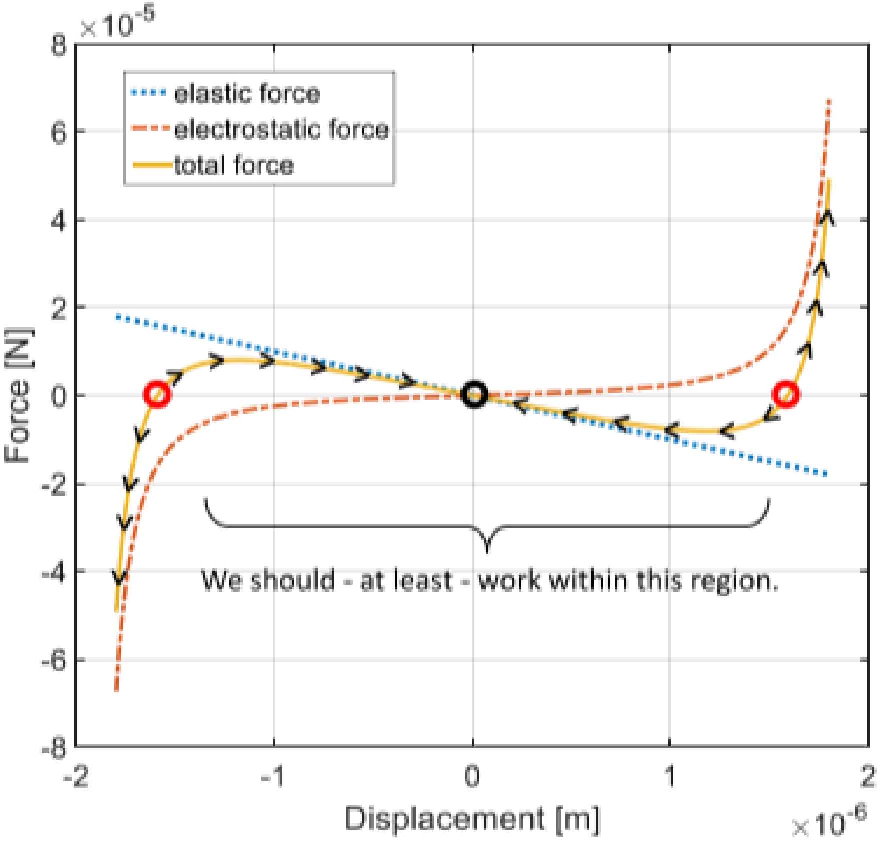


Figure 20: Red circles → UNSTABLE Black circle → STABLE

The equilibrium conditions can significantly change if one tries to either:

- **increase** the facing **area** (thus the total capacitance at rest);
- **decrease** the **gap** between plates;
- **decrease** the **stiffness**;
- **increase** the bias **voltage**.

Once a certain condition is reached, there is no longer any stable equilibrium point! The system undergoes an instability known as **pull-in**: the rotor plates snap onto the stator ones.
From the shown graphs one can infer that the slope of the two curves around the origin should be compared to identify the critical condition. Let us thus solve the equation for small displacements ($x \ll g$):

$$kx = \frac{V_{DD}^2}{2}\frac{\varepsilon_0 AN}{(g-x)^2} - \frac{V_{DD}^2}{2}\frac{\varepsilon_0 AN}{(g+x)^2}$$

$$kx = \frac{V_{DD}^2}{2}\frac{C_0}{g}\left[\frac{1}{(1-x/g)^2} - \frac{1}{(1+x/g)^2}\right]$$

$$kx = \frac{V_{DD}^2}{2}\frac{C_0}{g}\left[\frac{1}{1+\cancel{(x/g)^2}-2x/g} - \frac{1}{1+\cancel{(x/g)^2}+2x/g}\right]$$

$$kx = \frac{V_{DD}^2}{2}\frac{C_0}{g}\left[\frac{1}{1-2x/g} - \frac{1}{1+2x/g}\right] = \frac{V_{DD}^2}{2}\frac{C_0}{g}\left[\frac{4x/g}{1-\cancel{(2x/g)^2}}\right]$$

$$kx = \frac{V_{DD}^2}{2}\frac{C_0}{g}\left[\frac{1}{(1-x/g)^2} - \frac{1}{(1+x/g)^2}\right] \xrightarrow{\text{for small displacement}} kx = 2\,V_{DD}^2\frac{C_0}{g^2}x$$

For small displacements, the electrostatic force is linear with the displacement, with an opposite sign w.r.t. the elastic force.

$$k - 2\,V_{DD}^2\frac{C_0}{g^2} = 0 \xrightarrow{\text{definition of electrostatic stiffness}} k + k_{elec} = 0$$

The term $-2\,V_{DD}^2\frac{C_0}{g^2}$ takes thus the name of (negative) **electrostatic equivalent stiffness**, k_{elec}. Let's calculate now the pull-in voltage:

$$k - 2\,V_{DD}^2\frac{C_0}{g^2} = 0 \implies V_{DD}^2 = \frac{k\,g^2}{2\,C_0} \implies V_{DD,PI} = \sqrt{\frac{g^3\,k}{2\,\varepsilon_0\,A\,N}}$$

There is no pull-in if and only if the elastic stiffness is, in modulus, larger than the electrostatic equivalent stiffness! Again from a graphical point of view, we note that an equilibrium point exists only if the elastic force slope around zero is larger than electrostatic force slope, which is the electrostatic stiffness.

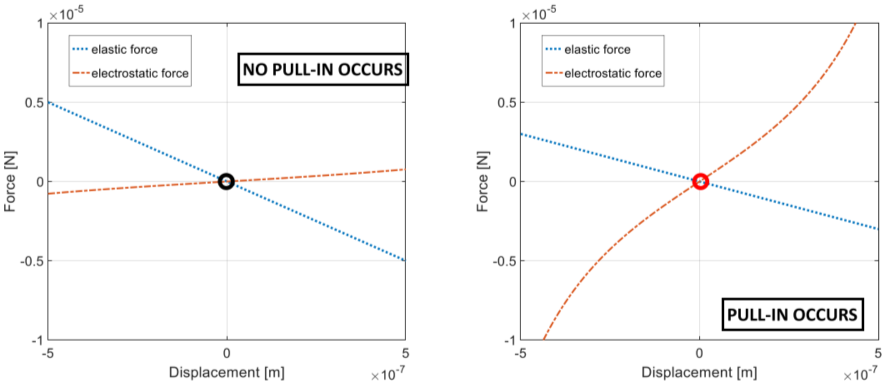


Figure 21: **DX**: NO PULL-IN occours **SX**: PULL-IN occours

Our ultimate goal is to calculate an acceleration, so we remove now the hypothesis of zero acceleration, and assume $K_{el} \ll k$.
For **small accelerations** we have a shift of the equilibrium point (this is the displacement we want to measure to recover the acceleration value).
After a certain acceleration value however, no more stable point exists! Too large accelerations can cause instability in PP MEMS accelerometers, even at biasing values lower than the pull-in voltage!

$$\frac{V_{DD}^2}{2}\frac{C_0}{g}\left[\frac{1}{(1-x/g)^2}-\frac{1}{(1+x/g)^2}\right]+m\,a_{ext}-kx=0$$

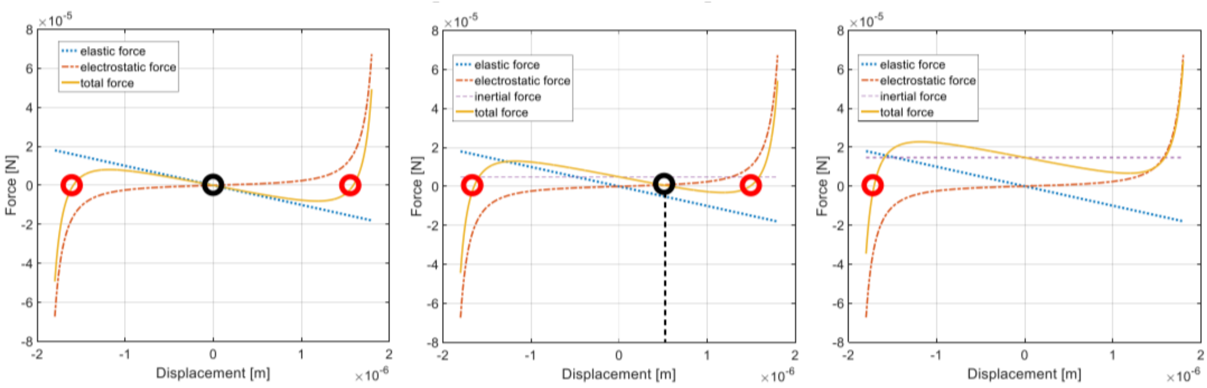


Figure 22: **DX**: $a_{ext} = 0$ **CENTER**: small a_{ext} **SX**: Unstable condition

Mechanical stoppers are rigid elements used to avoid the complete snap between rotor and stators plates, which can cause adhesion. Stoppers are positioned at a distance from the rotor which is smaller than the gap between plates. These are biased at the same voltage as the suspended mass so to avoid electrostatic forces and are usually of relatively small contact area, to minimize the probability of adhesion.

3.2.4 Parallel-Plate Accelerometer Sensitivity

We now solve the stationary condition, for small displacements, to find out the output voltage vs the input acceleration.

$$kx = ma_{ext} + 2V_{DD}^2 \frac{C_0}{g^2} x$$

$$x = \frac{m}{\left(k - 2V_{DD}^2 C_0/g^2\right)} a_{ext} = \frac{m}{k + k_{elec}} a_{ext} = \frac{m}{k_{tot}} a_{ext} = \frac{1}{\omega_0^2} a_{ext}$$

The MEMS resonance frequency changes as the elastic stiffness changes!
Indeed it is $\omega_0 = \sqrt{\frac{k_{tot}}{m}}$! This effect is named electrostatic softening or tuning of the resonance frequency.
A MEMS accelerometer is well-designed only if it effectively undergoes small displacements even for accelerations corresponding to the full-scale value!
By putting together the two found expressions below:

$$\Delta V_{out} = 2\frac{V_{DD}}{C_f}\frac{C_0}{g} x \qquad\qquad x = \frac{1}{\omega_0^2} a_{ext}$$

We can finally evaluate the overall sensitivity of a differential, parallel-plate MEMS axel readout through a charge amplifier:

$$\frac{\Delta V_{out}}{a_{ext}} = 2\frac{V_{DD}}{C_f}\frac{C_0}{g}\frac{1}{\omega_0^2} = 2\frac{V_{DD}}{C_f}\frac{C_0}{g}\frac{m}{\left(k - 2V_{DD}^2 \frac{C_0}{g^2}\right)}$$

As we can see:

- A **small gap** enhances the sensitivity but it is unfavorable for pull-in issues
  - remember that: $V_{DD,PI} = \sqrt{\frac{g^3\, k}{2\, \varepsilon_0\, A\, N}}$;
- A **large mass** enhances the sensitivity but take care of bandwidth and area limits
  - remember that: $\omega_0 = \sqrt{\frac{k_{tot}}{m}}$;
- A **small overall stiffness** enhances the sensitivity but it is unfavorable for pull-in issues and max bandwidth;
- A **large bias voltage** enhances the sensitivity but it facilitates pull-in and is limited by the consumption of the IC.
- A **large thickness** enhances the sensitivity!
  - Indeed A, m and k grow with h (pull-in and resonance are independent of h!)

3.2.5 Linearity of a PP configuration

The sources of non-linearity in the axel response are two:

- the nonlinear response of differential PPs;
- nonlinear effects from electrostatic forces.

We consider only the first one, assuming a device which is quite safe from pull-in instabilities (low k_{elec}). We can define the % linearity error of the sensitivity, relative to the full-scale:[8]

$$\varepsilon_{\%}(x) = \frac{2\,C_0\frac{x}{g}\left[\frac{1}{1-(x/g)^2}\right] - 2\,C_0\frac{x}{g}}{\Delta C_{FSR}} \cdot 100 = \frac{2\,C_0\frac{x}{g}\left[\frac{1}{1-(x/g)^2}\right] - 2\,C_0\frac{x}{g}}{2\,C_0\frac{x_{max}}{g}\left[\frac{1}{1-(x_{max}/g)^2}\right]} \cdot 100 =$$

$$= \frac{x\left[\frac{1}{1-(x/g)^2}\right] - x}{x_{max}\left[\frac{1}{1-(x_{max}/g)^2}\right]} \cdot 100$$

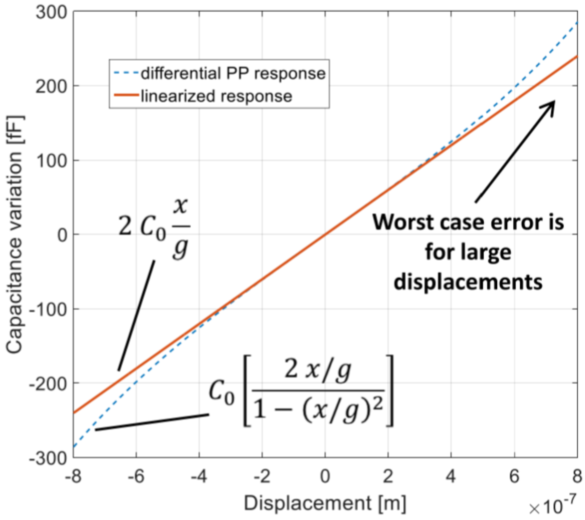


Figure 23: Linearization of PP response

Let's evaluate the **worst case** (i.e. at $x = x_{max}$):

$$\varepsilon_{\%_{max}} = \frac{x_{max}\left[\frac{1}{1-(x_{max}/g)^2}\right] - x_{max}}{x_{max}\left[\frac{1}{1-(x_{max}/g)^2}\right]} \cdot 100 = \left(1 - \frac{1}{\left[\frac{1}{1-(x_{max}/g)^2}\right]}\right) \cdot 100 =$$

$$= \left(\frac{x_{max}}{g}\right)^2 \cdot 100$$

---

[8] Less formally $\frac{\text{True}\Delta C - \text{Linearized}\Delta C}{\Delta C|_{a=a_{FSR}}}$

3.2.6 Design trade-offs

We have seen that there are marked trade-offs between:

1. **Sensitivity and pull-in** Pull-in effects require the device stiffness to be relatively large. As a consequence, the resonance frequency grows and the sensitivity (which goes with $1/f_0^2$) decreases.

$$V_{DD_{PI}} \propto \sqrt{k} \qquad\qquad \frac{\Delta V_{out}}{a_{ext}} \propto \frac{1}{k_{tot}}$$

2. **Sensitivity and full-scale range** Non-linearity is limited by the "small displacement" requirements of differential parallel plates.

$$a_{max} = \frac{x_{max}}{dx/da} = x_{max} \cdot \omega_0^2$$

Let's see if there are any other configuration that can avoid the pull-in and non-linearity.

3.3 Comb-finger (CF) configuration

Comb-finger capacitors are **area-varying** type:

$$C = \frac{2\varepsilon_0 AN}{g} = \frac{2\varepsilon_0 h(x+x_0)N}{g} \qquad\qquad \frac{dC}{dx} = \frac{2\varepsilon_0 hN}{g}$$

The capacitance variation is linear with x!

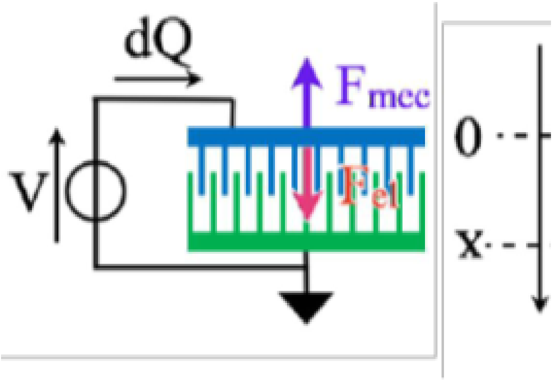


Figure 24: Comb-figure configuration

The electrostatic force is calculated exactly as for the parallel-plate case but the result has the important difference that the force is independent of the displacement x. So, in a differential configuration, the force is null even in presence of displacements!

$$F_{elec} = \frac{V^2}{2}\frac{dC}{dx} = \frac{V^2 \varepsilon_0 hN}{g}$$

$$|F_{elec,1}| = |F_{elec,2}| = \frac{V^2 \varepsilon_0 hN}{g} \qquad \Longrightarrow \qquad m\ddot{x} + b\dot{x} + kx = m\, a_{ext} = F_{ext} + \cancel{F_{elec}}$$

A solution to linearity and FSR issues could be the implementation of CF-based readout. For the computation of the sensitivity we could follow these steps:

1. evaluate the displacement per unit acceleration;
2. evaluate the single-ended capacitance variation per unit displacement;
3. evaluate the output voltage per unit differential capacitance variation.

$$\frac{\Delta V_{out}}{a_{ext}} = 2\frac{dV}{dC}\frac{dC}{dx}\frac{dx}{da} = 2\frac{V_{DD}}{C_f}\frac{C_0}{x_0}\frac{1}{\omega_0^2} \rightarrow \text{linear with } \textbf{NO approximation}$$

Advantages:

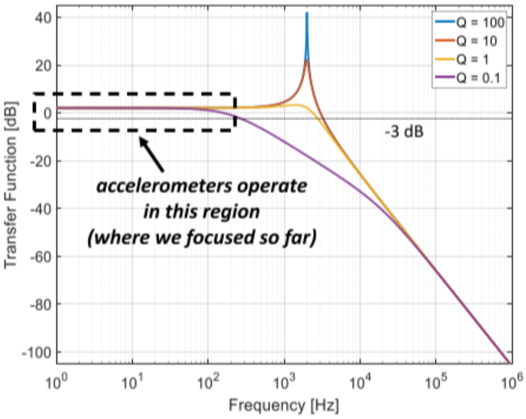
- The electrostatic force **is not a function of the displacement**. There is thus no electrostatic stiffness. In turn, there is no electro-static softening. No risk of pull-in exists.
- Ideally, there is no non-linearity in the capacitive readout. No trade-off exists between sensitivity and full-scale.

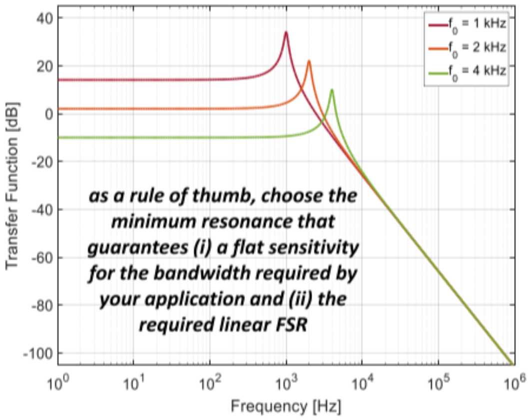
Drawbacks:

- For the same resonance, gap, and voltage, the number of fingers that one can fit in a given area does not allow to reach the same dC/dx as for PP devices. About a factor 5-10, as a rule of thumb, due to:
  - $\mathbf{x_0} \gg g$ (to avoid fringe effects);
  - for the same available area, $C_{0,CF} < C_{0,PP}$.

3.4 Accelerometer bandwidth and operating region

In 2.3, we studied the transfer function between external force and displacement. It is valid as well between **external acceleration** (linear with applied force) and **output voltage** (linear with x in *CF*, and in *PP* with small-displacement approximation).





The Fourier domain transfer function modulus can be indeed written as below, by combining the formulas of previous paragraphs:

$$|T_{VA}(j\omega)| = \left|\frac{X(j\omega)}{F_{ext}(j\omega)}\right|\left|\frac{V(j\omega)}{X(j\omega)}\right|\left|\frac{F_{ext}(j\omega)}{A(j\omega)}\right| = m\frac{2V_{DD}}{C_f}\frac{C_0}{g}\frac{1/m}{\sqrt{(\omega_0^2-\omega^2)^2+\left(\frac{\omega\omega_0}{Q}\right)^2}} =$$

$$= \frac{2V_{DD}}{C_f}\frac{C_0}{g}\frac{1}{\sqrt{(\omega_0^2-\omega^2)^2+\left(\frac{\omega\omega_0}{Q}\right)^2}} \xrightarrow{\text{low-frequency: } \omega\ll\omega_0} |T_{VA}| = \frac{\Delta V_{out}}{a_{ext}} = 2\frac{V_{DD}}{C_f}\frac{C_0}{g}\frac{1}{\omega_0^2}$$

For frequencies close or beyond resonance, the low-frequency expression is no longer valid.
First considerations for "choosing" the Q factor:

- the Q factor has no effect on the DC response;
- a too low Q (< 0.5) lowers the -3 dB bandwidth.

Which value, possibly ≥ 0.5, should we take?
To answer this question, we need to make another consideration. The accelerometer can be subject to random shocks and vibrations, free falls (ecc...). In all such events, the resonance frequency is excited by a broadband spectral stimulus. We have seen (section 2.3) that the ring-down response time goes linearly with Q in these cases: $\tau = \frac{Q}{\pi f_0}$. Therefore, all information would be corrupted in the time interval after the event.
So, how do we define the accelerometer bandwidth?

- For $Q \leq 0.5$ (real split poles) just take the -3dB bandwidth value of the first pole.
  $\mathbf{Q = 0.5}$ seems the best choice to maximize the bandwidth. (real coincident poles)
- For $Q \geq 0.5$ take the minimum between the -3dB bandwidth and $\frac{1}{2\pi\tau}$.
  Since $\frac{1}{2\pi\tau} = \frac{f_0}{2Q}$, **$\mathbf{Q = 0.5}$ is the best choice to maximize the bandwidth.**

The preliminary conclusion is that optimum **Q is $\mathbf{\frac{1}{2}}$**. Such conclusion (which is often correct!) does not take into account two facts:

- Q is dependent on b, but also noise, as we will see in a while, depends on b;
- oscillations at resonance can be always filtered using electronic LPF.

3.5 Thermo-mechanical Noise

We have seen how the response of a MEMS to steps or pulses shows a ring-down behavior, during which its kinetic energy decreases. By the **fluctuation-dissipation theorem**, any dissipative mechanism that results in mechanical damping must be balanced by a fluctuation force to maintain macroscopic energy balance, hence thermal equilibrium. In other words, to prevent the unphysical result that the system temperature goes below the temperature of the surroundings, we include in the motion equation a term which we name *"fluctuation force"* F_n. This is the force through which the energy passes between the MEMS and the environment.
In the simplest vision: gas molecules, randomly excited by thermal agitation, hit the mass and apply a random force (mechanical noise) to it.
The goal is to find an expression of the power spectral density S_{F_n} of F_n. We will work through the velocity, as this is a physical quantity easy to relate to thermal agitation.
Let us start from the motion equation in the Laplace domain:

$$mX(s)s^2 + bX(s)s + kX(s) = F_n(s)$$

Let us write it as an explicit function of the velocity:

$$V(s) = X(s)s \longrightarrow \; mV(s)s + bV(s) + \frac{k}{s}V(s) = F_n(s)$$

Let us define the transfer function between the velocity and the applied force:

$$V(s)\left(m\,s + b + \frac{k}{s}\right) = F_n(s) \implies V(s) = \frac{1}{\left(m\,s + b + \frac{k}{s}\right)}F_n(s) = T_{VF}(s)F_n(s)$$

We now pass to the Fourier domain by putting $s = jw$, and we write the power spectral density of the velocity, $S_{vn}(\omega)$):

$$V(j\omega) = \frac{1}{\left(m\,j\omega + b + \frac{k}{j\omega}\right)}F_n(j\omega) = \frac{1}{b + j\left(m\omega - \frac{k}{\omega}\right)}F_n(j\omega)$$

$$S_{vn}(\omega) = |T_{VF}|^2 S_{Fn}(\omega) = \frac{1}{b^2 + \left(m\omega - \frac{k}{\omega}\right)^2}S_{Fn}(\omega) = \frac{1}{b^2}\frac{1}{1 + \left(\frac{m\omega}{b} - \frac{k}{b\omega}\right)^2}S_{Fn}(\omega)$$

We re-arrange the found expression by using the formulas for the resonance frequency $\left(\omega_0 = \sqrt{\frac{k}{m}}\right)$ and the quality factor $\left(Q = \omega_0\frac{m}{b} = \frac{k}{\omega_0 b}\right)$:

$$S_{vn}(\omega) = \frac{1}{b^2}\frac{1}{1 + \left(\frac{\omega}{\omega_0}Q - \frac{\omega_0}{\omega}Q\right)^2}S_{Fn}(\omega) = \frac{1}{b^2}\frac{1}{1 + Q^2\left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega}\right)^2}S_{Fn}(\omega)$$

To evaluate the *rms velocity noise*, we integrate the velocity spectrum over the entire frequency range ($df = d\omega/2\pi$):

$$\bar{v}_n^2 = \frac{1}{2\pi}\int_0^\infty S_{vn}(\omega)d\omega = \frac{1}{2\pi b^2}\int_0^\infty \frac{1}{1 + Q^2\left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega}\right)^2}S_{Fn}(\omega)d\omega =$$

$$= \frac{\omega_0}{2\pi b^2}\int_0^\infty \frac{1}{1 + Q^2\left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega}\right)^2}S_{Fn}(\omega)d\frac{\omega}{\omega_0}$$

Let us assume that the force noise spectrum is white $S_{Fn}(\omega) = S_{Fn}$:

$$\frac{\omega_0 S_{Fn}}{2\pi b^2}\int_0^\infty \frac{1}{1 + Q^2\left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega}\right)^2}d\frac{\omega}{\omega_0}$$

The finite integral below has a value that equals $\approx \pi/2$, regardless of the specific Q. With this we can conclude the calculation of the *rms velocity noise*:

$$\bar{v}_n^2 = \frac{\omega_0 S_{Fn}}{2\pi b^2}\frac{1}{Q}\frac{\pi}{2} = \frac{\omega_0 S_{Fn}}{4b^2 Q} = \frac{S_{Fn}}{4\,m\,b}$$

For a 1-degree-of-freedom system, we can write the corresponding rms noise in terms of kinetic energy:

$$\frac{1}{2}m\bar{v}_n^2 = \frac{1}{2}m\frac{S_{Fn}}{4\,m\,b}$$

We also know that the energy of a *1-DOF system* is related to the absolute temperature through the Boltzmann constant k_B:

$$\frac{1}{2}m\bar{v}_n^2 = \frac{1}{2}k_BT$$

So we can finally find our expression for the fluctuation force spectrum[9], thus in units of [N^2/Hz]:

$$\frac{\cancel{1}}{\cancel{2}}k_BT = \frac{\cancel{1}}{\cancel{2}}\cancel{m}\frac{S_{Fn}}{4\,\cancel{m}\,b} \longrightarrow S_{Fn} = 4\,k_B\,T\,b$$

The expression is independent of the frequency as we assumed white noise. The noise in terms of force density $[N^2/Hz]$ can be converted into acceleration $[(m/s^2)^2/Hz]$ noise density:

$$S_{An} = \frac{S_{Fn}}{m^2}$$

The **N**oise **E**quivalent **A**cceleration **D**ensity (**NEAD**), considering only thermo-mechanical noise, is therefore:

$$NEAD = \sqrt{S_{An}} = \sqrt{\frac{4k_BTb}{m^2}} = \sqrt{\frac{4k_BT\omega_0}{mQ}} \qquad \left[\frac{m/s^2}{\sqrt{Hz}}\right]$$

Noise improves:

- with low b (or high Q);
- with large mass values.

A note on white noise. We assumed $S_{F_n}(j\omega)$ as white noise. This is a reasonable approximation if we think at the thermal agitation in a molecular perspective. Indeed, we can model the corresponding force noise in **time domain** as some **impulses** due to gas molecules colliding on the mass. As a result, in the **frequency domain** we find the sum of constant values: i.e. **white noise**

---

[9] **This expression is valid for all Spring-Mass-damper systems!**

3.6 Electronic Noise

MEMS thermo-mechanical noise is not the only noise source in micro-electro-mechanical systems. Electronic noise should be also taken into account, typically deriving from resistance Johnson noise, operational amplifier white and 1/f noise. Let us make a quick example with the circuit of paragraph 3.2.1.

for $C_{par} \gg C_0$

$$S_{Vn,out} = S_{Vn}\left(1+\frac{s\,R_f C_{par}}{1+s\,R_f C_f}\right)^2 = $$
$$\approx S_{Vn}\left(1+\frac{C_{par}}{C_f}\right)^2$$

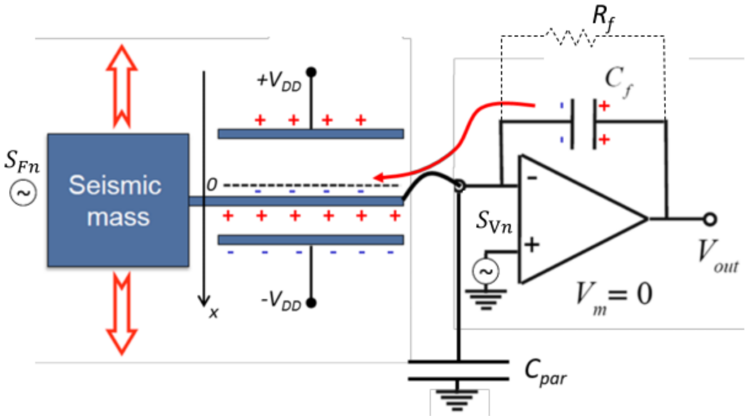


Figure 25: Electronic Noise Spectrum

R_f is the parasitic (non-infinite) resistance of the switch (e.g. an off MOSFET) across C_f.

$$\sqrt{S_{Vn,out}} \sim \sqrt{S_{Vn}\left(\frac{C_{par}}{C_f}\right)^2}$$

$$\sqrt{S_{a,eln}} = \frac{\sqrt{S_{Vn,out}}}{2\frac{V_{DD}}{C_f}\frac{C_0}{g}\frac{1}{\omega_0^2}} = \frac{\sqrt{S_{Vn}}\,C_{par}}{2V_{DD}\frac{C_0}{g}\frac{1}{\omega_0^2}} = $$
$$= 2.1\cdot 10^{-4}\frac{\frac{m}{s^2}}{\sqrt{Hz}} = 22\frac{\mu\hat{g}}{\sqrt{Hz}}$$

$$\sqrt{\overline{a_{eln}^2}} = \sqrt{S_{a,eln}}\cdot\sqrt{BW} = 0.34\,m\hat{g}$$

$\sqrt{S_{Vn}} = 20\frac{nV}{\sqrt{Hz}}$
$C_f = 100\,fF$
$V_{DD} = 1.5\,V$
$g = 2\,\mu m$
$f_0 = 2\,kHz$
$C_p = 10\,pF$

$$S_{Vn,out} = S_{Vn}\left(1+\frac{s\,R_{sw}C_{par}}{1+s\,R_{sw}C_f}\right)^2 \sim S_{Vn}\left(1+\frac{C_{par}}{C_f}\right)^2 \sim S_{Vn}\left(\frac{C_{par}}{C_f}\right)^2$$

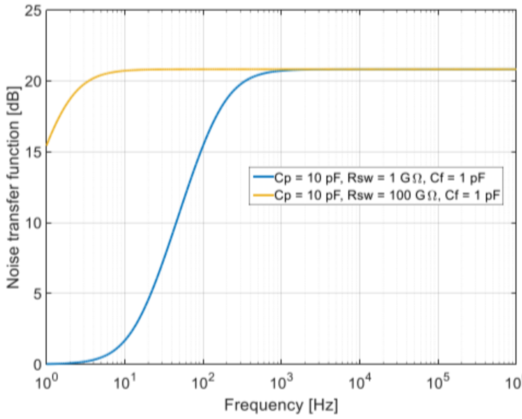
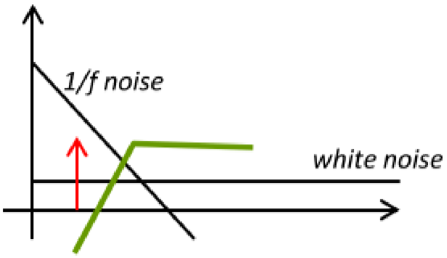


Figure 26: Example with real and typical values

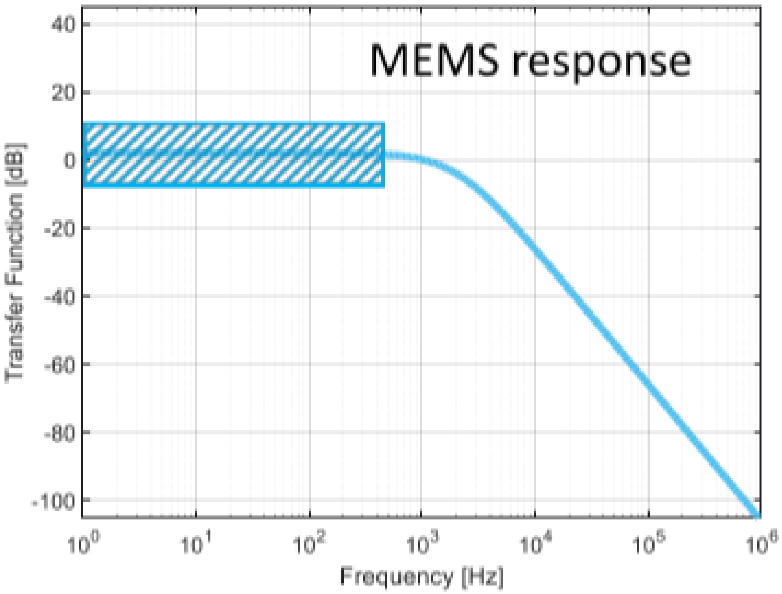
Noise density is amplified by roughly a factor ten (assuming a parasitic of 10 pF and feedback capacitance of 1 pF).

NOTES:
At low frequencies ($< 1 - 10kHz$), noise may be strongly limited by $1/f$ noise issues. For this reason the MEMS is often switched (with a certain frequency) between full bias (±VDD) and no bias, so to "modulate" the signal to higher frequencies. This allows to high-pass filter the 1/f noise contributions. Demodulation is then applied to bring back the signal to the baseband.

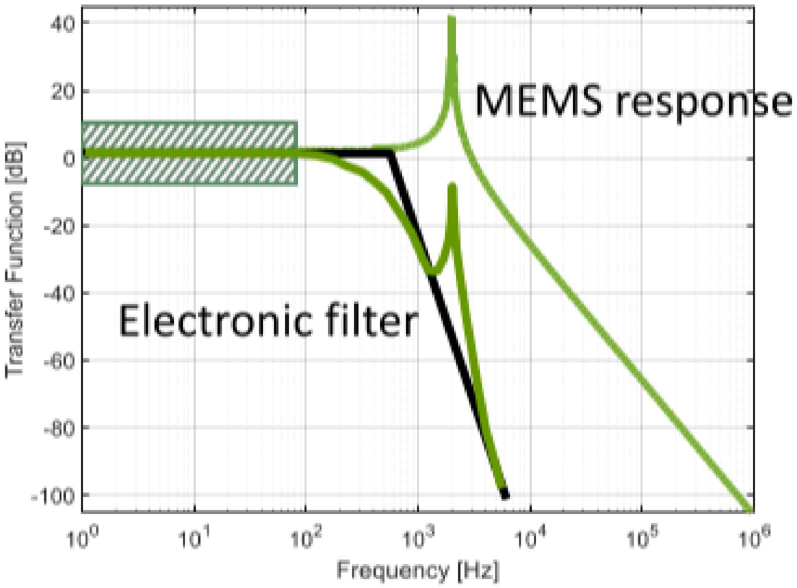


3.7 Optimum Q for application

- **Optimization of the bandwidth**
  It requires ideally a Q in the order of 0.5 (in between over-damped and under-damped conditions).
  This should be effectively pursued, as far as the electronic noise still dominates. Having a relatively high damping (low Q) does not represent an issue in this case.



- **Noise optimization**
  It requires high Q. This should be pursued as far as device noise dominates. This typically occurs for high-performance applications, where power dissipation is raised and electronic noise is low. Electronic low-pass filtering is required to filter long ring-downs.

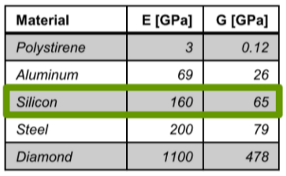


3.8 Flexural Springs

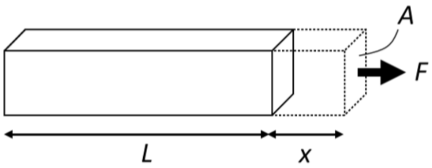
As we know a process is characterized by a certain thickness, defined usually by the epitaxial polySi growth. In this sense, MEMS processes are “planar”, as you can design the in-plane dimensions but you cannot choose the out-of-plane one. The elastic stiffness of a spring depends both on its material and on its geometry. The material parameter is the so called **Young’s modulus *E*** (or modulus of elasticity), defined as the ratio between the stress σ applied orthogonal to a material surface A (force per unit area) and the reversible strain ε shown by the material (relative elongation):

$$\sigma = E\,\varepsilon \qquad\qquad \frac{F}{A} = E\,\frac{x}{L}$$

It defines the material resistance to elastic deformations. Silicon and poly-silicon have Young’s modules in the order of 150-180 GPa, comparable to steel and twice the value of aluminum.



Material	E [GPa]	G [GPa]
Polystirene	*3*	*0.12*
Aluminum	*69*	*26*
Silicon	*160*	*65*
Steel	*200*	*79*
Diamond	*1100*	*478*



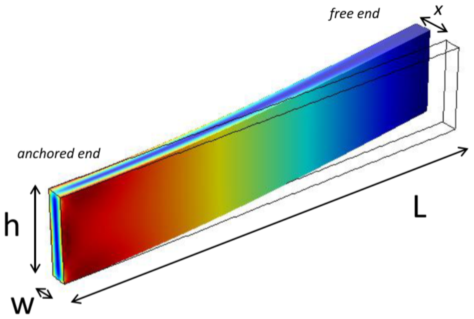
An elementary stiffness formula calculation can be done for the case above of **axial stiffness**. The stiffness is by definition the ratio of the displacement to the force, so we easily find its dependence on the Young's modulus and the geometry:

$$F = \frac{A\,E}{L}x \qquad \longrightarrow \qquad k = \frac{F}{x} = \frac{A\,E}{L} = \frac{E\,h\,w}{L}$$

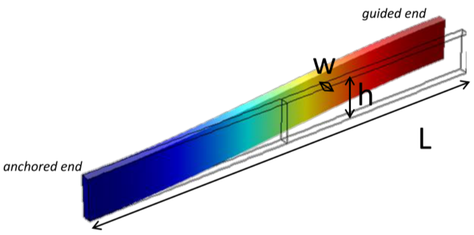
Heavy calculations lead to the generation of the so-called **stiffness matrix**, indicating the stiffness of a spring in one direction (in this case a beam of rectangular cross section) as a function of the applied force direction.
We will use only the (two) most useful formulas for our calculations, from which we will derive all configurations of interest for our goal.

1. The first configuration is named **In-plane clamped beam**, and analyzes a beam clamped at one end only to the substrate, with the second end free to move in any directions. As one intuitively can expect, the stiffness decreases with the beam length and increases with beam width and height. The exact dependence is described by the formula below:

$$k_X = \frac{F_X}{x} = \frac{E\,w^3\,h}{4L^3}$$



2. **Guided-end springs** have one end clamped to the substrate and the other end free to move along a plane (usually parallel to the clamped end plane). It is the most common case of in-plane (IP) springs in MEMS devices.
The difference w.r.t. the first configuration is that the free-end boundary moves parallel to the clamped end boundary.



To find out the stiffness of a guided-end spring, it is first convenient to have a look at how to find the stiffness of the series of **multiple springs**.

Let us assume two springs with arbitrary stiffness k_1 and k_2. The springs are connected in series and subject to a force F. So:

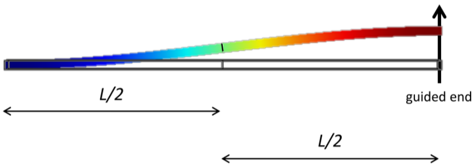
$$x_1 = \frac{F}{k_1} \qquad x_2 = \frac{F}{k_2}$$

$$x = x_1 + x_2 = \frac{F}{k_1} + \frac{F}{k_2} = F\left(\frac{1}{k_1} + \frac{1}{k_2}\right) = \frac{F}{k_x} \longrightarrow \frac{1}{k_x} = \frac{1}{k_1} + \frac{1}{k_2} \longrightarrow k_x = \frac{k_1 k_2}{k_1 + k_2}$$

The stiffness calculation can be easily done by assuming that the spring is formed by the series of two identical free-end springs, having half the length L each.

$$k_1 = k_2 = \frac{Ew^3h}{4(L/2)^3} = 2\frac{Ew^3h}{L^3}$$

$$k_X = \frac{k_1 k_2}{k_1 + k_2} = \frac{1}{2}k_1 = \frac{Ew^3h}{L^3}$$



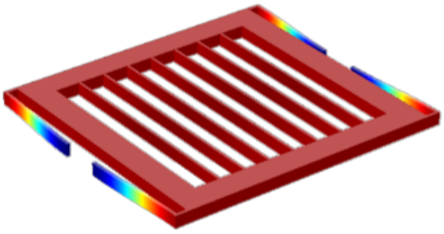
3.9 Parallel combination of spring

Devices have usually more than one spring, in a parallel configuration. The calculation for parallel springs is dual w.r.t. the calculation of series springs:

$$F_1 = k_1 x \qquad F_2 = k_2 x$$

$$F = F_1 + F_2 = (k_1 + k_2)x = k_x x \longrightarrow k_x = k_1 + k_2$$

We can now, as an example, calculate the overall stiffness of an accelerometer suspended by four springs of length L, in the shown configuration:



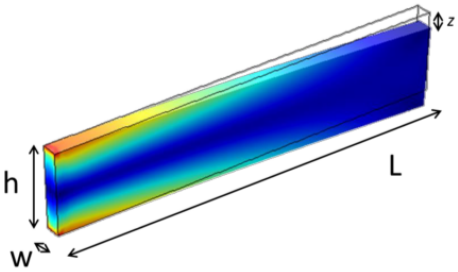
It is the parallel of 4 springs each of which is of the guided-end type. We thus put in parallel 4 identical springs, each having the stiffness calculated two slides ago:

$$K_{ix} = \frac{Ew^3h}{L^3} \qquad k_x = k_{1x} + k_{1x} + k_{1x} + k_{1x} = 4\,k_{ix} \qquad k_x = 4\frac{Ew^3h}{L^3}$$

3.10 Out-of-plane motion

Identical considerations hold for the calculation of the stiffness for a force acting along the vertical direction z, (in practice, you should exchange w and h).

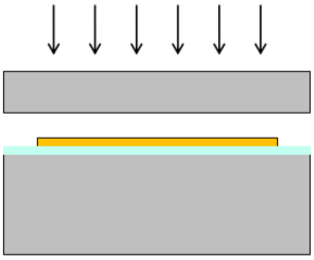
$$k_z = \frac{F_z}{z} = \frac{Eh^3w}{4L^3} \xrightarrow{\text{for 4 springs in parallel}} k_z = \frac{Eh^3w}{L^3}$$

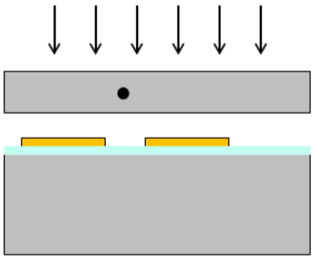


Note the cubic dependence on h, a parameter you cannot act on by design. This makes it difficult to obtain low stiffness values for z-axis devices based on this kind of spring (vertical translation).

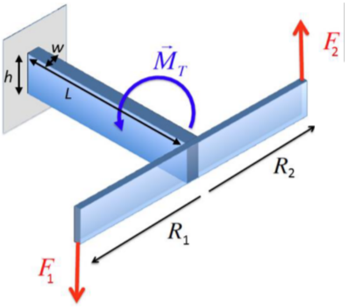
3.11 Torsional Springs

As we have seen the stiffness for z-axis translational motion is too large, we need to lower the stiffness by about two orders of magnitude! However, it's good to reject undesired out-of-plane motion in in-plane sensing devices. To solve the issue related to vertical motion, a common solution is to use torsional springs to allow out-of-plane (OOP) **rotations**. This gives a further advantage: while a translation in the vertical direction cannot have a differential readout, OOP rotations allow for a differential readout through a pair of electrodes placed beneath the suspended mass.


single-ended capacitance variation


differential capacitance variation

In general, an OOP rotation will be caused by a **torque**. In the course, we analyze only configurations similar to the one below.



The torque $\mathrm{M_T}$ will be the sum of the two individual torques, each defined as the product of the distance from the rotation center and the applied force.

$$M_T = R_1 F_1 + R_2 F_2$$

For the considered bar configuration, the torsional stiffness (in light blue in the main table) is given by:

$$k_\theta = G\,\frac{h\,w^3}{3L} \qquad [N \cdot m] \qquad (\text{G: } \textbf{Shear Modulus})$$

For the calculation of the **moment of inertia**, the torsional bar itself can be often neglected, compared to the contributions from other parts of the MEMS. Everything can be in general schematized like this:

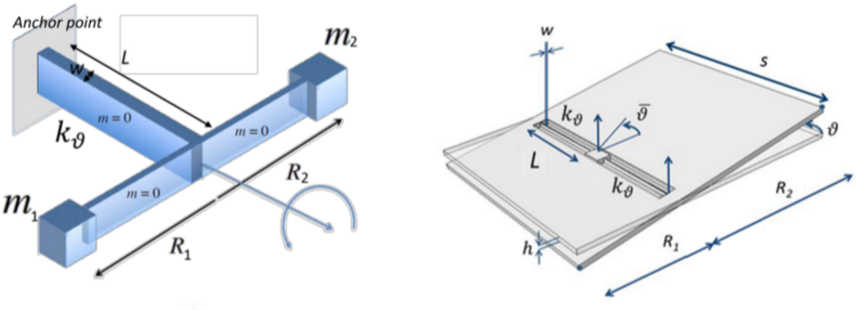


Figure 27: Vertical sensing accelerometer

This yields a moment of inertia which is, by definition:

$$I_1 = \int_{m_1} r_1^2 dm_1 = \int_0^{R_1} r_1^2 s\,(r_1)\,h\rho dr_1 = \frac{R_1^3 s\rho h}{3} = \frac{R_1^2 m_1}{3} \qquad I_2 = \frac{R_2^2 m_2}{3}$$

$$I_{TOT} = I_1 + I_2$$

3.12 Effects of process non-uniformities on springs

We need to solve the problem of **too narrow** width of in-plane springs for translation. The example below gives an idea of how important this issue can be:

- assume the considered spring width $w = 1.5\mu m$;
- assume a max variability in the width definition (by DRIE) $dw = \pm 0.15\mu m$;

$$k = E\frac{hw^3}{L^3} \qquad dk = E\frac{h}{L^3}3w^2dw \qquad \frac{dk}{k} = 3\frac{dw}{w}$$

$$f_0 = \frac{1}{2\pi}\sqrt{\frac{k}{m}} \qquad df_0 = \frac{1}{2\pi\sqrt{m}}\frac{-1}{2\sqrt{k}}dk \qquad \left|\frac{df_0}{f_0}\right| = \frac{1}{2}\frac{dk}{k} = \frac{3}{2}\frac{dw}{w}$$

$$S = \frac{\Delta V_{\text{out}}}{a_{\text{ext}}} = \alpha\frac{1}{f_0^2} \qquad dS = -2\alpha\frac{1}{f_0^3}df_0 \qquad \left|\frac{dS}{S}\right| = 2\frac{df_0}{f_0} = 3\frac{dw}{w}$$

As we can see a ± 10% uncertainty on the **width** turns directly into a ± 30% uncertainty on the **stiffness** and on the **sensitivity** (and on a ± 15% **bandwidth** uncertainty).
If we try increasing w we reduce the percentage uncertainty, but the stiffness increases by a factor 8 and the sensitivity decreases by the same factor! It is **not a convenient approach**!
A way to obtain low stiffness without increasing L (so the device area) and narrowing too much w is to arrange in series more springs of relatively large stiffness. It is the concept of **spring folding**: instead of a single-fold, narrow spring we put in series more folds with a larger w.

$$k_{1-\text{fold}} = E\frac{hw_1^3}{L^3} \qquad k_{N-\text{fold}} = \frac{1}{N}E\frac{hw_N^3}{L^3} \qquad \xrightarrow{k_{1-fold}=k_{N-fold}} \qquad w_N = \sqrt[3]{N}w_1$$

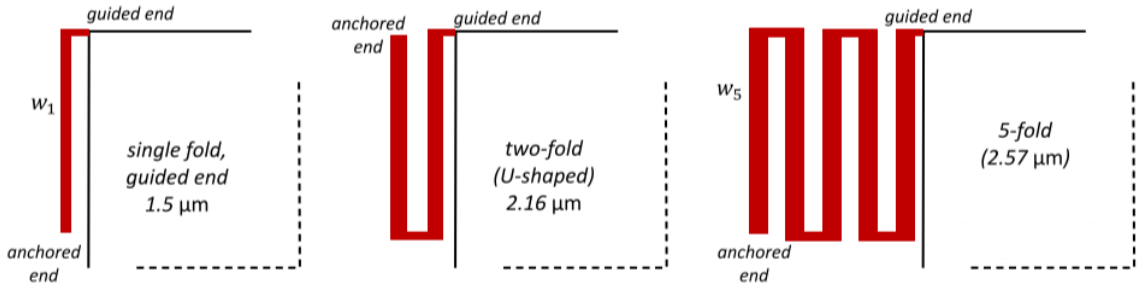


Figure 28: Different folded springs (e.g. $w_5 = \sqrt[3]{5}w_1 = 1.75w_1$)

3.13 Mechanical offset

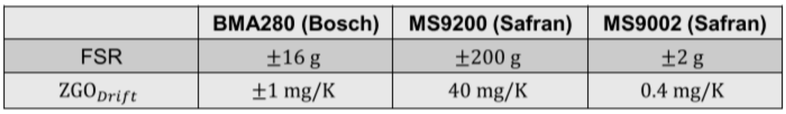
The offset (or **ZGO, zero-g-output**) is defined as the output value that you measure on an accelerometer axis when **no accelerations occur** in that direction. It has both electronic and mechanical origin.
We begin from the mechanical one: the mass can be offset w.r.t. the nominal centered rest position, due to process tolerances such as:

- uniformity of gaps between the two sets of differential parallel plates;
- residual mechanical stresses induced by wafer bending (temperatures gradients during process and/or operation, or stresses of wafer bonding).

The offset can vary with the temperature. This is a big issue because a simple initial calibration does not solve the offset drift problems. A mechanical offset of e.g. 10 nm only, causes an output variation which is about 600 times the 1-m$\hat{g}$ resolution and about 1/14 of the FSR!

$$S = 2\frac{dx}{da_{ext}}\frac{dC}{dx}\frac{dV}{dC} \qquad x_{max} = a_{FSR}\cdot\frac{dx}{da_{ext}} \qquad a_{ZGO}(T) = \frac{x_{os}(T)}{\frac{dx}{da_{ext}}}$$

State of the art capacitive MEMS accelerometers suffer from a clear **trade-off between ZGO drift and FSR**, as these terms depend in an opposite manner from the gain $dx/da_{ext} = 1/\omega_0^2$: accelerometers with lower resonance frequency show indeed a better ZGO drift but a lower FSR.



	BMA280 (Bosch)	MS9200 (Safran)	MS9002 (Safran)
FSR	±16 g	±200 g	±2 g
ZGO$_{Drift}$	±1 mg/K	40 mg/K	0.4 mg/K

A mechanical offset drift also causes the non-linearity to drift! Let us compare:

- The capacitance variation with no mechanical offset:

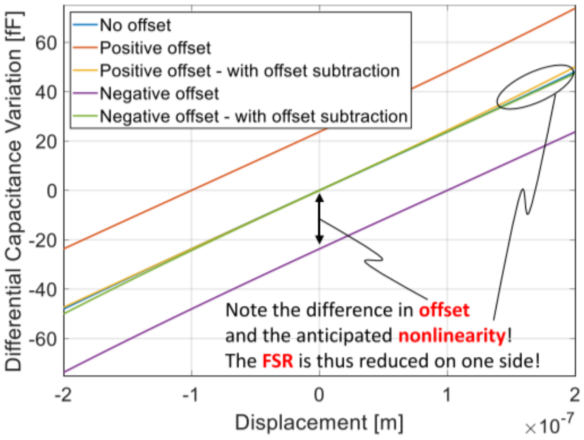
$$\Delta C_{diff} = 2C_0\frac{x}{g}\left[\frac{1}{1-\left(\frac{x}{g}\right)^2}\right]$$

- the capacitance variation in presence of a mechanical offset x_{os}:

$$\Delta C_{diff} = C_0\left(\frac{-1}{1+x/\left(g-x_{os}\right)} + \frac{1}{1-x/\left(g+x_{os}\right)}\right) \approx 2C_0\frac{x}{g}\left[\frac{1}{1-\left(\frac{x-x_{os}}{g}\right)^2}\right]$$

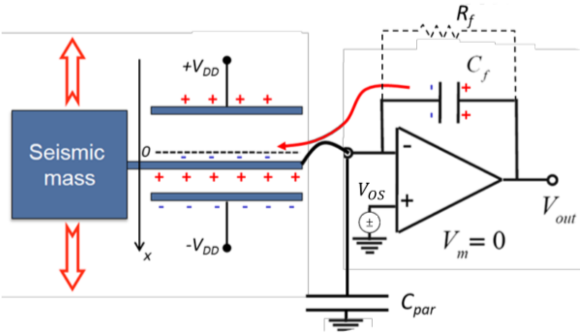
→ The linearization condition becomes $x - x_{os} \ll g$ which is dependent to x_{os}

The figure plots the capacitance variation versus displacement for a 100-nm mechanical (positive or negative) offset on a 1.5 µm gap.

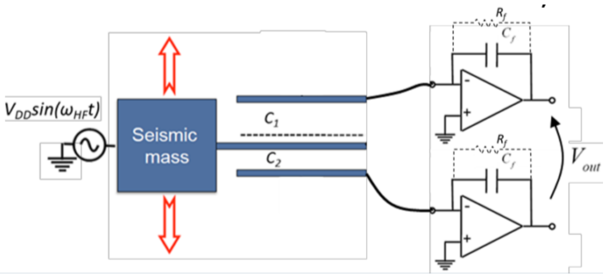


3.14 Electronic Offset

We have seen that the studied topology is not the best choice to sense DC accelerations (at low frequency the feedback is dominated by R_f).



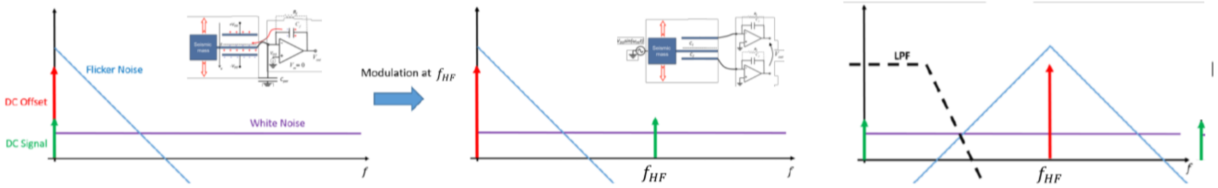
An additional issue of the discussed topology is that the offset of the operational amplifier will be seen at the output. Its drifts will therefore induce a net output offset drift that can be referred in terms of acceleration (same consideration holds for bias currents). A solution to all these issues is to use a **modulated signal**, as in the circuit below.



The calculation leads to a sensitivity similar to what seen so far:

$$\frac{\Delta V_{\mathrm{out}}}{a_{\mathrm{ext}}} = 2\frac{V_{DD}}{C_f}\frac{C_0}{g}\frac{1}{\omega_0^2}\sin\left(\omega_{HF}t\right) \qquad \Delta V_{out} = 2\frac{V_{DD}}{C_f}\frac{C_0}{g}\frac{1}{\omega_0^2}\sin\left(\omega_{HF}t\right)a_{\mathrm{ext}} + V_{OS}(T)$$

However, the signal at the amplifiers outputs is now modulated around a high frequency ($\omega_{HF} \gg \omega_0$, to excite motion with electrostatic forces). The amplifiers output spectrum becomes:



By multiplying it by a tone with same frequency and phase, one gets the signal back to baseband. Electronic offset, its drift, and 1/f noise are sent to high frequency and filtered through a low-pass stage.[10]

$$\begin{aligned}
\Delta V_{\mathrm{demod}} &= \left[2\frac{V_{DD}}{C_f}\frac{C_0}{g}\frac{1}{\omega_0^2}\sin\left(\omega_{HF}t\right)a_{\mathrm{ext}} + V_{OS}(T)\right]\sin\left(\omega_{HF}t\right) = \\
&= \left[2\frac{V_{DD}}{C_f}\frac{C_0}{g}\frac{1}{\omega_0^2}\sin^2\left(\omega_{HF}t\right)a_{\mathrm{ext}}\right] + V_{OS}(T)\sin\left(\omega_{HF}t\right) = {}^{11} \\
&= \left[\frac{V_{DD}}{C_f}\frac{C_0}{g}\frac{1}{\omega_0^2}a_{\mathrm{ext}} - \frac{V_{DD}}{C_f}\frac{C_0}{g}\frac{1}{\omega_0^2}a_{\mathrm{ext}}\cos\left(2\omega_{HF}t\right)\right] + V_{OS}(T)\sin\left(\omega_{HF}t\right) = {}^{12} \\
&= \frac{1}{2}\left[2\frac{V_{DD}}{C_f}\frac{C_0}{g}\frac{1}{\omega_0^2}a_{ext} + \cancel{2\frac{V_{DD}}{C_f}\frac{C_0}{g}\frac{1}{\omega_0^2}a_{ext}\sin\left(2\omega_{HF}t - \frac{\pi}{2}\right)}\right] + \cancel{V_{OS}(T)\sin\left(\omega_{HF}t\right)}
\end{aligned}$$

$$\Longrightarrow \frac{\Delta V_{demod}}{a_{ext}} = 2\frac{V_{DD}}{C_f}\frac{C_0}{g}\frac{1}{\omega_0^2}$$

Assume now that:

- you have a mechanical offset a_{os} that (ideally) is not drifting with T;
- the amplifier offset and its drifts are bypassed by the demodulation.

this looks a nice situation, in principle, to have low offset drift.
However if the feedback capacitances drift, the sensitivity drifts and this induces also a drift of the mechanical offset term!
Offset is critical whenever you need to integrate the signal to recover the position of a moving object in real time, starting from its initial position and the integration of the 3-axis accelerometer signals. This is the case of inertial-based navigation (integrating an offset leads to a diverging calculated position!). State-of-the-art accelerometers cannot cope with offset stability vs temperature specifications required by this kind of applications. However a lot of research is ongoing, and in a few years we expect big improvements in this field.

---

[10] Note: in the following computation the voltage drops by a factor 2, so we assumed a LPF with a $G = 2$ in the last equation

[11] $1 - 2\sin^2(x) = \cos(2x) \rightarrow 2\sin^2(x) = 1 - cos(2x)$

[12] $\cos(x) = \sin(x - \frac{\pi}{2})$

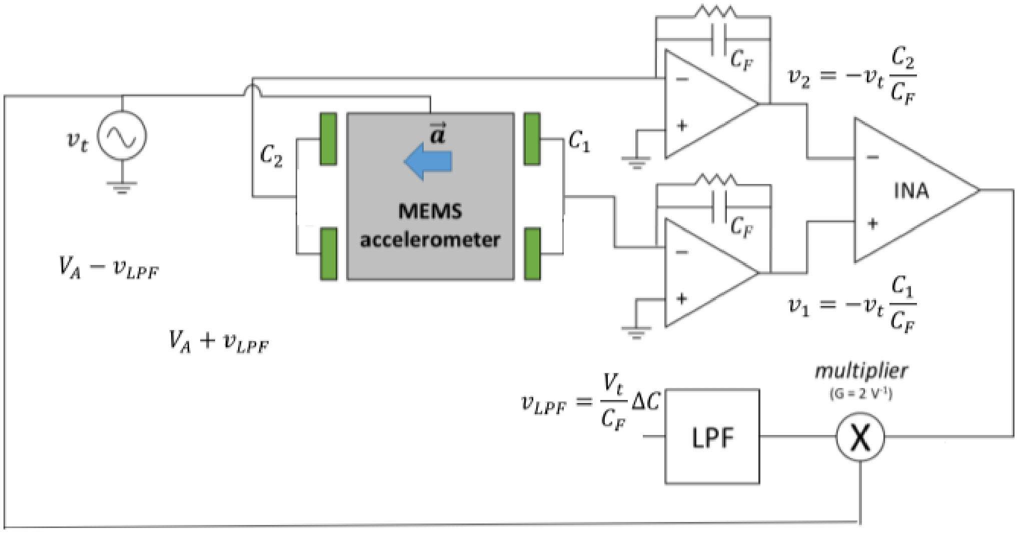
3.15 Alternative readout topologies for next axel generation

To lower costs, companies must fit more devices in every single wafer, therefore, their footprint should be small. In order to do that, there is the need to change the readout. In fact, voltage-controlled readout (i.e. applying a known voltage to a MEMS capacitor to read out the generated charge) limits the area scaling.
Miniaturization implies lower area, which implies lower mass and so a lower stiffness to keep the same resonance. If k is left as is, the resonance grows and sensitivity decreases. Similarly, we cannot decrease V_{DD}, because this causes a decrease in sensitivity, and we can't act on the gap to re-boost it because this would lead to anticipated pull-in, and to heavier non-linear effects. Therefore, this problem can't be resolved unless we change the readout physics or working principle.

3.15.1 Force-feedback working principle

Let us follow the phase of the sinusoidal signals in the circuit below:

1. assume a 0° reference phase for the sinusoidal signal $v_t = V_t sin(\omega_{HF} t)$;
2. assume that the inertial force pushes the mass leftwards;
3. C_2 increases while C_1 decreases, thus the INA output is in phase with respect to v_t;
4. the low-pass filter output after the demodulation is thus positive;
5. thus far, this is our «standard» operation.



What if we exploit this voltage to react on the mass motion and keep it in the central position? The LPF output is duplicated with opposite sign and sent to the new electrodes; they apply a force that reacts on the motion induced by a_{ext}!

The electrostatic force applied by the right electrode is larger than the one applied by the left electrode. The **loop gain is thus negative**, as it reacts to the mass motion and tends to keep it in the central position. If the loop gain is large enough, the electrostatic force will balance the inertial action. In this way we find the sensitivity:

$$F_{\text{elec}} = \left[\frac{(V_A + V_{\text{out}}\ )^2}{2} - \frac{(V_A - V_{\text{out}}\ )^2}{2} \right] \frac{C_{0d}}{g} = \frac{2V_A V_{\text{out}}\ C_{0d}}{g} = m a_{ext}$$

$$\frac{V_{out}}{a_{\text{ext}}} = \frac{mg}{2V_A C_{0d}} = \frac{kg}{\omega_0^2 2V_A C_{0d}}$$

Advantage: as the MEMS is always close to the central position, linearity and thus FSR are extended!
Drawback: additional circuit blocks mean larger consumption!

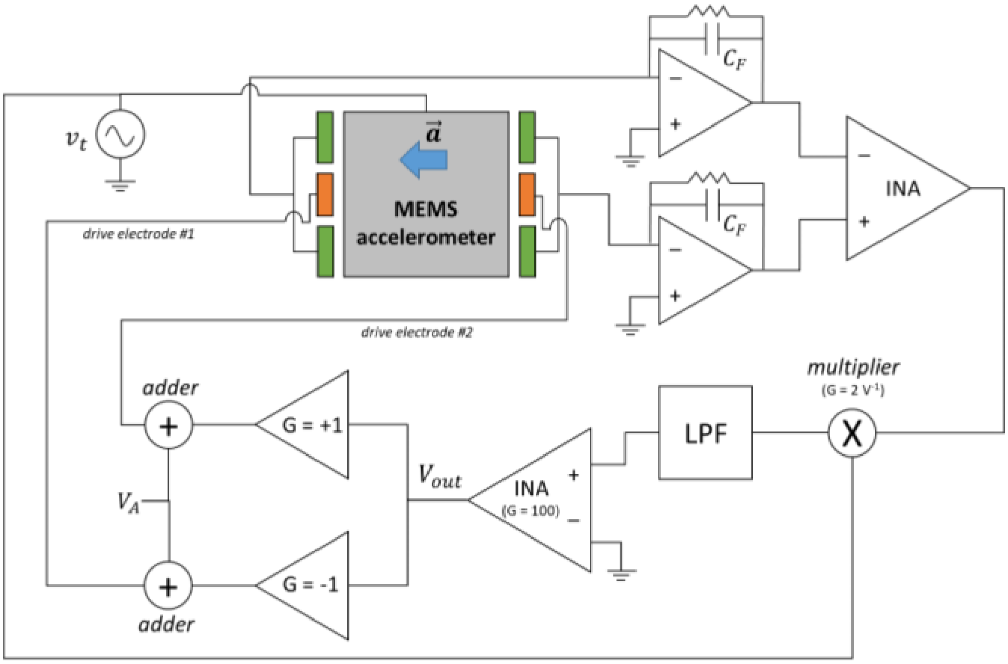


Figure 29: Example of force-feedback circuit

3.15.2 Charge-controlled readout

The starting idea is to apply a known charge amount Q_1 (e.g. via a current) and readout the corresponding voltage

$$\cancel{\Delta Q = \Delta C V_{DD}} \qquad \Delta V = \frac{Q_1}{\Delta C} \quad \text{where Q}_1 \text{ is the controlled quantity}$$

Let us start, for the sake of simplicity, from a single-ended MEMS capacitor. Let us see the difference between voltage control and charge control (C_S is a generic MEMS sense capacitance).

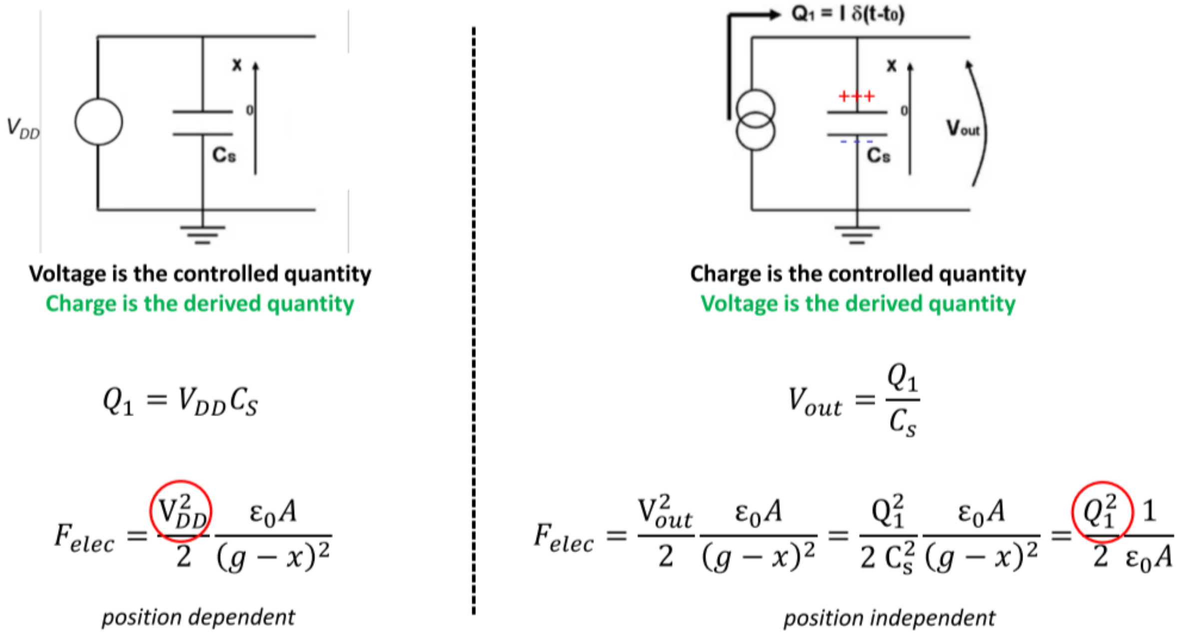
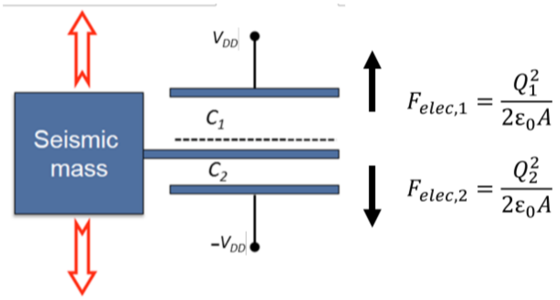


Figure 30: **Right:** Voltage-controlled **Left:** Charge-controlled

As we can see, the electrostatic force generated by a given charge quantity is **independent** of the position. If we use a differential charge-controlled system, as below, due to charge neutrality on the mass, we always have $Q_1 = Q_2$: so we can ideally obtain a **null electrostatic force** on the suspended mass! No pull-in phenomena will thus occur in this situation!



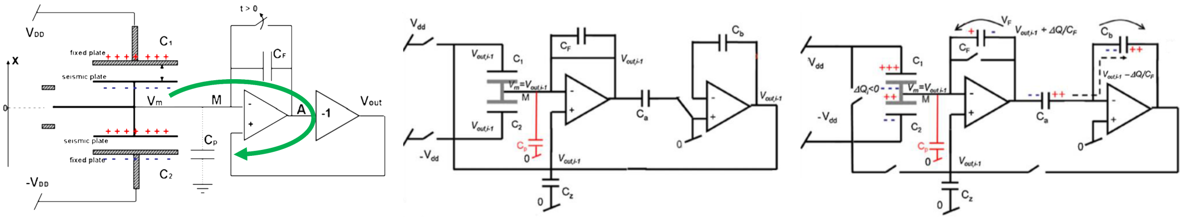
$$F_{elec,tot} = 0 \qquad if \;\; |Q_1| = |Q_2|$$

Let's compute the ideal voltage on the rotor V_m as a function of the displacement assuming the **charge neutrality condition**:

$$Q_1 + Q_2 = 0 \quad \Longrightarrow \quad -\left(V_{DD} - V_m\right) C_1 + \left(V_m + V_{DD}\right) C_2 = 0$$

$$\Longrightarrow \quad -\frac{(V_{DD} - V_m)}{g+x} + \frac{(V_m + V_{DD})}{g-x} = 0 \quad \Longrightarrow \quad \frac{V_m}{g+x} + \frac{V_m}{g-x} = \frac{V_{DD}}{g+x} - \frac{V_{DD}}{g-x}$$

$$\Longrightarrow \quad V_m \left(\frac{2g}{g^2 - x^2}\right) = V_{DD} \left(\frac{-2x}{g^2 - x^2}\right) \quad \Longrightarrow \quad V_{out} = V_m = -V_{DD} \frac{x}{g}$$



(a) Basic feedback configuration. (b) Reset phase. (c) Readout phase.

Figure 31: Charge-controlled readout axels in basic (a) and advanced (b,c) implementations.

However, up to now we have neglected the presence of **parasitic capacitances** affecting the suspended mass node of the charge-controlled configuration. As already mentioned in previous paragraphs, this presence is unavoidable, and, at least for consumer applications, the parasitic is much larger than the MEMS capacitance itself.

To sum up, the basic implementation shown in figure 31a can be synthesized in this way:

1. through a feedback mechanism, the circuit action is such that the voltage V_m is not held at a constant value: it is adjusted so that the voltage difference $|V_{DD} - V_m|$ is diminished across the capacitance whose value increased, and the voltage difference $|V_{DD} - V_m|$ is increased across the capacitance whose value diminished. This keeps the charges on C_1 and C_2 equal and opposite;
2. the quantity which ultimately generates the feedback is the rotor charge, so that the circuit is considered charge controlled;
3. the result is a minimization of the overall electrostatic force on the seismic mass.

However, we expect a dependency on the parasitic capacitance C_p in the transfer function:

$$V_m = V_{DD}\frac{x}{g}\frac{1}{1+\frac{C_P(g^2-x^2)}{2\epsilon_0 Ag}}$$

As **continuous-time** charge-controlled readout of accelerometers is critically affected by parasitics, the effective implementation is obtained using a switched-capacitor circuit operating in two phases.

The idea is to separate the moment when charge on the parasitic is updated by the feedback loop (so to cope with the desired value of V_m) from the moment when charge is readout from the rotor, using a 2-phase, switching circuit:

1. in a first phase ("**reset**"), the circuit does not apply any voltage across the MEMS capacitances, but at the same time updates the charge on the parasitic (figure 31b);
2. in a second phase ("**readout**"), it applies voltages for the readout without leaving to the system enough time for charge sharing (for motion) (figure 31c);
3. the phases are **repeated rapidly** (at a frequency » than resonance) so that between consecutive cycles the MEMS can be considered stationary.

4 RESONATOR

The next inertial sensor we would like to study is the gyroscope, which is often coupled to accelerometers in small inertial measurement units (IMUs). However, the working principle of most gyroscopes relies, as we will see, on the stable oscillation of a proof mass at its resonance, so to provide the velocity that generates the Coriolis force. As a consequence, it is convenient to analyze first **MEMS resonators** (the vibrating element) and **oscillators** (the complete system with the sustaining circuit) before gyros. MEMS-based oscillators are also used as time sensors (i.e. **clocks**).

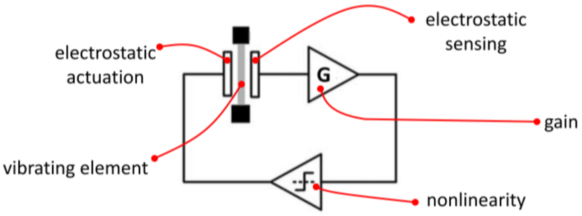
The applications are varied:

- **Timing** measurement applications (real-time clocks);
- **Synchronization** applications (HF clocks);
- **Drive** circuit in MEMS gyroscopes;
- Resonant (FM) **physical sensors**;
- Resonant (FM) **chemical sensors**.

4.1 Comb Resonator

The basic idea to implement an oscillator based on a MEMS resonator is shown below:

- one **stator** (fixed electrode) to **actuate**;
- a **suspended** resonant **element**;
- one **stator** (fixed electrode) to **sense** the motion;



The characteristic parameters of the resonator are given by the already analyzed formulas:

$$\omega_0 = \sqrt{\frac{k}{m}} \qquad Q = \omega_0 \frac{m}{b} = \frac{k}{\omega_0 b}$$

The loop oscillates (at resonance) provided that it satisfies the **Barkhausen criteria** on gain and phase:[13]

$$|G_{loop}(j\omega_0)| = 1 \qquad \angle\left(G_{\text{loop}}\ (j\omega_0)\right) = 0^\circ$$

---

[13] This will be deepened in paragraph 4.4

We first analyze the comb-finger resonator topology because it grants:

- linearity for large displacement, hence immunity to accelerations;
- absence of squeezed-film damping[14] (larger achievable Q).

Firstly, we could define the **parameters** that we are going to use:

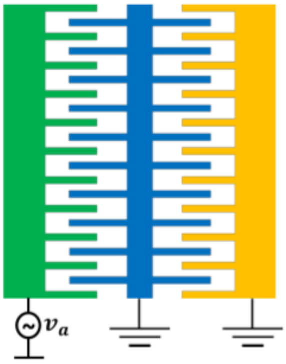
- $\mathbf{N_{CF}}$ = n. of rotor fingers (per side);
- **g** = gap between rotor and stator fingers;
- $\mathbf{L_{ov}}$ = fingers overlap length at rest;
- **m** = effective mass of the resonant element;
- **k** = effective stiffness;
- **b** = damping coefficient;
- **h** = process height;
- **A** = $h \cdot L_{ov}$, initial facing fingers area;
- $\mathbf{V_p}$ = rotor voltage;
- $\mathbf{V_A}$ = actuation voltage;
- $\mathbf{V_s}$ = sensing voltage.

Let us first calculate the obtainable actuation force applying only a voltage signal at the device resonance frequency:

$$V_A = v_a\, sin(2\pi f_o t) = v_a\, sin(\omega_0 t) \qquad V_p = 0\ V \qquad V_s = 0\ V$$

$$C_A = \frac{2\varepsilon_0 h\,(x + L_{ov})\,N_{CF}}{g} \qquad \mathrm{d}C_A = \frac{2\varepsilon_0 h N_{CF}}{g}dx$$

$$|F_{\text{elec}}| = \left|\frac{V_A^2}{2}\frac{dC_A}{dx}\right| = \frac{v_a^2\varepsilon_0 h N_{CF}}{g}\sin^2(\omega_0 t) = \frac{v_a^2\varepsilon_0 h N_{CF}}{g}\frac{1-\cos(2\omega_0 t)}{2}$$



Due to the quadratic dependence between voltage and force, we obtain an actuation at twice the resonant frequency, which will not excite the MEMS at resonance!

---

[14] Compression and decompression of an air film between moving surfaces causes a damping proportional to the facing areas.

One technique to **linearize the actuation force** is to superimpose a small AC signal to a large DC value as shown in figure 32a:

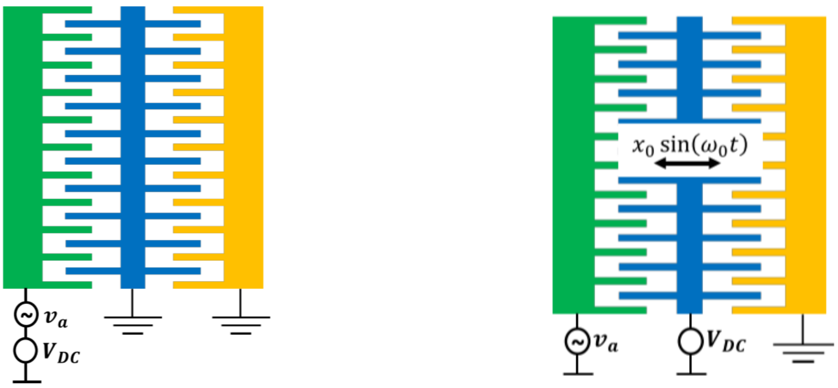
$$V_A = v_a\ sin(2\pi f_o t) + V_{DC} \qquad V_p = 0\ V \qquad V_s = 0\ V$$

$$|F_{\text{elec}}| = \frac{\left(v_a \sin\left(\omega_0 t\right) + V_{DC}\right)^2 \varepsilon_0 h N_{CF}}{g} = \frac{\varepsilon_0 h N_{CF}}{g}\left[\left(v_a \sin\left(\omega_0 t\right)\right)^2 + 2V_{DC} v_a \sin\left(\omega_0 t\right) + V_{DC}^2\right]$$

Provided that the small-signal condition is met, we obtain a linearized expression:

$$\frac{v_a^2}{2} \ll 2V_{DC} v_a \quad \rightarrow \quad v_a \ll 4V_{DC}$$

$$|F_{elec}| = \frac{\varepsilon_0 h N_{CF}}{g}\left[2V_{DC} v_a \sin\left(\omega_0 t\right) + V_{DC}^2\right]$$



(a) DC voltage applied on actuation to gain linearity.

(b) DC voltage applied on rotor to avoid offsets.

Figure 32: Basic MEMS comb resonators.

However, the force has still a **DC term** that breaks the resonator symmetry. This contribution is caused by the DC difference between the rotor and the stator. If we can have the same DC force on the opposite side, we can compensate the DC shift. This can be obtained simply by applying the **DC bias** to the **rotor**[15] as shown in figure 32b:

$$V_A = v_a\ sin(2\pi f_o t) \qquad V_p = V_{DC} \qquad V_s = 0\ V$$

$$|F_{\text{elec}}| = \left|\frac{\varepsilon_0 h N_{CF}}{g}\left[\left(v_a \sin\left(\omega_0 t\right)\right)^2 - 2V_{DC} v_a \sin\left(\omega_0 t\right) + V_{DC}^2 - V_{DC}^2\right]\right| \sim$$

$$\sim \left|\frac{\varepsilon_0 h N_{CF}}{g} 2V_{DC} v_a \sin\left(\omega_0 t\right)\right|$$

---

[15] We can get the same result biasing the sensor plate, but it is inconvenient because a large DC bias could cause disturbances in sensor's electronics, whereas the stator has no other electronics connected to it.

With the last configuration, the electrostatic force excites the MEMS at resonance, so with the maximum gain between force and displacement, without any displacement offset.
We can rearrange the terms in the found expression in the following way:

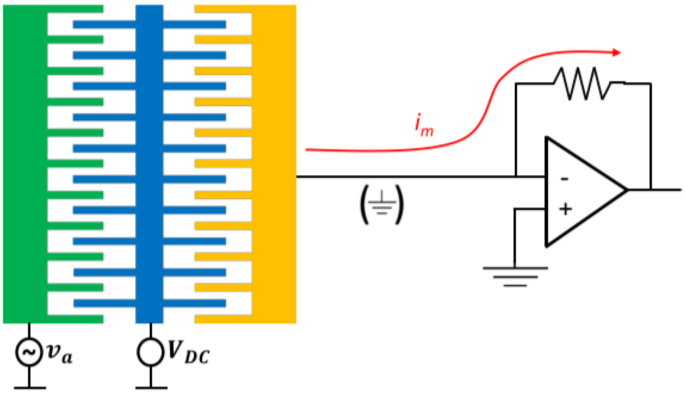
$$|F_{\text{elec}}| = \frac{2V_{DC}\varepsilon_0 h N_{CF}}{g} V_a = V_{DC}\frac{dC_A}{dx}V_a$$

So that we can define the so-called **electromechanical transduction factor** for the actuation port η_A, which relates directly the applied small voltage signal and the electrostatic force.

$$|F_{elec}| = \eta_A \, V_a \qquad \frac{F_{elec}(s)}{V_a(s)} = \eta_A$$

- η_A is a function of the resonator geometry and polarization;
- the higher η_A, the better the "**actuation**" capability of my driving stator.

To readout the rotor displacement, so to provide a feedback signal that can generate v_a and close the loop, we need to sense the current induced by the capacitance variation, while keeping the ground potential at the sense node.



We re-arrange the expression of the current using the relationship between displacement and velocity for a sinusoidal motion:

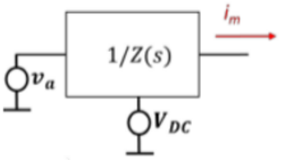
$$\dot{x} = \frac{d}{dt}\left(x_0 \sin(\omega_0 t)\right) = \omega_0 x_0 \cos(\omega_0 t)$$
$$i_m = \frac{dQ}{dt} = \frac{d(CV)}{dt} = \frac{d(V_{DC}C_S)}{dt} = V_{DC}\frac{dC_S}{dt} = V_{DC}\frac{dC_S}{dx}\frac{dx}{dt} = V_{DC}\frac{dC_S}{dx}\dot{x}$$

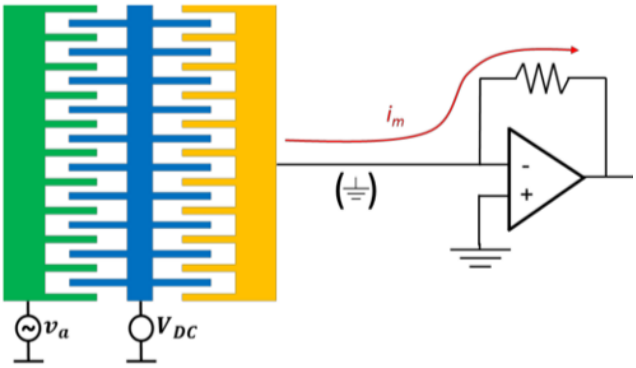
We can define an **electromechanical transduction factor** also for the sense port η_S, which, in this case, has an identical expression as the one for the drive port. Here it relates the velocity of the mass with the motional current through the sense port:

$$i_m = V_{DC}\frac{dC_s}{dx}\dot{x} = \eta_s \, \dot{x} \qquad \frac{i_m(s)}{sX(s)} = \eta_s$$

4.2 Electrical Admittance

We can define the **electrical admittance** $1/Z(s)$ between the applied voltage v_a and the output current i_m.





Note that the system has **three ports**. The model should still take into account the third port, as this has a role in determining the admittance parameters, as we will see soon. We first evaluate the force-displacement law at ω_0:

$$\frac{F_{\text{elec}}(j\omega_0)}{V_a(j\omega_0)} = \eta_A \qquad \frac{X(j\omega_0)}{F_{\text{elec}}(j\omega_0)} = \frac{Q}{jk} \text{ (see section 2.3)} \qquad \frac{i_m(j\omega_0)}{j\omega_0 X(j\omega_0)} = \eta_S$$

$$i_m(j\omega_0) = \eta_S j\omega_0 X(j\omega_0) = \eta_S j\omega_0 \frac{Q}{jk} F_{\text{elec}}(j\omega_0) = \eta_S \omega_0 \frac{Q}{k} \eta_A V_a(j\omega_0)$$

$$\frac{i_m(j\omega_0)}{V_a(j\omega_0)} = \eta_S \omega_0 \frac{Q}{k} \eta_A$$

For a symmetric resonator we have $\eta_s = \eta_A = \eta$ thus:

$$\frac{i_m(j\omega_0)}{V_a(j\omega_0)} = \eta^2 \omega_0 \frac{Q}{k} \qquad Q = \omega_0 \frac{m}{b}, \qquad \omega_0 = \sqrt{\frac{k}{m}}$$

$$\Longrightarrow \frac{i_m(j\omega_0)}{V_a(j\omega_0)} = \frac{\eta^2}{b}$$

The admittance (inverse of impedance) at resonance is real and depends only on the **damping coefficient** (no dependence on m or k)!
If we take into account the complete expression of the force to displacement transfer function, we can calculate the electrical admittance at any frequency:

$$\frac{X(s)}{F_{\text{elec}}(s)} = \frac{1/m}{\left(s^2 + s\frac{\omega_0}{Q} + \omega_0^2\right)} = \frac{1}{(ms^2 + bs + k)}$$

$$\frac{i_m(s)}{V_a(s)} = \eta^2 s \frac{1/m}{\left(s^2 + \frac{b}{m}s + \frac{k}{m}\right)} = \eta^2 s \frac{1/m}{\left(s^2 + s\frac{\omega_0}{Q} + \omega_0^2\right)}$$

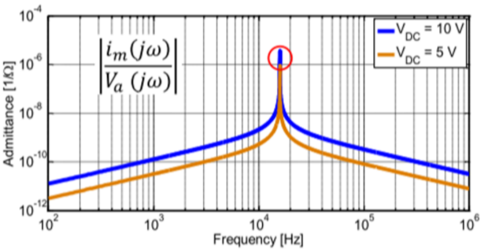
Doing some approximation we can get:

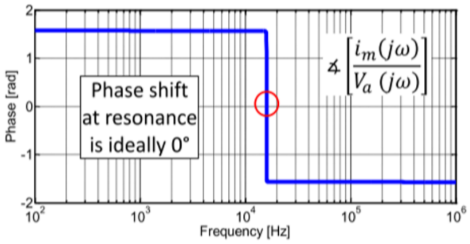
$$\Rightarrow \omega \ll \omega_0 \qquad \longrightarrow \qquad \frac{i_m(s)}{V_a} = \frac{\eta^2}{k}s$$

$$\Rightarrow \omega = \omega_0 \qquad \longrightarrow \qquad \frac{i_m(s)}{V_a} = \frac{\eta^2}{b}$$

$$\Rightarrow \omega \gg \omega_0 \qquad \longrightarrow \qquad \frac{i_m(s)}{V_a} = \frac{\eta^2}{m}\frac{1}{s}$$

The capacitive readout introduces a zero in the origin. This is followed by two complex conjugate poles at resonance (for resonators, we always assume **large quality factors**). The admittance modulus is maximized at resonance and depends strongly on V_{DC}, whereas the phase is substantially independent of V_{DC}.





4.3 Electrical equivalent model

The analysis done so far suggests the resonator to be modeled by an electrical equivalent circuit. We have indeed noticed that the behavior before, at and after resonance can be described in terms of the following expressions, each of which reminds us of a passive equivalent electrical component:

$$\frac{i_m(s)}{V_a(s)} = \frac{\eta^2}{k}s \quad \longrightarrow \quad \frac{1}{Z} = \frac{\eta^2}{k}s = sC_{eq} \quad Z = \frac{1}{sC_{eq}} \quad \longrightarrow \quad C_{eq} = \frac{\eta^2}{k}$$

$$\frac{i_m(s)}{V_a(s)} = \frac{\eta^2}{b} \quad \longrightarrow \quad \frac{1}{Z} = \frac{\eta^2}{b} = \frac{1}{R_{eq}} \quad Z = R_{eq} \quad \longrightarrow \quad R_{eq} = \frac{b}{\eta^2}$$

$$\frac{i_m(s)}{V_a(s)} = \frac{\eta^2}{m}\frac{1}{s} \quad \longrightarrow \quad \frac{1}{Z} = \frac{\eta^2}{m}\frac{1}{s} = \frac{1}{sL_{eq}} \quad Z = sL_{eq} \quad \longrightarrow \quad L_{eq} = \frac{m}{\eta^2}$$

So, we can write:

$$\frac{i_m(s)}{V_a(s)} = \frac{1}{\left(\frac{m}{\eta^2}s + \frac{b}{\eta^2} + \frac{k}{s\eta^2}\right)} = \frac{1}{\left(L_{eq}s + R_{eq} + \frac{1}{sC_{eq}}\right)}$$

The 3-port resonator can be fully modeled by an electrical equivalent 2-port model (series RLC)! All the parameters are a function of V_{DC} which represents the third (hidden) port.

It is interesting to note that there is a direct relationship between the dissipative terms in the mechanical domain (dissipation by damping) and in the electrical domain (Joule effect dissipation in the resistance).

$$\frac{i_m(s)}{V_a(s)} = \frac{s}{\left(\frac{m}{\eta^2}s^2 + \frac{b}{\eta^2}s + \frac{k}{\eta^2}\right)} = \frac{s}{\left(L_{eq}s^2 + R_{eq}s + \frac{1}{C_{eq}}\right)}$$

We will see that what we need to compensate, to build up an oscillator, is represented by its losses, i.e. its dissipation, so its equivalent resistance. The sustaining circuit will need a resistive gain equal to $1/R_{eq}$ in the regime condition. The importance of an accurate modeling of the damping coefficient b (and thus of the Q factor) becomes thus evident for the correct design of a resonator.

The equivalent electrical model is very useful when we need to couple the MEMS with an oscillating circuit in a simulation phase.

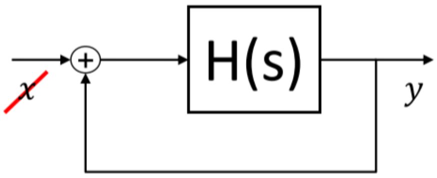
The 3-port resonator seems not that different from an accelerometer from the point of view of the structure(two fixed electrodes and a moving frame). However, huge differences are in the design parameters (e.g. no need for large mass, generally higher resonance frequencies, very large quality factors) and in the operation (oscillation within an electronic loop).

Note again the parallelism between **resistance** and **damping** and between **capacitance** and **spring stiffness**.

4.4 Oscillator conditions

After the introduction of the resonant element, implemented as a MEMS transducer, we can introduce the sustaining circuits to form a full oscillator. We have seen before that the loop will oscillate provided that it satisfies the **Barkhausen criteria** on gain modulus and phase.

Let's assume the feedback loop shown below, where $H(s)$ represents the combined **loop gain** from the resonator and the electronics.



We know that for this to be an oscillator, it must be self-sustained without any input signal x!

$$y = H(s)(\cancel{x} + y) \quad \Longrightarrow \quad y = H(s)y \quad \Longrightarrow \quad H(s) = 1$$

This condition is satisfied only if $H(s)$ has unitary modulus and no phase shift.

In the regime operation, the loop gain shall satisfy the conditions:[16]

$$\begin{cases} \left|G_{loop}\left(j\omega_0\right)\right| = 1 = 0dB \\ \angle\left(G_{loop}\left(j\omega_0\right)\right) = 0^\circ \end{cases}$$

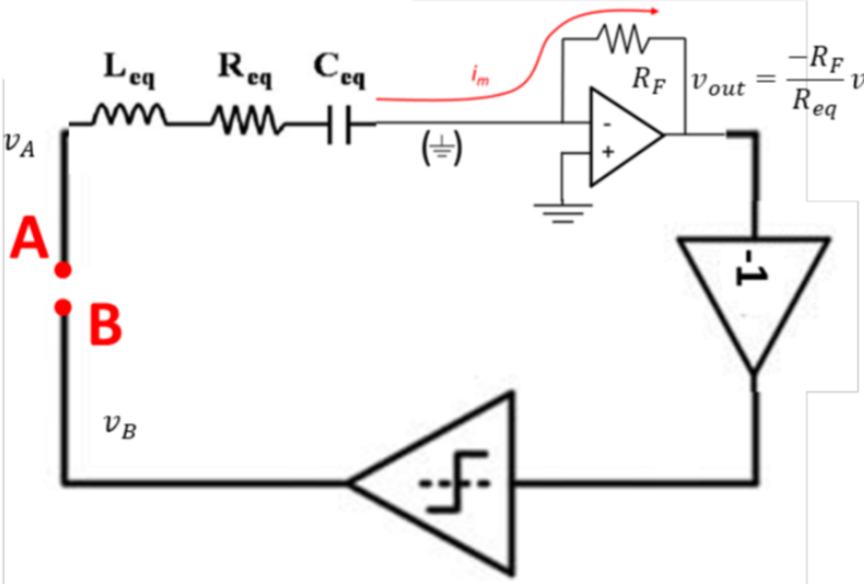
---

[16] Note that any frequency that satisfies this criterion is self-sustained, thus we can exploit an high Q to have just ω_0 to self-sustain.

However, in the initial phase for the oscillation to build up, the conditions should be:

$$\begin{cases} \left|G_{loop}\left(j\omega_0\right)\right| > 1 > 0dB \\ \angle\left(G_{loop}\left(j\omega_0\right)\right) = 0° \end{cases}$$

We have seen how a resonator can be modeled as an **equivalent RLC** circuit. The dissipative term is represented by the equivalent resistance R_{eq}.



Looking at the figure above that includes the electrical model, and considering that at resonance the model simplifies into the resistive term, we see that the gain between nodes A and B **at resonance** using a transimpedance stage is:

$$G_{\text{loop}}\ (j\omega_0) = \frac{v_B}{v_A} = \frac{R_F}{R_{eq}}$$

The condition $|G_{loop}(j\omega_0)| > 1$ at the startup is satisfied only if the overall circuit gain is initially larger than R_{eq}.
However, it's not possible to exactly match $|G_{loop}| = 1$ in a linear circuit and we know that:

- with $|G_{loop}(j\omega_0)| < 1$ the oscillation **never starts**;
- with $|G_{loop}(j\omega_0)| > 1$ the oscillation **diverges**.

The idea is to have $|G_{loop}(j\omega_0)| > 1$ at the start-up and then to add a **non-linearity** (i.e. a change in gain as a function of signal amplitude) in the electronics, so to hold $|G_{loop}(j\omega_0)| = 1$. The non-linear stage has a gain which is not constant as a function of the input signal amplitude. For example, a **buffer** (unitary gain stage), which saturates after the oscillation reaches a specific amplitude, can be considered as a non-linear stage, and will be widely adopted in the following of this presentation. Also a **comparator** (e.g. implemented as an open-loop operational amplifier) or a **high-gain amplifier** stage have as well the behavior described by the graphs in figure 33, but with a very high gain in a narrow linear region.

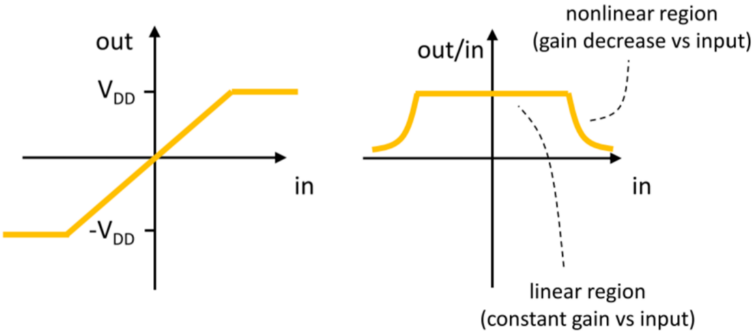
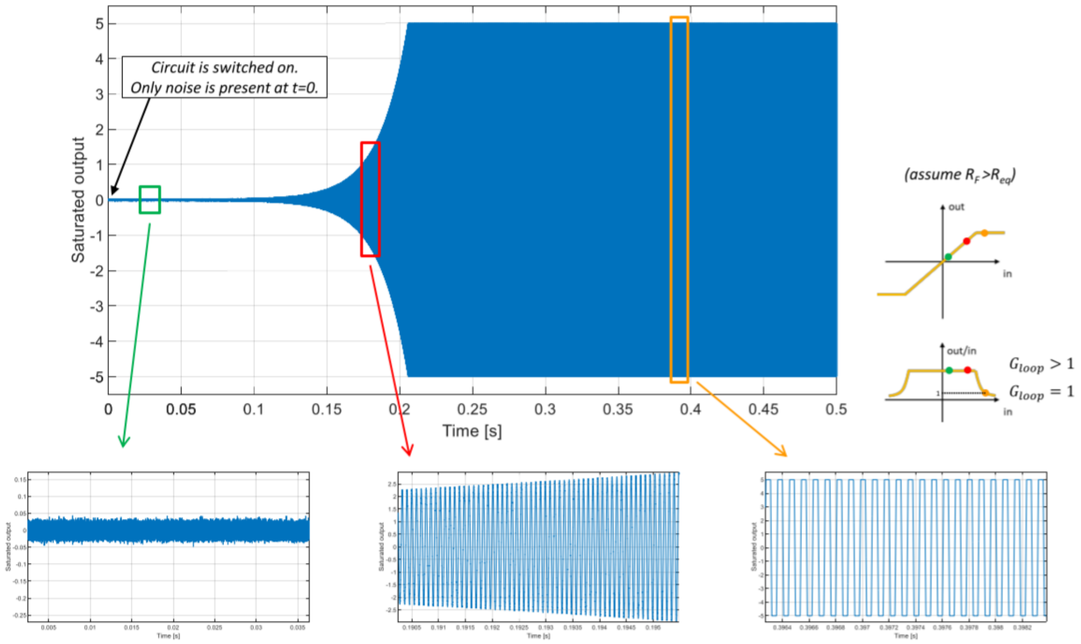


Figure 33: Gain behavior as the input varies

We can summarize everything we have said in this picture:



Noise is dominant. Within all the frequency components, the one corresponding to ω_0 begins to grow, hidden in a generic noisy signal.

The component at resonance begins to dominate over the others (it is the only one satisfying the start-up condition). The circuit loop gain is still > 1.

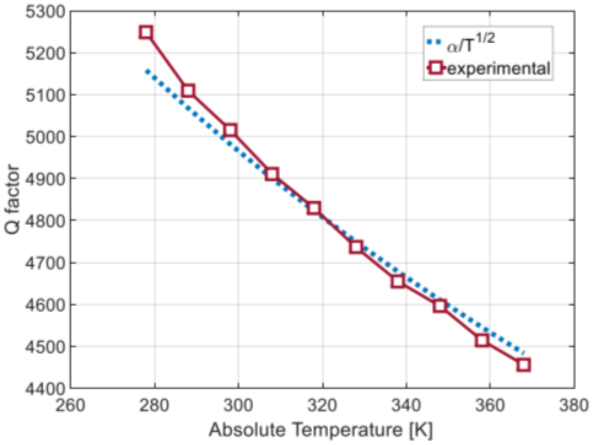
The circuit reaches the saturation at one node. The circuit gain decreases from "> 1" to "= 1" and the oscillation stabilizes. Turn-on time depends on Q.

Note that this kind of resonance is possible thanks to resonator's transfer function that has one zero at the origin and the two complex conjugate poles at ω_0. Therefore, if the quality factor Q is high enough, all frequencies except ω_0 are rejected, leading to the desired condition.

Note: Q variation with temperature The damping coefficient and thus R_{eq} and the Q factor are functions of the absolute temperature. In particular, we will see that the Q factor decreases with T, with a power law that depends on the pressure regime. In the most common situation, Q decreases with the square root inverse of the absolute temperature:

$$Q \propto \frac{1}{\sqrt{T}} \qquad b = \frac{\omega_0 m}{Q} \qquad R_{eq} = \frac{b}{\eta^2}$$

The relevant consequence is that, if a resonator needs to operate over a given T range, its equivalent resistance (which changes with b) must be compensated for every T value. The worst condition is the highest T (lowest Q, highest R_{eq}).



4.5 Circuit based on linear amplification and saturation

The circuit solution shown below, based on a transresistance front-end, is probably the simplest approach to a "feedback" oscillator.

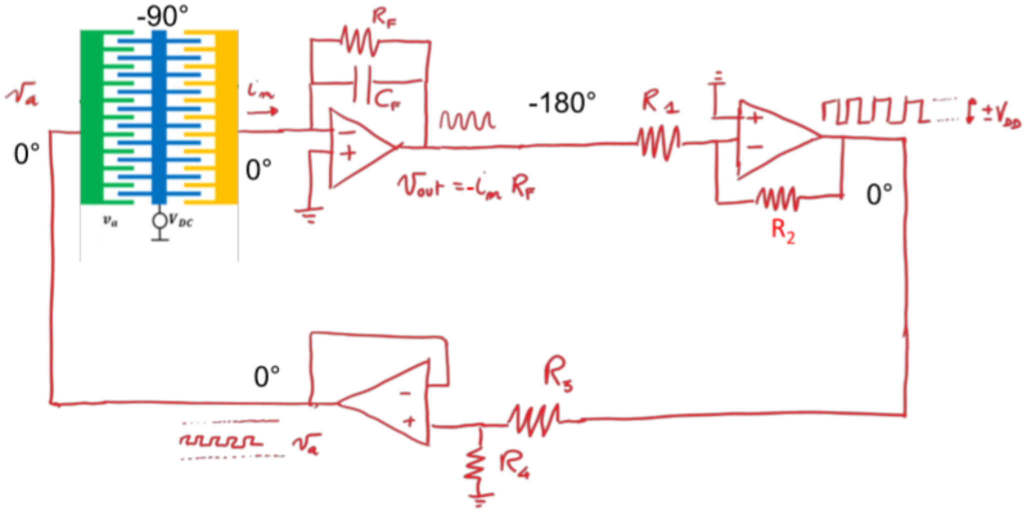
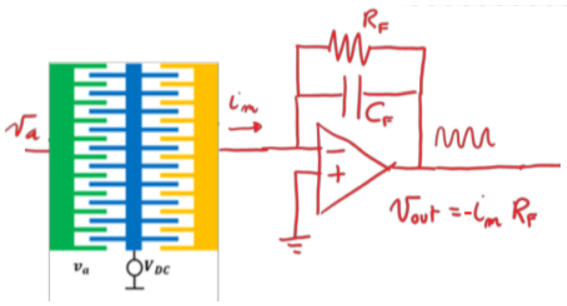


Figure 34: The shown phase lags refer to signals at resonance.

The **TIA** (*TransImpedence Amplifier*) has the advantage of giving a 180° shift from v_a to v_{out}, easy to complete with an inverting stage so to reach the 360° lag. If the gain given by R_F is not enough to compensate R_{eq}, a **high-gain stage** (linear at the start-up) allows the oscillation to rise up from noise. Remember that noise has all frequency components: only those corresponding to ω_0 are amplified by the quality factor and see a loop gain larger than one at the start-up.

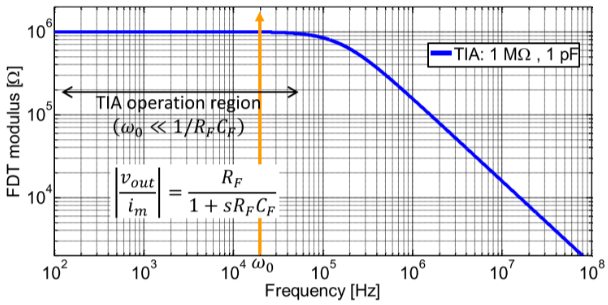
Let's now analyse the transfer function of the 3 stages one by one:

- **Transimpedance Amplifier (TIA)**
  We have to guarantee that the poles of the network goes after the resonance frequency ω_0.

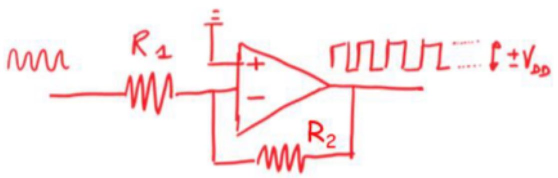


$$\frac{v_{\text{out}}}{i_m}(s) = \frac{-R_F}{1 + sR_FC_F} \qquad\qquad \frac{v_{\text{out}}}{v_a}(\omega_0) \approx -\frac{R_F}{R_{eq}}$$

R_F represents the gain of the stage. In principle a large value is convenient to minimize Johnson noise and the impact of noise of the following stages. C_F (real or parasitic) introduces a pole, whose relevance will be clarified partly later when discussing feedthrough effects. Below we could see the Bode diagram of FdT modulus:



- **High-gain stage**



As R_F only may not be enough to compensate the resonator losses, the overall (startup) loop gain is now set by:

$$\frac{v_{\text{out },2}}{i_m}(s) = \frac{R_F}{1 + sR_FC_F}\frac{R_2}{R_1} \qquad\qquad \frac{v_{\text{out},2}}{v_a}(\omega_0) \approx \frac{R_F}{R_{eq}}\frac{R_2}{R_1} > 1$$

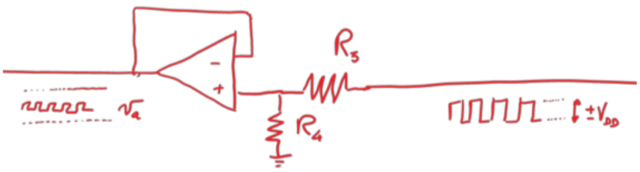
Assuming that this condition is met, when we switch on the circuit, at time $t = 0$ we have noise only, i.e. harmonic components at every frequency. The harmonic corresponding to ω_0 is the only one which is amplified by a positive loop gain larger than one. The oscillation at ω_0 begins to increase. At a certain point, the high-gain stage saturates its output. The loop gain thus decreases, stabilizing to 1.

- **Amplitude limitation stage**
  We have seen how the MEMS resonator can be linearized only when having:

$$\frac{v_a^2}{2} \ll 2\, V_{DC}\, v_a \qquad \Longrightarrow \qquad \frac{v_a}{4} \ll V_{DC}$$

  We need thus to lower a bit the driving amplitude, e.g. by simply using a resistive divider. Note that the loop design needs to take into account this lowering:

$$G_{\text{loop}}\ (\omega_0) = \frac{R_F}{R_{eq}} \frac{R_2}{R_1} \frac{R_4}{R_3 + R_4} > 1$$



  The buffer is used to drive the MEMS with a low output impedance.

Finally the overall gain with a TIA configuration is:

$$G_{\text{loop}}\ (s) = \frac{1}{L_{eq}s + R_{eq} + \frac{1}{sC_{eq}}} \; \frac{R_F}{1 + sR_F C_F} \; \frac{R_2}{R_1} \; \frac{R_4}{R_3 + R_4}$$

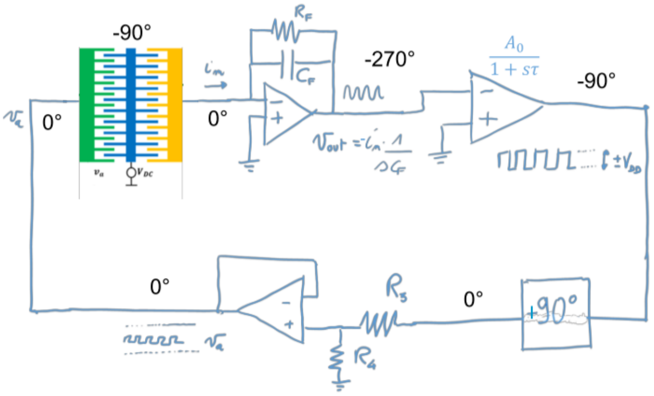
For power and noise constraints, it would be nice to have a large R_F:

$$\sqrt{\frac{4k_BT}{R_F}\, BW} = i_{\text{noise}} \qquad \text{compared with} \qquad i_m$$

However, large integrated resistances are not feasible (large → too much area). One possible solution could be to use a MOS transistor. The idea is to use a MOS **switched off** and exploit the large resistance of the channel. Unfortunately, its value is hard to predict and highly depends on the bias voltage which may fluctuate. This is a problem for **TIA-based oscillators**, as it directly causes gain changes for large R_F values. Often, a **charge amplifier** approach is thus preferred.

4.6 Circuits based on Comparators

In the figure below the high-gain is substituted by a **comparator**, designed as an open-loop OpAmp and thus having just the gain A_0.

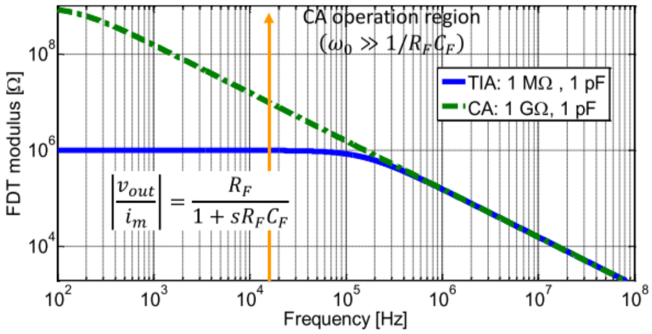


We adopt the approach based on a **charge amplifier** (integrator): we thus need to recover a 90° phase shift to satisfy the Barkhausen condition. Let's analyze the 3 stages separately:

- **Charge Amplifier**
  In this case the stages is designed to have the feedback pole before the resonance frequency ω_0. A resistance is still needed to bias the OpAmp and avoid saturation due to integration of its bias currents, but it must have a large value.

$$\frac{v_{\text{out}}}{i_m}(s) = -\frac{R_F}{1+sR_FC_F} \qquad\qquad \frac{v_{\text{out}}}{v_a}(\omega_0) \approx -\frac{1}{sC_FR_{eq}}$$

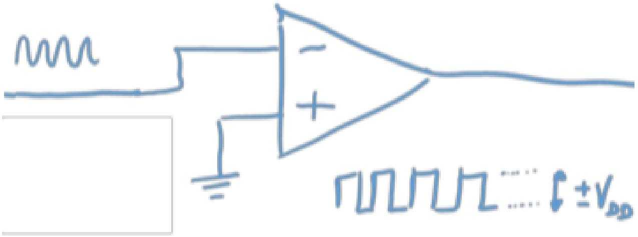
The feedback resistance can be made large through a MOS transistor in off state so that its noise contribution becomes negligible and noise is dominated by the OpAmp.



We have seen that resistances implemented through MOS are not easily predictable nor repeatable. However, this has no impact on the CA circuit gain at resonance, if the following condition is verified:

$$pole = \frac{1}{2\pi R_FC_F} \ll f_0$$

- **Comparator**
  The comparator can be just implemented as an operational amplifier with no feedback. The rise and fall times of the square wave will be determined by the OpAmp slew rate.



  The advantage using a comparator is that it can consume very little power compared to a full high-gain operational amplifier. Nevertheless a comparator needs a low input offset to correctly square an input sinusoid (and to provide correct start-up from noise).

- **Amplitude limitation and phase adjustment**
  Amplitude limitation is still required, exactly for the same reasons seen in the other case. A +90° phase shift is also required to match the phase condition (Barkhausen criteria), as the charge amplifier introduces -90° more phase shift. This phase shift can be obtained using a **derivator** (or an integrator plus an inverting stage). As an alternative, a *phase locked loop* (PLL) block is used to add the required phase shift.

To conclude, an oscillator is a system that relies on a resonating element and on suitable sustaining circuit. Its optimization is another perfect example of co-design between the device and the electronics.
We could underline some trade-offs arising in the co-design:

- a **high Q** would be desirable as the electronics **gain** could be **lower** (power dissipation would correspondingly decrease). However, a very high Q usually demands for low pressures, which become **poorly repeatable**;
- a **high bias voltage on the rotor** would **increase** the **transduction** factors: however, the generation of high-voltage references (e.g. through charge pumps) requires power consumption;
- **device constraints** on the electronics are also imposed by the need to operate at relatively low displacements to **avoid non-linearities**.

4.7 Feedthrough capacitance

The resonator model that we used so far does not account for possible (often unavoidable) **parasitic electrical components** affecting the nodes where the MEMS is electrically connected to the sustaining circuit. These nodes see large parasitic capacitances to ground and above all, depending on the goodness of the design, may see also a direct coupling in between them.
To make the ideal model more realistic, we take into account parasitic electrical elements that affect the resonator:

- **capacitances** (parasitic and drive/sense capacitances at rest) between ground and another low impedance (they can be typically neglected);
- **parasitic resistance** to ground, typically large values (they can be neglected);
- **parasitic capacitance** directly coupling the actuation port to the sense port (**feedthrough capacitance**). Usually, even small values (usually 0.1 fF to 10 fF) are sometimes critical for the resonator behavior.

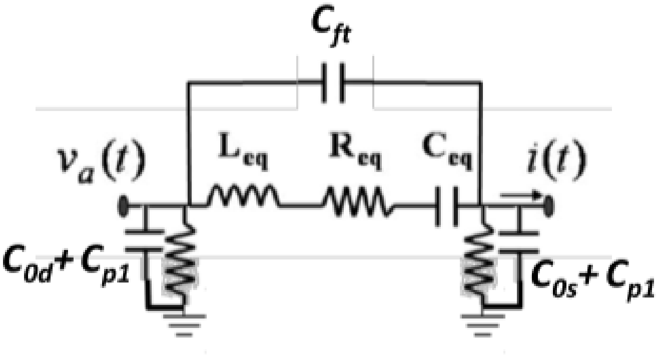


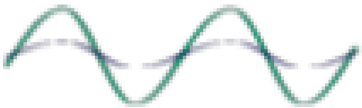
Figure 35: Representation of the feedthrough capacitance

We therefore need to modify the found analytical expression of the electrical equivalent admittance, taking into account the parasitic feedthrough capacitance:

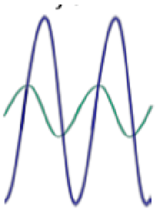
$$\frac{i(s)}{V_a(s)} = \frac{i_m(s) + i_{ft}(s)}{V_a(s)} = \frac{1}{\left(L_{eq}s + R_{eq} + \frac{1}{sC_{eq}}\right)} + sC_{ft}$$

As we can see the feedthrough capacitance adds a contribution which grows with ω and becomes eventually dominant for large frequencies. Therefore, the value of the feedthrough capacitance can be either negligible or relevant, depending on the frequency.

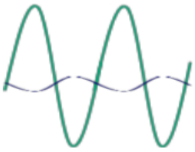
- **Low Frequencies**
  The feedthrough is usually larger than the equivalent capacitance. The **phase is the same** as for the ideal model.
  $$\frac{i(s)}{V_a(s)} = s\,C_{eq} + s\,C_{ft}$$



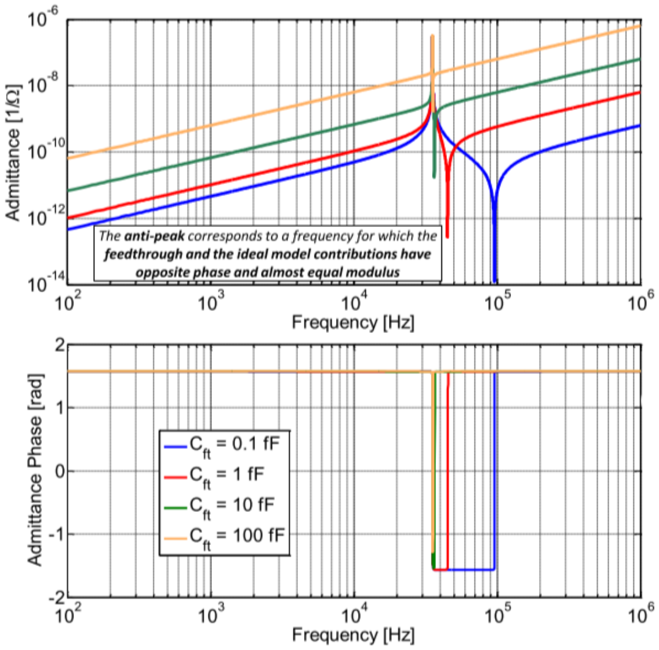
- **Around the resonance frequency**
  As far as the **modulus** of the feedthrough contribution is **smaller than** $1/R_{eq}$, the phase goes as in the ideal model. After ω_0, when $|sC_{ft}|$ is again larger than $1/R_{eq}$, the phase is again dominated by the feedthrough.
  $$\frac{i(s)}{V_a(s)} = \frac{1}{R_{eq}} + sC_{ft}$$



- **High frequency**
  The phase is dominated by the feedthrough as its contribution is much larger than $\frac{1}{s\,L_{eq}}$.
  $$\frac{i(s)}{V_a(s)} = \frac{1}{s\,L_{eq}} + sC_{ft}$$



Now we can see the graphical result with the Bode plot of the modulus and phase:



As we can see, the **peak** value almost does not change but there is a characteristic **anti-peak** (when FT and RLC sum with opposite phase and similar modulus). The phase returns to π and for large C_{ft} values, the phase drop may be lower than π (risk of no oscillation if it does not cross 0).

4.8 Effects of feedthrough on electronic circuits

In a realistic situation, the analyzed circuits will thus see a parasitic capacitance directly coupling drive and sense ports.

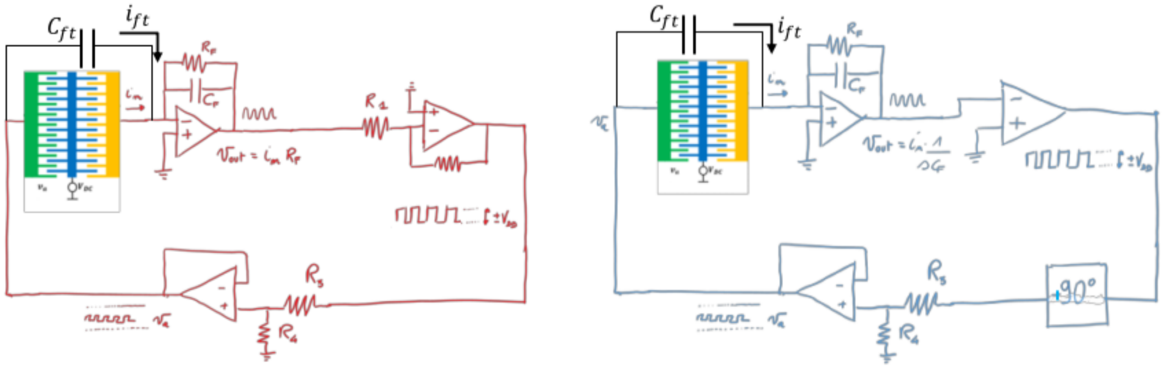


Figure 36: **Right:** Oscillator with TIA **Left:** Oscillator whit CA

We have seen how the C_{ft} introduces a contribution i_{ft} that couples with the motional current i_m, increasing with the frequency. We now analyze the effects of the capacitance C_{ft} on the loop gain and phase of the whole oscillator. To calculate it, it's convenient to **open the loop** in a point which sees a high impedance, so to avoid the need to reconstruct the impedance after opening. For both the analyzed topologies, a convenient point is the positive input of the buffer.

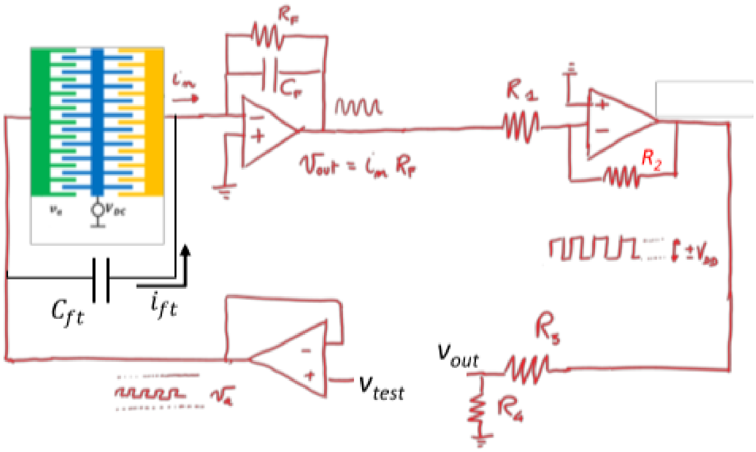


Figure 37: Open loop for the computation of the gain loop

$$G_{\text{loop}}(s) = \frac{v_{\text{out}}}{v_{\text{test}}} = \left(\frac{\eta^2}{m} \frac{s}{\left(s^2 + \frac{b}{m}s + \frac{k}{m}\right)} + sC_{ft} \right) \frac{-R_F}{1 + sR_F C_F} \frac{-R_2}{R_1} \frac{R_4}{R_3 + R_4}$$

Plotting the loop gain we can notice that $s\,C_{ft}$ (green curve) generates an anti-peak in the modulus and a return to π of the phase.

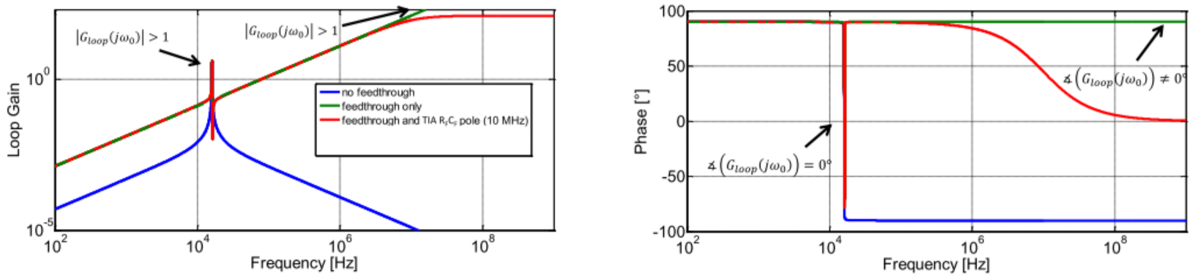


Figure 38: Bode plot of modulus and phase of the gain loop

As a consequence $G_{loop} > 1$ for several frequencies other than ω_0. However, the phase condition is not met, so there is no risk of spurious oscillations. Yet, the circuit will have (at least) the TIA pole (see the red curve now). This generates a -90° phase shift which leads the phase to 0° in a region where the loop gain is much larger than 1. Therefore, a lot of other frequencies can oscillate: the circuit will not work as desired!

A way to limit this effect is the introduction of additional poles in the circuit given e.g. by R_2C_2 (where C_2 is placed in parallel to R_2): this is convenient for low power applications, as no extra dissipation is added. The idea is to lower the loop gain right after the resonance peak, to avoid chances that it grows larger than 1 for other frequencies. As the phase shift introduced by an RC electronic pole (-90°) begins much before the pole frequency, care should be taken to avoid effects on the phase lag at resonance (i.e. the pole should be ideally placed at least one decade beyond resonance). The loop-gain expression now includes the TIA and 2nd stage poles:

$$G_{\text{loop}}\left(s\right)=\frac{v_{\text{out}}}{v_{\text{test}}}=\left(\frac{\eta^2}{m}\frac{s}{\left(s^2+\frac{b}{m}s+\frac{k}{m}\right)}+sC_{ft}\right)\frac{-R_F}{1+sR_FC_F}\frac{-R_2}{R_1\left(1+sR_2C_2\right)}\frac{R_4}{R_3+R_4}$$

In this example we see how challenging this is by setting a pole pair at 10 MHz (TIA) and at 1 MHz (added), we still have a clear oscillation at $\sim$ 3 MHz (in purple).

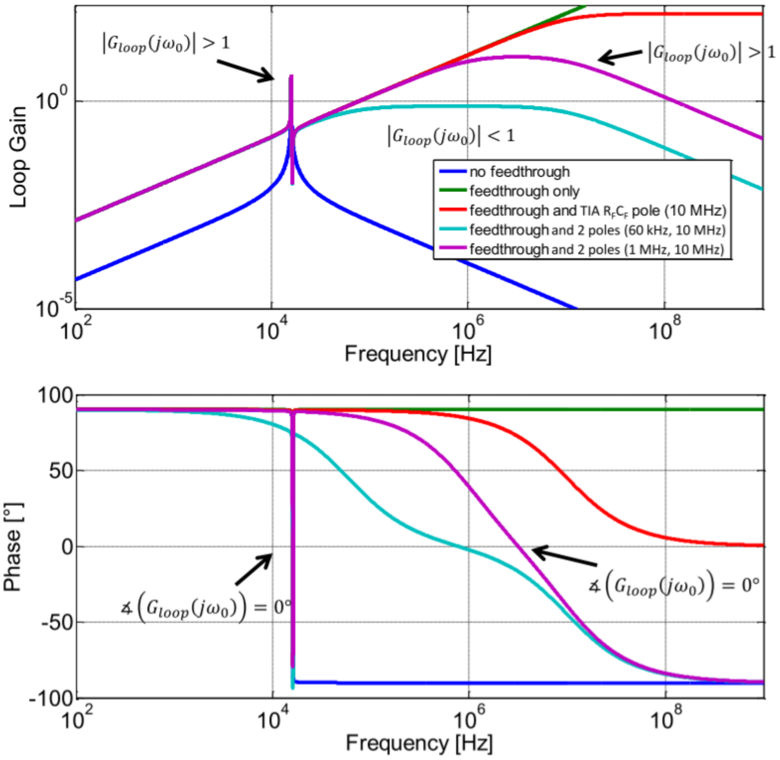


Figure 39: Bode plot of modulus and phase of the gain loop

Only if we set one pole at 60 kHz, not that far from resonance, we get the overall gain $<$ 1 where the phase crosses the 0° value. However, this causes a **slight phase shift** at resonance, and the circuit oscillates at a frequency different from the peak one.

The phase shift induced by the proximity of the poles is visible on the phase (not symmetric with respect to $\pm\pi$ around resonance). This makes the phase cross the $0°$ point at a frequency different from resonance. Provided that the shift is small, so that the loop gain is still larger than 1 (0 dB), the circuit still oscillates, though at a frequency slightly different from ω_0.

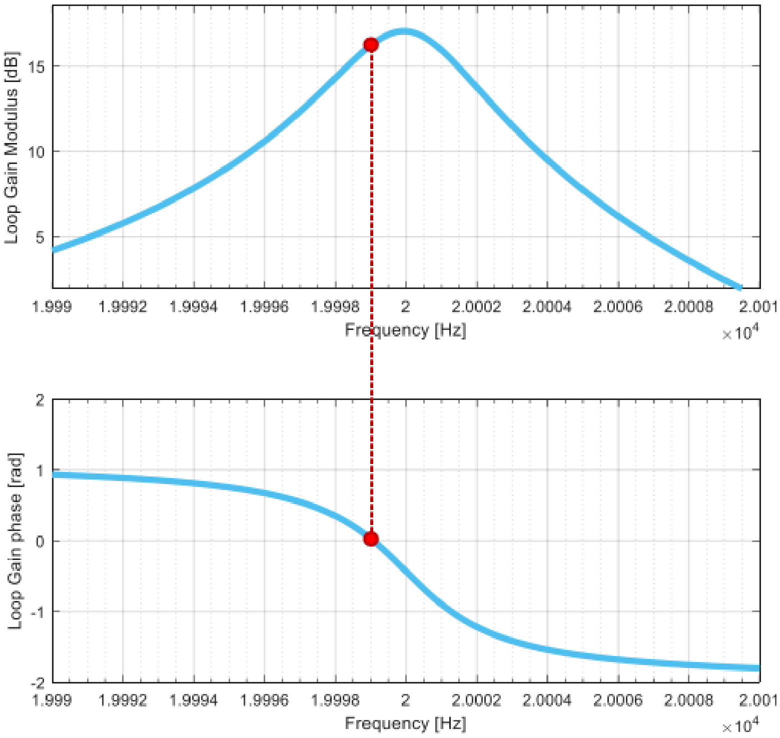


Figure 40: Representation of the phase shift

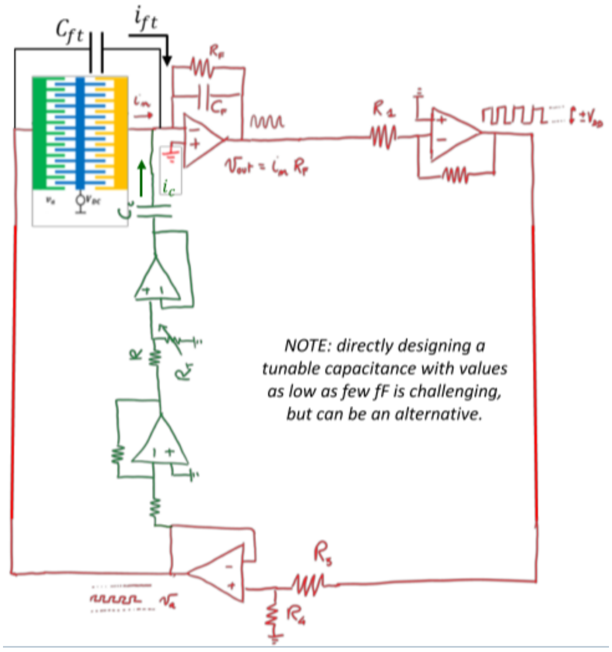
An alternative solution for feedthrough is to compensate it at the origin, i.e. to sum at the sense node a current opposite to the feedthrough contribution. This can be obtained by suitably tuning a capacitive gain, excited by a current which is in anti-phase with respect to the feedthrough signal:

$$i_C = C_C \frac{dv_a}{dt}\left(-\frac{R_T}{R_T+R}\right)$$

Where the tunable resistance R_T is selected to obtain:

$$i_C = C_C \frac{dv_a}{dt}\left(-\frac{R_T}{R_T+R}\right) = -i_{ft} = -C_{ft}\frac{dv_a}{dt}$$

$$\longrightarrow C_C\left(\frac{R_T}{R_T+R}\right) = C_{ft}$$



NOTE: directly designing a tunable capacitance with values as low as few fF is challenging, but can be an alternative.

Figure 41: Feedthrough compensation circuit

The figure below shows the effect of tuning the feedthrough compensation capacitance:

- **Under-compensation**: still an anti-peak after the resonance peak;
- **Over-compensation**: the anti-peak anticipates the resonance peak.

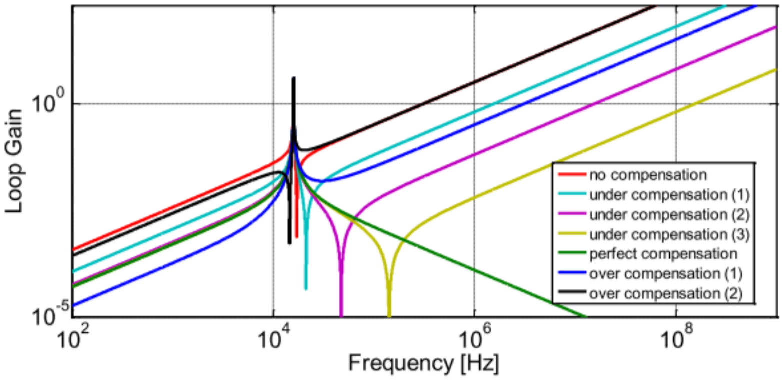
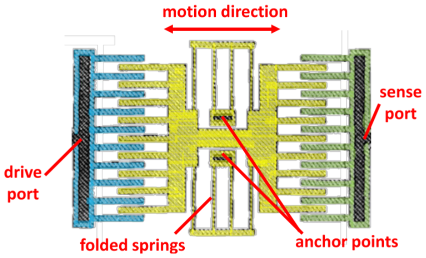


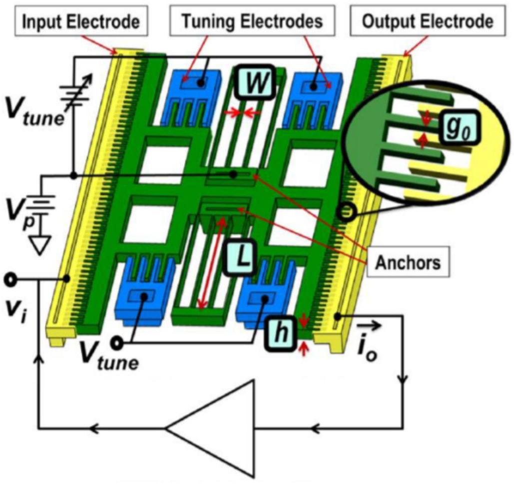
Figure 42: Examples of Under and Over Compensation

4.9 Tang Resonator

A common resonator at relatively low frequency (< 100 kHz) is the so-called **Tang configuration**. It is based on what we have just seen. The springs are designed as four 2-fold beams. A motivation for this type of suspension is its large capability for relief of built-in residual stress and temperature stress in the structural film: the folded beams are anchored near the center, thus allowing expansion or contraction of the four beams along the y-axis.



For several applications, it is mandatory to accurately match a specific frequency. We have seen how DRIE etching may induce larger/smaller dimensions than expected, which result in beam width and resonance frequency deviations. A way to **tune** a resonator frequency is to **exploit electrostatic softening** (only down-tuning is possible), so the design typically targets larger frequencies and a downward tuning. This tuning is obtained by means of parallel plates, which are not used for actuation and sensing but just to tune the frequency.



Designing N_P parallel plates to obtain an overall capacitance C_P, having a gap g_{PP} from the rotor arms, the obtainable frequency shift is given by the formula aside and can be selected by changing the voltage difference between rotor (whose voltage is chosen to match the desired h) and parallel-plate tuning stators.

$$\Delta\omega = \sqrt{\frac{k_{tot}}{m}} - \sqrt{\frac{k - 2\left(V_{DC} - V_{tune}\right)^2 \frac{C_{PP}}{g_{PP}^2}}{m}}$$

5 GYROSCOPE

While accelerometers are sensors of linear motion, gyroscopes represent the category of inertial sensors of **rotational motion**. Though there exist gyroscopes that measure directly an angle, most MEMS gyros measure (or at least give an estimate of) the angular rate (or angular velocity), relying on the **Coriolis acceleration** (angular accelerations, instead, give rise to too small signals). The application are varied:

- Automotive (ESC, roll-over detection, smart tyre);
- Consumer (gaming, VR);
- Imaging/industrial (camera motion stabilization, platform stabilization);
- Mapping (indoor navigation);
- Medical (vestibular disease rehabilitation, motion monitoring);
- Military/aerospace;
- Aircraft active control.

5.1 Coriolis Force

We have seen in paragraph 2.1 that the acceleration of an object is given by these contributions:

$$\vec{a}_{Pa} = \vec{a}_{Pr} + \vec{a}_{O'a} + \cancel{(\overrightarrow{\dot{\Omega}_{O'a}} \times \vec{r}_{Pr})} + \cancel{\vec{\Omega}_{O'a} \times (\vec{\Omega}_{O'a} \times \vec{r}_{Pr})} + 2(\vec{\Omega}_{O'a} \times \vec{v}_{Pr})$$

Where in typical use cases, centripetal acceleration and angular acceleration are negligible compared to the Coriolis acceleration.

Let's assume a system which is **rotating only**. Newton's law in our inertial frame (or any other inertial reference) is:

$$\vec{F}_{Pa} = m\,\vec{a}_{Pa}$$

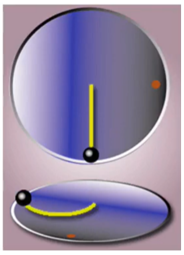


Figure 43: **Top**: Trajectory in fixed system **Bottom**: Trajectory in red dot system

In the **non-inertial frame** (rotation only):

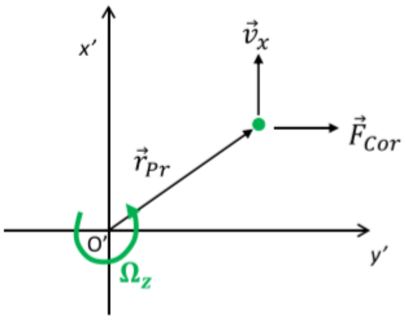
$$\vec{\boldsymbol{F}}_{\boldsymbol{Pr}} = m\vec{a}_{Pr} = m\vec{a}_{Pa} - m\vec{a}_{O'a} - 2m\left(\vec{\Omega}_{O'a} \times \vec{v}_{Pr}\right) = \vec{\boldsymbol{F}}_{Pa} + \vec{\boldsymbol{F}}_{Acc} + \vec{\boldsymbol{F}}_{\text{Cor}}$$

The MEMS mass is is suspended through springs and packaged in a gaseous encapsulation environment (ambient air). Thus, the true forces to consider for the point-like mass are the elastic force and the damping force. Assuming that effects of accelerations are rejected, in a gyroscope we obtain:[17]

$$m\ddot{\vec{x}}_{Pr} + b\dot{\vec{x}}_{Pr} + k\vec{x}_{Pr} = -\cancel{m\vec{a}_{O'a}} - 2m(\vec{\Omega}_{O'a} \times \dot{\vec{y}}_{Pr})$$

5.1.1 How to measure the angular rate

Let's assume a velocity $\vec{v}_x$ and, for sake of simplicity, $\vec{a}_{O'a} = 0$ (our gyroscope will be designed to be immune to acceleration). The measurement of $\vec{F}_y = \vec{F}_{Cor}$ (obtained by measuring the displacement y) allows to determine $\vec{\Omega}_z$ if the value of the velocity $\vec{v}_x$ in the relative system is known.



$$\vec{F}_y = -2m\left(\vec{\Omega}_z \times \vec{v}_x\right)$$

$$m\ddot{y} + b\dot{y} + ky = -2m\Omega_z v_x$$

The system will therefore feature motion along two directions, and thus two spring-mass-damper systems (two modes of interest): one per motion direction.

5.2 General architecture of a single-mass gyroscope

There are two motion directions of interest in a MEMS gyroscope:

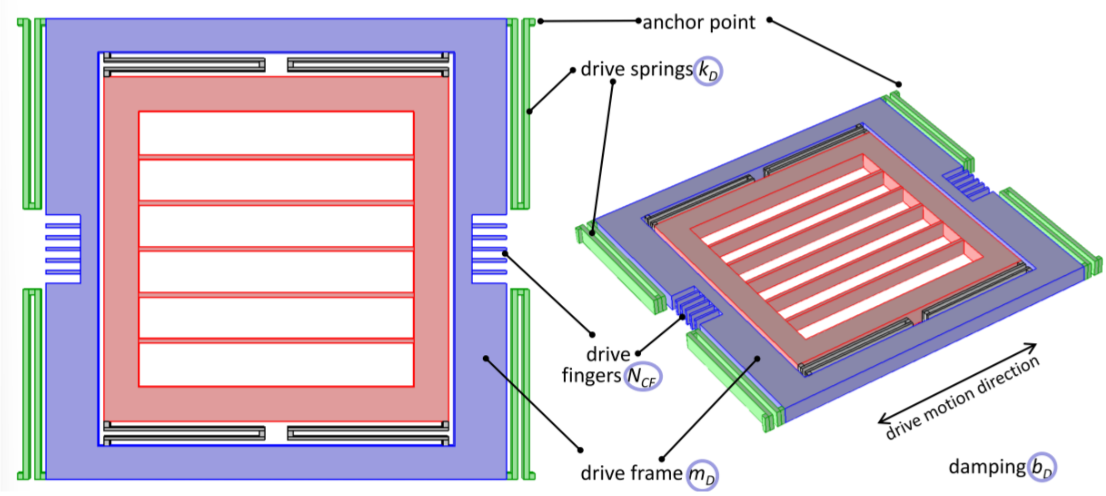
- The **drive mode direction**, along which the device is excited by an electrostatic oscillation to create the desired **drive velocity** v_x (which will be sinusoidal);
- The **sense mode direction**, along which the device (or part of it) moves (only under angular rates) to sense the **Coriolis force** $\vec{F}_y$ (which also will be sinusoidal);

Note that both directions are orthogonal to the angular rate sensing direction (z).

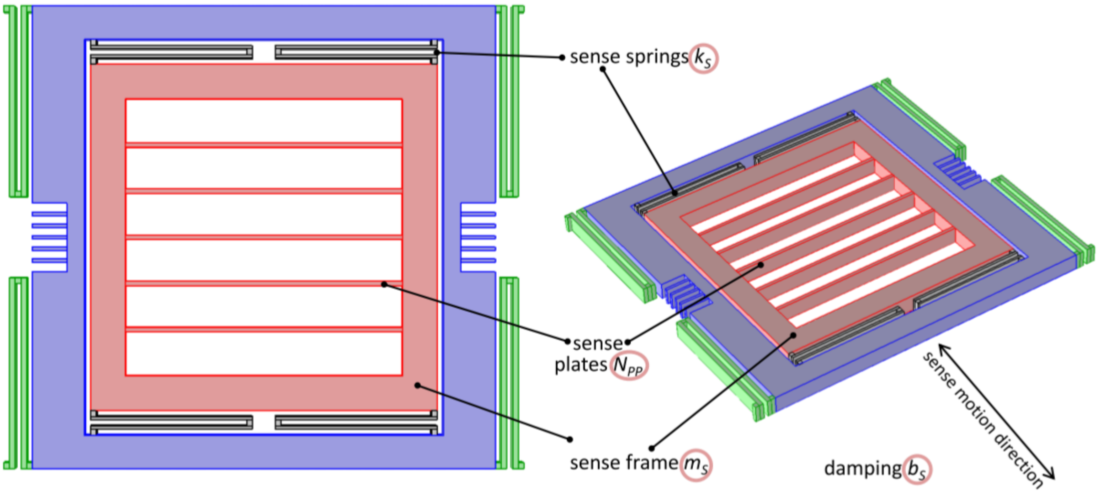
---

[17] Note that the Coriolis force is perpendicular to the mass' velocity (cross-product), thus we need at least 2 degrees of freedom and 3 dimensions for our model.

In the simplest implementation, the **drive mode** is anchored to the substrate by means of **drive springs**, which allow motion in the x-direction and are quite rigid in the y direction. The drive mode is kept in **resonance oscillation**, usually via comb fingers. The main block of the drive mode is called the **drive frame**.



The **sense frame** is attached to the drive frame through another set of **decoupling springs**. These springs are quite rigid in the drive direction, so that the sense mode is perfectly dragged by the drive mode. However they allow the sense frame to move also in the orthogonal (y) direction.



The working principle at a glance is the following:

- The **drive mode** is kept in oscillation along the **X-direction**: both frames ($m_d + m_s$) move together with a sinusoidal velocity $\vec{v}_{Pr}$;
- In presence of an angular rate in the Z-direction, both frames experience a Coriolis force in the **Y-direction**; however the drive frame is rigid in that direction and will not move, whereas the **sense frame** displacement will be revealed (e.g. via capacitive sensing) and will provide the information on the Coriolis force.

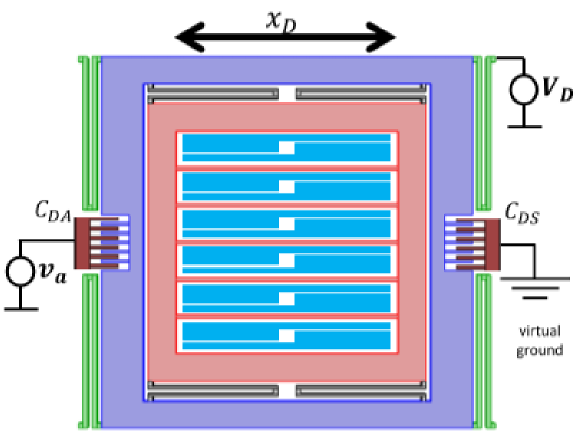
In a simplified vision, a gyroscope can be seen as a combination of a comb-driven resonator and a parallel-plate accelerometer of the Coriolis acceleration.

5.3 Sensitivity Calculation

The calculation of the sensitivity follows these consecutive steps:

5.3.1 Calculation of the drive electrostatic force

The calculation of the electrostatic force follows exactly what we have seen for a comb-driven, comb-sensed resonator, biased with a rotor DC voltage much larger than the actuation AC voltage. note that the sense frame is consequently biased at the same voltage (frames are short circuited through springs!).



$$C_{DA} = \frac{2\varepsilon_0 h\left(L_{ov} - x\right)N_{CF}}{g} \qquad C_{DS} = \frac{2\varepsilon_0 h\left(L_{ov} + x\right)N_{CF}}{g} \qquad |F_{elec}| = \frac{\Delta V^2}{2}\frac{\partial C}{\partial x}$$

$$|F_{elec}| = \frac{\varepsilon_0 h N_{CF}}{g}\left[\left(v_a \sin\left(\omega_0 t\right)\right)^2 + 2V_{DC}v_a \sin\left(\omega_0 t\right) + V_{DC}^2 - V_{DC}^2\right] \sim \frac{\varepsilon_0 h N_{CF}}{g} 2V_{DC}v_a \sin\left(\omega_0 t\right)$$

For the sake of simplicity, just consider the phasor amplitude:

$$F_{elec,0} = \frac{2\varepsilon_0 h N_{CF}}{g} V_{DC}\, v_a$$

5.3.2 Calculation of the drive displacement and velocity at resonance and the Coriolis Force

We assume now that the drive mode resonator is self-sustained at resonance (e.g. through an oscillator circuit). We thus assume that the transfer function between force on the drive mode and drive displacement is fully amplified by the quality factor: $|T_D(j\omega_0)| = \frac{Q_D}{k_D}$.

$$x_{D,0} = F_{\text{elec },0}\frac{Q_D}{k_D} = \frac{Q_D}{k_D}\frac{\varepsilon_0\ h\ N_{CF}}{g}\ 2\ V_{DC}\ v_a$$

We can now calculate the time derivative[18] to obtain the velocity and, in turn, the Coriolis force. We simplify the cross-product as we consider the angular rate component orthogonal to drive motion.

$$v_D = \dot{x}_D \quad \rightarrow \quad v_{D,0} = \omega_0 x_{D,0} = \omega_0 \frac{Q_D}{k_D}\frac{\varepsilon_0\; h\; N_{CF}}{g}\; 2\; V_{DC}\; v_a$$

$$F_{\text{Cor}} = m_S \cdot a_{Cor} = m_S \cdot 2(\vec{v}_D \times \vec{\Omega}) = 2m_S \cdot v_D \cdot \Omega$$

$$\rightarrow \quad F_{\text{Cor, 0}} = 2m_S \cdot \omega_0 \frac{Q_D}{k_D}\frac{\varepsilon_0\; h\; N_{CF}}{g}\; 2\; V_{DC}\; v_a \cdot \Omega = 2\; m_S\; x_{D,0}\; \omega_0\; \cdot\; \Omega$$

5.3.3 Calculation of the sense mode displacement

We are now interested in calculating the motion in the Y-direction caused by the Coriolis force on the sense frame. As a first case-study, we assume the sense mode to be at resonance frequency ω_0 (matched with the drive mode). The sense displacement is thus amplified by the quality factor of the sense mode Q_S:

$$y_{S,0} = \frac{F_{Cor,0}}{k_S} Q_S = 2m_S x_{D,0}\omega_0 \frac{Q_S}{k_S} \cdot \Omega \qquad \xrightarrow{Q_S = \frac{k_S}{\omega_0 b_S}} \qquad y_{S,0} = 2\; m_S\; x_{D,0}\; \frac{1}{b_S} \cdot \Omega$$

We can finally evaluate the expression for the sensitivity:

$$\frac{y_{S,0}}{\Omega} = \frac{x_{D,0}}{b_S/2m_S}$$

We neglected the sign as we are just looking at the AC amplitude and we will see that this quantity $\frac{\omega_0}{2Q} = \frac{b}{2m} = \Delta\omega_{BW}$ corresponds to the sensing bandwidth at resonance. Once we take into account the assumptions given upon we finally get:

$$\frac{y_{S,0}}{\Omega} = \frac{x_{D,0}}{\Delta\omega_{BW}}$$

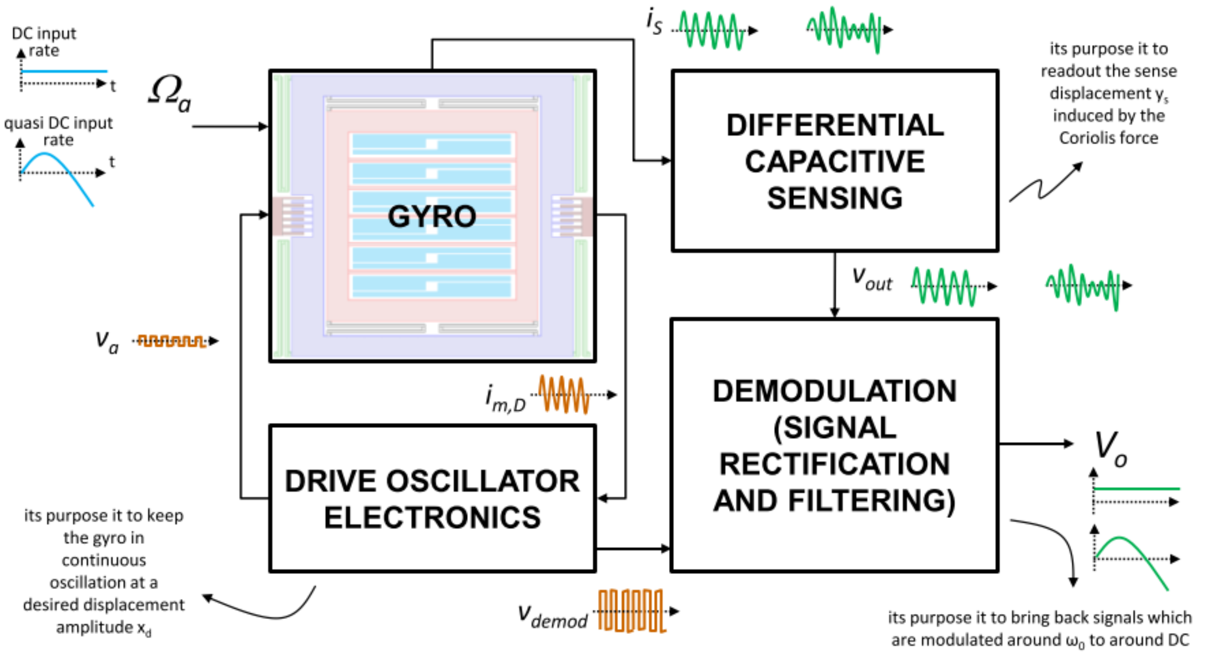
Note that we thus look for large drive motion amplitude and we see a first trade-off between sensitivity and bandwidth.

---

[18] Since we are considering amplitudes of sinusoidal signals, the derivative is done just by multiplying the amplitude by the signals' pulse.

5.3.4 Drive Loop

Regarding the electronic circuit we can expect an overall amplitude-modulated system formed by the following building blocks:

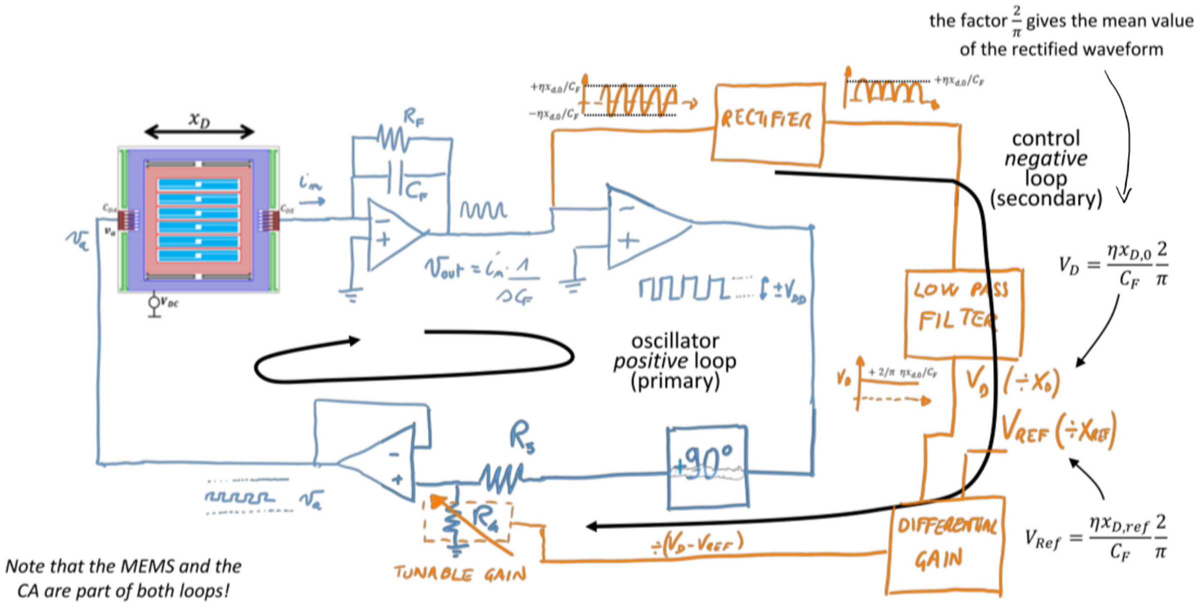


We have seen that the X-axis drive motion $x_{D,0}$ depends on the drive quality factor Q_D, which is relatively strong function of the absolute temperature. Therefore, also the drive displacement $x_{D,0}$ and thus the overall sensitivity suffer from changes caused by temperature variations (even in operation!). It is thus compulsory to build an amplitude control circuit that keeps the motion amplitude $x_{D,0}$ stable against Q factor changes. As a result we need an electronics topology with an actuation proportional to the displacement so that we can operate an amplitude modulation. Since the **CA topology** (paragraph 4.6) gives an output proportional to the displacement[19] x_D, we choose this one and we couple it with another negative loop to stabilize $x_{D,0}$. In particular, we have to:

- **take the AC** (sine) signal from the **CA output** (proportional to $x_{D,0}$);
- **rectify and low-pass** filter → we now have a DC signal proportional to $x_{D,0}$;
- **compare it** with a reference V_{REF}, related to the displacement x_{REF} you want to set;
- use the **difference between** x_{REF} and $x_{D,0}$ to change a tunable gain in the primary loop (e.g. change the voltage divider, acting on a tunable resistor);
- the **negative feedback** brings $x_{D,0}$ to the "*virtual ground*" x_{REF} through G_{loop2}.

---

[19] First stage's output tension is proportional to velocity for transimpedance amplifier topologies ($v_a = \eta \dot{x}_D R_F \propto \dot{x}_D$), while it is proportional to displacement for charge amplifier topologies ($v_a = \eta \frac{\dot{x}_D}{\omega C_F} = \eta \frac{x_D}{C_F} \propto x_D$)



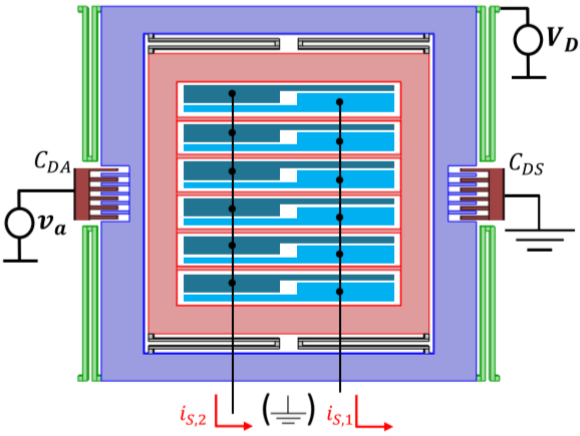
5.3.5 Calculation of the sense capacitance variation

Let us verify the choice of using parallel plates for the gyroscope sensing through a numerical count on the maximum sense displacement. Remember that parallel-plate sensing will be preferable to comb-finger if displacements $y_{S,0}$ are small compared to the gap (linearity).

$$y_{S,0} = \frac{x_{D,0}}{\Delta\omega_{BW}}\,\Omega$$

As the rotor is kept to V_{DC}, we keep the virtual ground of the readout stages to 0 V. The current flowing through each sense port can be written as:

$$i_{S,0i} = V_{DC}\frac{dC_{S,i}}{dt} = V_{DC}\frac{dC_{S,i}}{dy}\frac{dy}{dt} = V_{DC}\frac{dC_{S,i}}{dy}y_{S,0}\omega_0$$



We can calculate the sense capacitance variation for small displacements assuming N_{PP} sense cells of area A_{PP} and gap g:

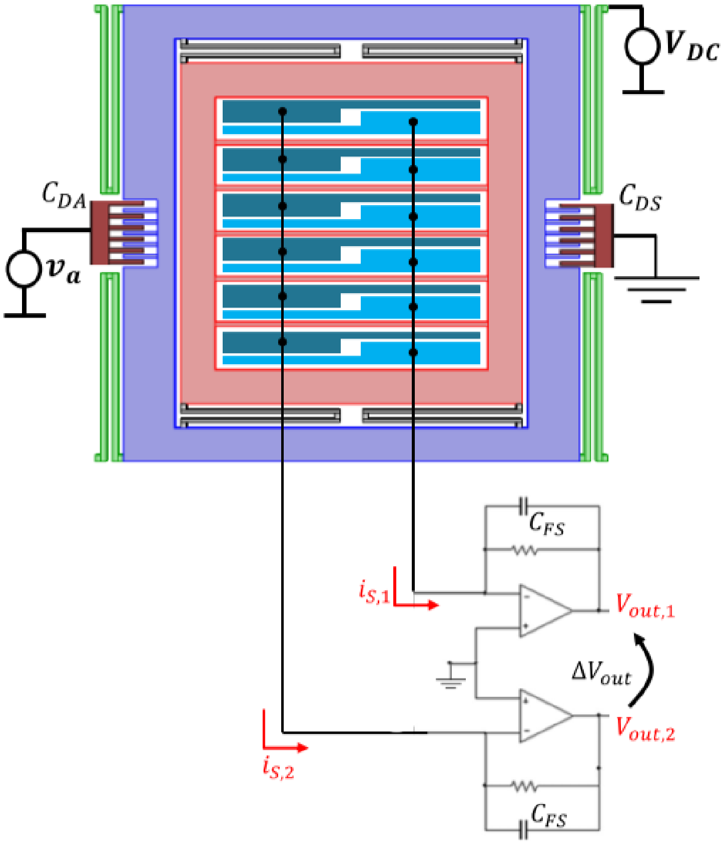
$$C_{S,1} = \frac{\varepsilon_0 A_{PP} N_{PP}}{g + y} \qquad\qquad C_{S,2} = \frac{\varepsilon_0 A_{PP} N_{PP}}{g - y}$$

$$\left| \frac{dC_{S,i}}{dy} \right| \approx \frac{\varepsilon_0 A_{PP} N_{PP}}{g^2} = \frac{C_S}{g}$$

5.3.6 Calculation of the output voltage per unit rate (sensitivity)

The current is integrated by each charge amplifier:

$$V_{\text{out },0i} = \left| \frac{i_{S,0i}}{j\omega_0 C_{FS}} \right| = V_{DC} \frac{dC_{S,i}}{dy} \frac{y_{S,0}}{C_{FS}} = \frac{V_{DC}}{C_{FS}} \frac{C_S}{g} y_{S,0}$$



So that we can evaluate the final expression for the (differential) sensitivity in small-displacement approximation:

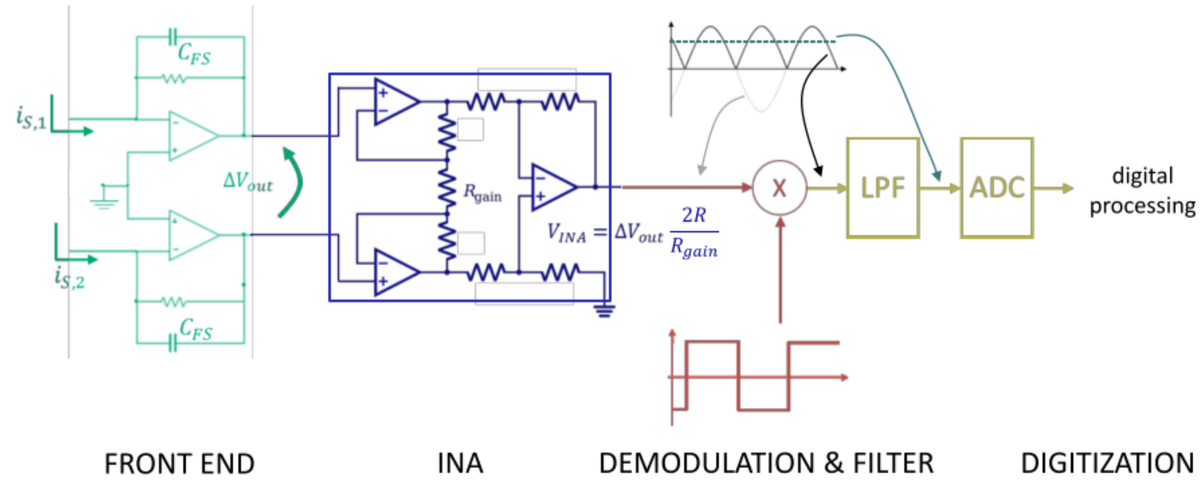
$$\frac{\Delta V_{\text{out },0}}{\Omega} = 2 \frac{V_{DC}}{C_{FS}} \frac{C_S}{g} \frac{x_{D,0}}{\Delta\omega_{BW}}$$

Note that at this point $\Delta V_{out,0}$ is still a signal modulated at $\omega_0 = \omega_D$!

5.4 Demodulation, filtering and ADC

Before being processed by the software that uses the angular rate information (e.g. apps in a smartphone, or the electronic control unit in a car), the signal needs to return to a DC value, which will be digitized by an ADC.
This occurs through the following 4 steps:

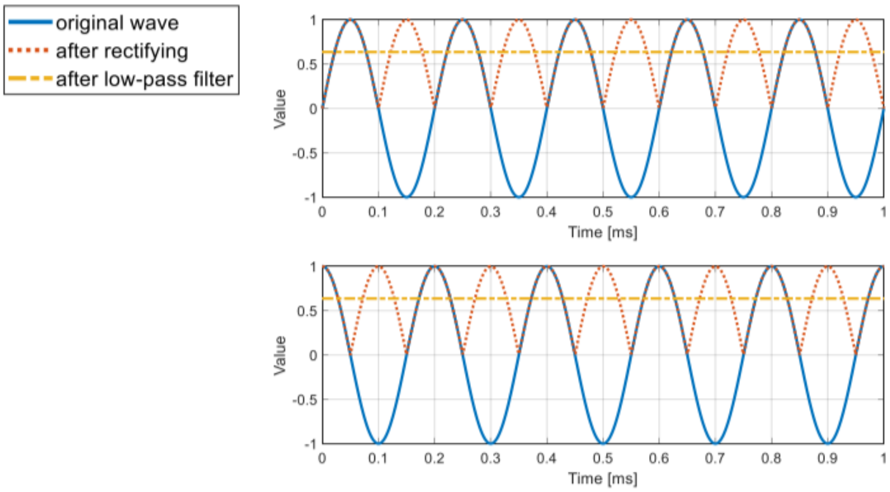
1. turn the differential AC signal into a single-ended AC signal, using e.g. an instrumentation amplifier (INA);
2. multiply the AC signal by a square wave at the same frequency and phase using the square wave from the drive loop;
3. filter the high-frequency component to save the DC value;
4. acquire the so obtained DC value with the ADC. The required number of bit is given by the ratio of ± FSR/resolution.



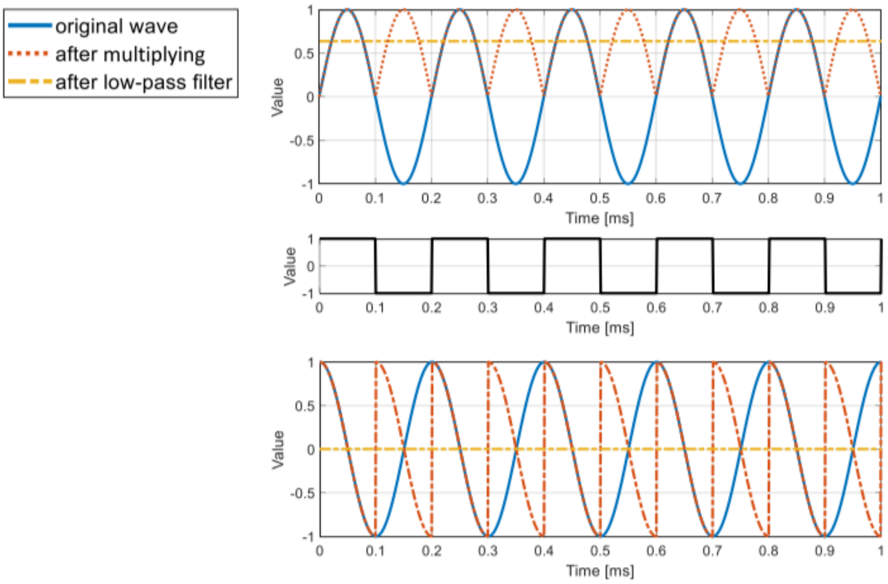
Why demodulation? The demodulation brings an AC signal modulated at ω_0 back to base-band (i.e. around DC), holding information about its amplitude. In the figure above, it is operated a synchronous demodulation, which is explained in the next page.

There are two ways of demodulating a signal:

- **Asynchronous** Demodulation
  It consists in rectifying (taking the absolute value) and filtering the signal, but it doesn't take into account the phase of the signal to demodulate (same result for different phases, i.e. also for different signs). As we can see from the figure below the two blue waves are phase shifted by 90° and the demodulated signal is identical.

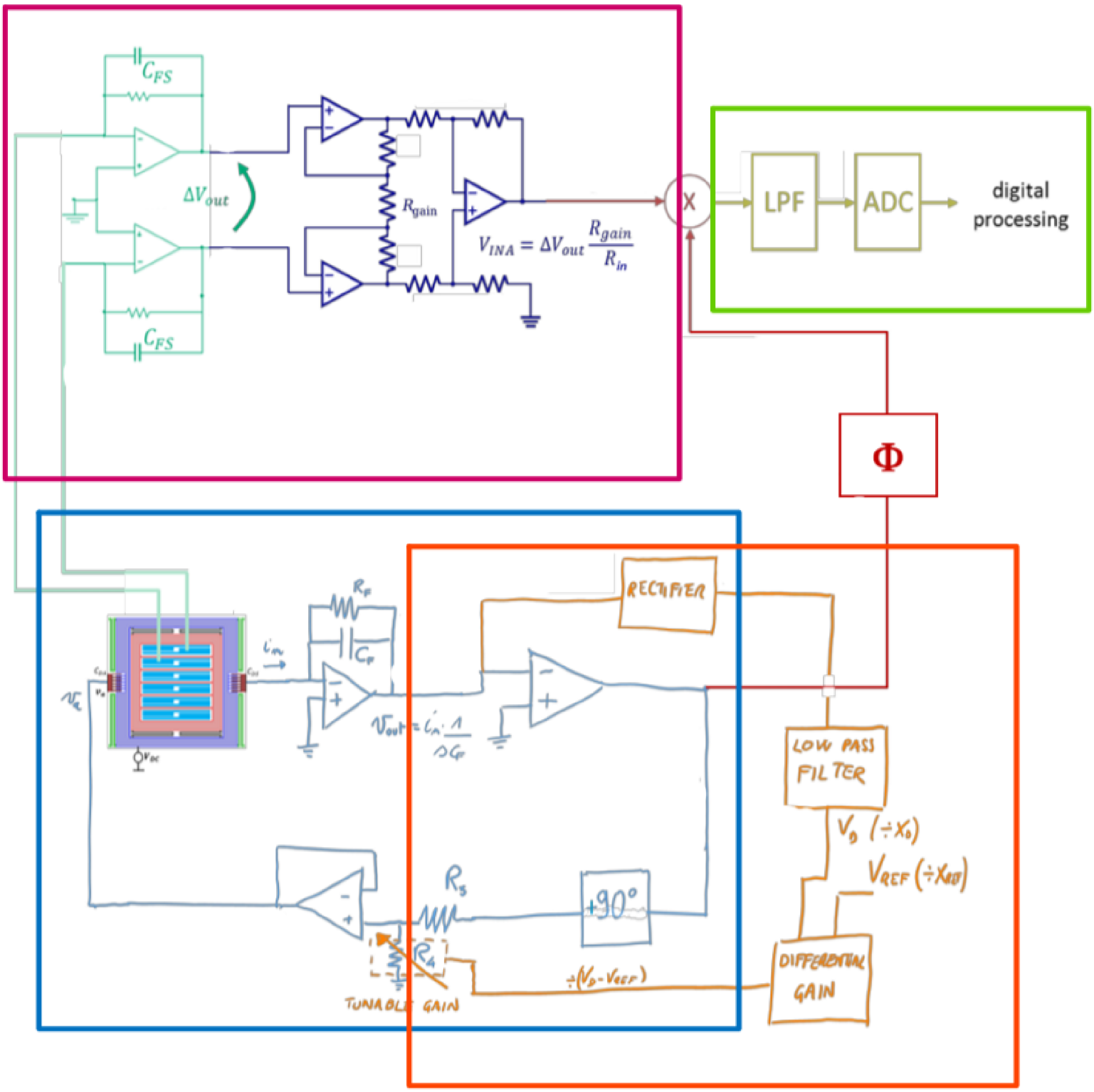


- **Synchronous** Demodulation
  It consists in multiplication by a square wave and filtering the signal. As you can see in the figure below, it takes into account the phase of the signal to demodulate (can distinguish the amplitude of different phase components of a signal at a given frequency, and saves the sign). Also note that quadrature disturbances are rejected (see par. 5.13).



Finally, by summarizing all the blocks we have described, we can draw the system in its entirety:

1. Primary loop (**oscillator**);
2. Secondary loop (**amplitude gain control**);
3. **Sense readout** and **demodulation chain**;
4. Analog to Digital Conversion (**ADC**).



5.5 Effect of accelerations and vibrations

So far, we neglected the fact that a linear acceleration (and not only an angular rate) can occur on the suspended mass. Let us check how problematic this can be:

- The effect of a consumer MEMS FSR acceleration (about 18 $\hat{g}$) on a gyro mass with a resonance e.g. at 5 kHz:

$$y_S = \frac{FSR}{\omega_0^2} = 176\ nm$$

- The effect of a Coriolis force (at the resolution limit, e.g. 100 mdps, about 2 mrad/s) for a device moving 5 µm and having a 200 Hz bandwidth:

$$y_{S,0} = \frac{5\mu m}{2\pi \cdot 200Hz} = 8\ pm$$

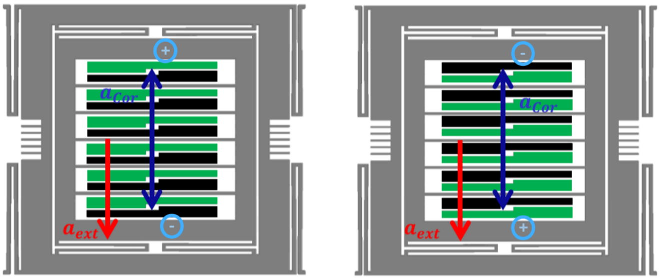
We conclude that accelerations are effectively disturbing. The example shows an effect of >20000 times the resolution to be measured and, even if the acceleration is not modulated at ω_0 and could be thus partially filtered, its effects are huge and it would be better to avoid such signals in the readout chain. We further note that the calculation done before assumes that the acceleration is not at resonance with the gyro modes. Otherwise, the unwanted acceleration signal would be even amplified by orders of magnitude! Thus, it is important that the frequencies of the gyro modes do not fall in a range where other disturbances (e.g. sounds or vibrations) are common! For this reason, gyro modes are usually designed to be:

- above 15-20 kHz for consumer applications (above the audio range);
- above 30-50 kHz for automotive, aerospace and military applications (above vibrations range);

Note that such increase of ω_0 gives also less sensitivity to accelerations. Often this is not enough (e.g. shocks) and more **advanced gyro architectures** should be designed, as described below.

5.6 Dual-mass Gyroscope

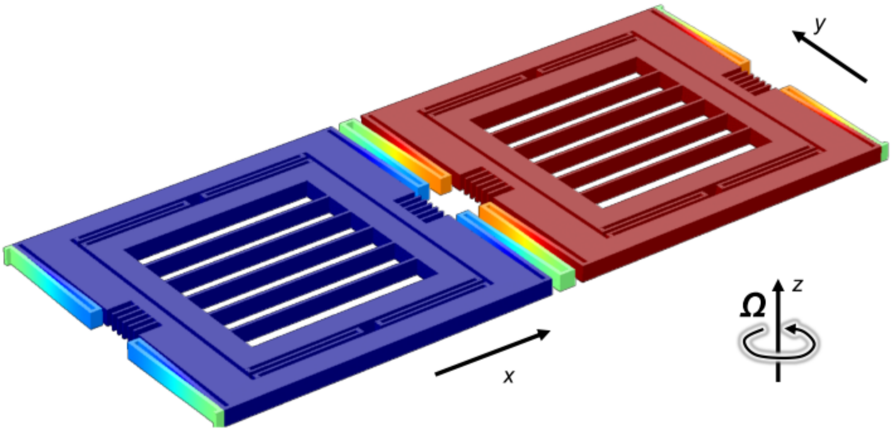
A basic idea could be to use two gyroscopes instead of one. The gyroscopes are **actuated** with the drive modes in opposite direction and, as a consequence, the Coriolis accelerations have opposite directions.



Whereas, as we can see from the picture, linear accelerations have the same direction on both sense masses and can thus be eliminated.

5.6.1 Tuning fork

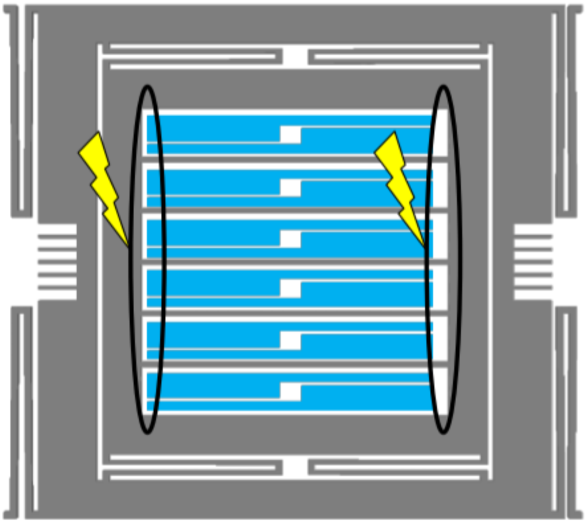
Rather than designing two separate devices, it is useful to couple them through a spring called **tuning fork** (TF). This ensures a single frequency for the anti-phase drive mode, and avoids the chance that two separate drive modes have different frequencies due to process non-uniformities: $\omega_{0,1} = \omega_{0,2}$. The Coriolis force and the sense mode are in anti-phase too, with a suitable arrangement of PP stators, accelerations will be rejected as a common mode signal.



The sensitivity calculation does not change because if on one side we have double the area occupied by the other side we have double the capacitance.

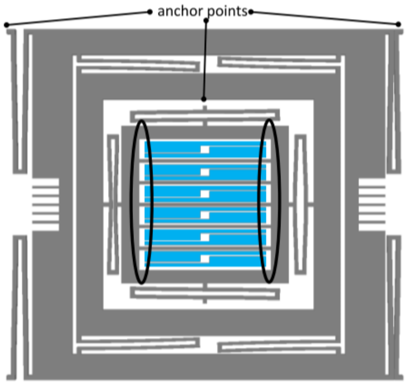
5.6.2 Doubly-decoupled gyroscope

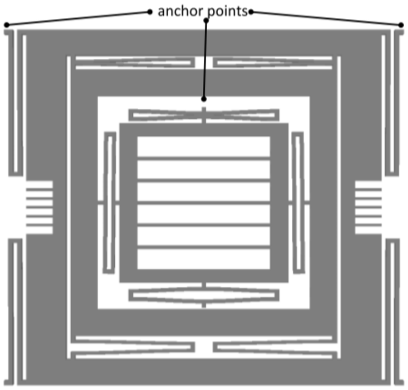
Another problem occurs because the sense mode, though moving parallel to the stators during drive motion, can see small (but non negligible) capacitance changes because of fringe effects. This suggest a further decoupling and originates an intermediate frame, called Coriolis (or decoupling) frame.



With this further decoupling, we have now three frames:

- **drive frame** → moving along X-direction
- **Coriolis frame** (decoupling frame) → dragged in the X-direction by the drive mode, pushed in the Y-direction by the Coriolis force;
- **sense frame** → not moving along the drive direction (note the new anchor point), pushed along the sense direction by the Coriolis frame.





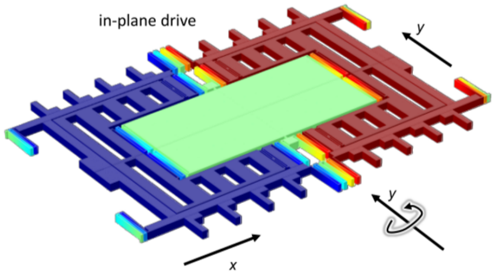
Obviously, also this solution can be implemented as a tuning-fork, dual-mass, configuration.

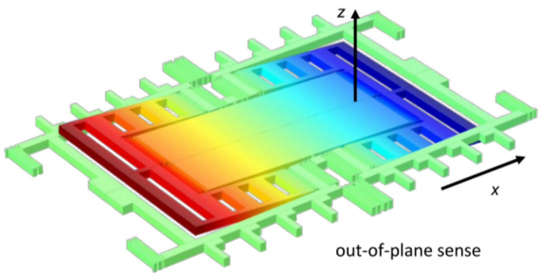
5.7 X and Y-axis gyroscopes

Y- (or X-) axis gyroscopes have usually a drive mode along the X(or Y-) axis, with the sense mode along the vertical axis. They usually feature the following characteristics:

- **rotation** rather than translation – as seen for accelerometers;
- **readout capacitive plates** designed beneath the structure;
- anti-phase, dual-mass, **tuning-fork drive mode**;
- **differential motion of the sense** mode to reject accelerations (like in Z-axis).

The other axis is obtained just by tilting this device by 90°.





5.8 Mode-matched gyroscope bandwidth

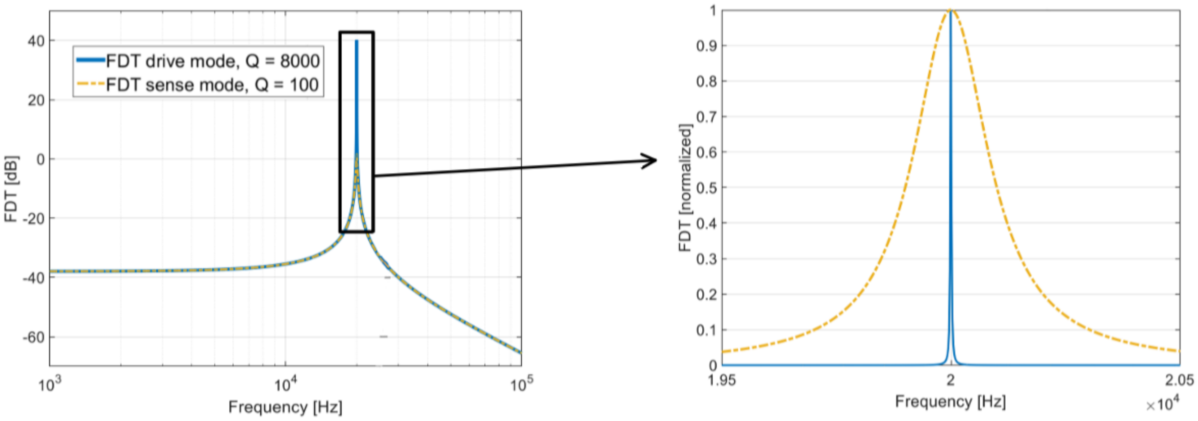
We saw that the sensitivity is proportional to the drive displacement amplitude, which is itself proportional to the applied AC voltage (actuator), the applied DC voltage (rotor) and the drive quality factor.

$$\frac{\Delta V_{\text{out },0}}{\Omega}=2\frac{V_{DC}}{C_{FS}}\frac{C_S}{g}\frac{x_{D,0}}{\Delta\omega_{BW}}=2\frac{V_{DC}}{C_{FS}}\frac{C_S}{g}\frac{x_{\text{ref}}}{\Delta\omega_{BW}}$$

So, if we want to **maximize the sensitivity** while keeping relatively low voltages (v_a should be within typical IC power supplies, and $\ll V_{DC}$), it is useful to exploit high drive Q factors to reach large displacements $x_{D,0}$. [20] This is why it is preferable to use comb-finger actuation and parallel plate sensing. On the other side, the sense mode Q factor does not explicitly appear in the sensitivity expression, it is "hidden" in the expression of the mechanical bandwidth: $\Delta\omega_{BW}\propto b_S\propto 1/Q_S$

$$\frac{\Delta V_{\text{out },0}}{\Omega}=2\frac{V_{DC}}{C_{FS}}\frac{C_S}{g}\frac{x_{D,0}}{\Delta\omega_{BW}}$$

The choice of a PP readout was so far motivated by the small displacement approximation.
We now report sample transfer functions of the two modes of a **mode matched** (i.e. with $\omega_{0,D}=\omega_{0,S}=\omega_0$) gyro, with Q factors in the order of several 1000s (drive) and few 100s (sense).



Note that:

- The system response is given by the sense mode transfer function (yellow), excited by the Coriolis force (blue) at a frequency corresponding to the drive mode frequency ω_0 times the frequency of the angular rate ω_Ω:

$$F_{\text{cor },0}\propto v_{D,0}\left(\omega_0\right)\cdot\Omega\left(\omega_\Omega\right)$$

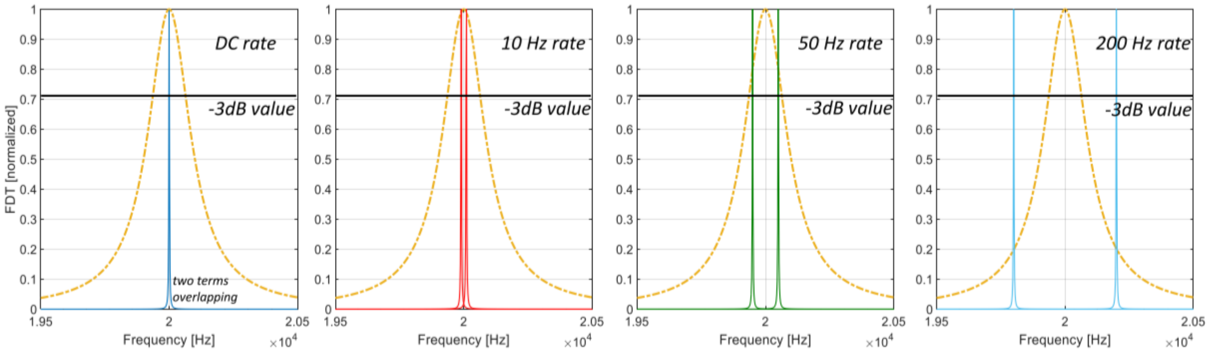
- The situation shown above is valid for perfect mode matching ($\omega_S=\omega_D=\omega_0$) and for a DC angular rate only ($\omega_\Omega=0$).

---

[20] Remember that for large displacement the parallel plate detection would be increasingly non-linear

The **sensing bandwidth** of a gyroscope represents the maximum frequency of the angular rate that a gyro can measure (-3 dB loss). Assume that we have an AC angular rate (e.g. cosinusoidal), the frequency of Coriolis force components is given by the sum/difference of the Coriolis force AC frequency ω_Ω and of the drive mode at ω_0.

$$\begin{aligned} F_{\text{Cor}} &= 2m_s v_{D,0}\cos(\omega_0 t)\,\Omega(t) = 2m_s v_{D,0}\cos(\omega_0 t)\,\Omega_0\cos(\omega_\Omega t) \\ F_{\text{Cor}} &= 2m_s v_{D,0}\frac{1}{2}\left[\cos\left([\omega_0+\omega_\Omega]\,t\right)+\cos\left([\omega_0-\omega_\Omega]\,t\right)\right]\cdot\Omega_0 \end{aligned}$$



There is a relationship between the -3dB value of a resonant transfer function and its quality factor, which can be found considering the displacement transfer function:

$$\left|\frac{y_S(j\omega)}{F_{\text{Cor}}(j\omega)}\right| = \left|\frac{1}{m_S}\frac{1}{(j\omega)^2+\frac{\omega_0}{Q_S}j\omega+\omega_0^2}\right| = \frac{1}{m_s}\frac{1}{\sqrt{\left(\omega_0^2-\omega^2\right)^2+\left(\frac{\omega\omega_0}{Q_s}\right)^2}}$$

At resonance:

$$\left|\frac{y_s(j\omega_0)}{F_{Cor}(j\omega_0)}\right| = \frac{1}{m_s}\frac{1}{\frac{\omega_0^2}{Q_s}} = \frac{Q_s}{k_s}$$

If we assume $\omega_\Omega = \Delta\omega_{BW} = \omega_0/2Q$ and calculate[21] the transfer function at $\omega_0 \pm \Delta\omega_{BW}$

$$\begin{aligned} \left|\frac{y_s(j\omega_0(1\pm 1/2Q))}{F_{\text{Cor}}(j\omega_0(1\pm 1/2Q))}\right| &= \frac{1}{m_s}\frac{1}{\sqrt{\left(\cancel{\omega_0^2}-\cancel{\omega_0^2}-\cancel{\frac{\omega_0^2}{4Q_s^2}}\pm\frac{\omega_0^2}{Q_s}\right)^2+\left(\frac{\omega_0^2}{Q_s}\pm\cancel{\frac{\omega_0^2}{2Q_s^2}}\right)^2}} = \\ &= \frac{1}{m_s^2}\frac{1}{\sqrt{2\left(\frac{\omega_0^2}{Q_s}\right)^2}} = \frac{1}{\sqrt{2}}\frac{Q_s}{k_s} \end{aligned}$$

As we can see, the assumed $\Delta\omega_{BW}$ is the value for which the sense transfer function looses 3 dB.

---

[21] In the computation we consider: $Q_s \gg 1 \quad \rightarrow \quad \frac{\omega_0^2}{Q_s^2} \ll \frac{\omega_0^2}{Q_s}$

As usually done in the field of electronics, the sensing bandwidth is given by the AC frequency for which the response value drops by 3 dB (0.5 in power scale, 0.707 in linear scale). For a second order system with complex conjugate poles, the -3dB value is thus calculated through the formulas that we obtained above.

$$\Delta\omega_{BW} = \frac{\omega_0}{2Q_S} = \frac{\omega_0 b_S}{2\omega_0 m_S} = \frac{b_S}{2m_S}$$

$$\Delta f_{BW} = \frac{f_0}{2Q} \qquad Q = \frac{f_0}{2\Delta f_{BW}}$$

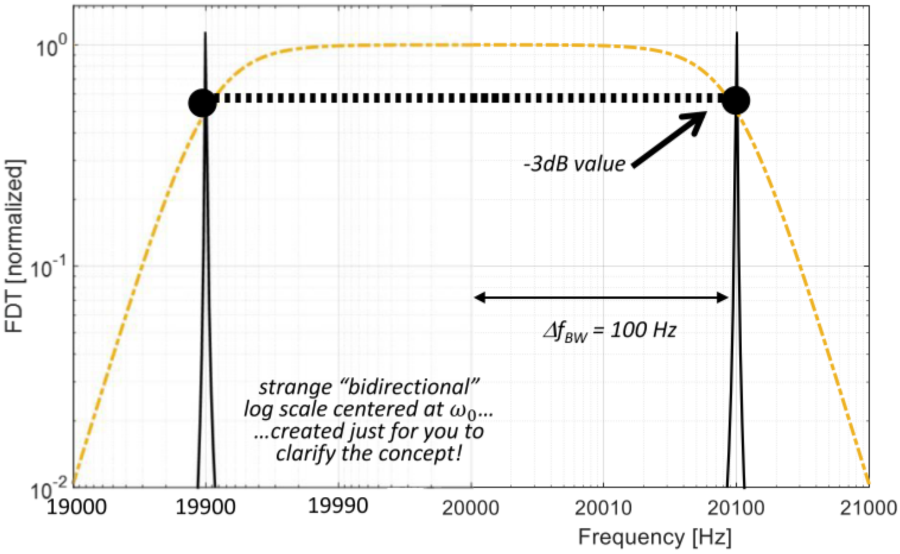


Figure 44: In this example: $\Delta f_{BW} = \frac{20\,kHz}{2\cdot 100} = 100\ Hz$

Therefore, gyros that operate at resonance show a clear sensitivity-bandwidth trade-off:

$$\frac{\Delta V_{\text{out },0}}{\Omega} = 2\frac{V_{DC}}{C_{FS}}\frac{C_S}{g}\frac{x_{D,0}}{\Delta\omega_{BW}} \qquad \frac{\Delta V_{\text{out, }0}}{\Omega}\Delta\omega_{BW} = \text{ constant} \qquad \Delta\omega_{BW} = \frac{b_S}{2m_S}$$

Or, in other words we can say that the gain-bandwidth product is constant. So, we have seen why a too small damping coefficient b_s in the sense mode (too large Q_S) is not suitable, as it would determine high sensitivity but a too low bandwidth.
For example, if we assume the same Q factor as for the drive mode, we get a bandwidth which is not compatible with almost any applications:

$$\Delta f_{BW} = \frac{f_0}{2\,Q} = \frac{20000\ Hz}{2\cdot 8000} = 1.25\ Hz$$

5.9 Mode-matched gyroscope noise

5.9.1 Thermo-mechanical noise of the gyro frames

Let us now calculate the thermo-mechanical noise for a gyroscope. We have two decoupled modes, so we should consider the thermo-mechanical noise contribution from each of them.

- **Drive Mode Noise**
  The Brownian noise in terms of force on the drive frame:

  $$S_{Fn} = 4\ k_B\ T\ b_D \qquad [\frac{N^2}{Hz}]$$

  This noise is transferred into a drive displacement noise through the drive mode transfer function, thus becoming a "sensitivity noise". Assuming, for sake of simplicity, that this contribution is constant over the drive mode peak width $\Delta f_{BW,D}$, we obtain:[22]

  $$S_{\mathrm{X}n} = 4k_BTb_D \cdot \left(\frac{Q_D}{k_D}\right)^2 \qquad \left[\frac{m^2}{Hz}\right]$$

  $$\sigma_{x,n} = \sqrt{S_{\mathrm{X}n} \cdot \Delta f_{BW,D}} = \sqrt{4k_BTb_D \left(\frac{Q_D}{k_D}\right)^2 \frac{f_0}{2Q_D}} = \sqrt{4k_BT\ b_D \frac{Q_D}{k_D^2}\frac{\omega_0}{4\pi}} =$$

  $$= \sqrt{k_BT\frac{k_D}{\omega_0 Q_D}\frac{Q_D}{k_D^2}\frac{\omega_0}{\pi}} = \sqrt{\frac{k_BT}{\pi}\frac{1}{k_D}}$$

  If we do a numerical example, we can notice that the value has a **negligible** (for $k_D = 50N/m \longrightarrow 5pm$) effect on the sensitivity (about 1 part per million).

- **Sense Mode Noise**
  We do the same calculation:

  $$S_{Fn} = 4k_BTb_S \quad \left[\frac{N^2}{Hz}\right] \quad \xrightarrow{|T(j\omega_0|^2} \quad S_{Yn} = 4k_BTb_S \cdot \left(\frac{Q_S}{k_S}\right)^2 \quad \left[\frac{m^2}{Hz}\right]$$

  $$\sigma_{y,n} = \sqrt{S_{Yn} \cdot \Delta f_{BW}} = \sqrt{4k_BTb_S \left(\frac{Q_S}{k_S}\right)^2 \frac{f_0}{2Q_S}} = \sqrt{4k_BT\frac{\omega_0 m_S}{Q_S}\left(\frac{Q_S}{k_S}\right)^2 \frac{f_0}{2Q_S}} =$$

  $$= \sqrt{4k_BT\frac{\omega_0 m_S}{k_S^2}\frac{\omega_0}{4\pi}} = \sqrt{\frac{k_BT}{\pi}\frac{1}{k_S}}$$

  As the sense frame is usually smaller (lower mass, consider only m_S) than the driven frames, its stiffness is usually a bit smaller too, to get the same resonance. If we do a numerical example for typical $k_S = 30N/m \longrightarrow 7pm$ (**not negligible** for typical sense mode displacements).

---

[22] This approximations is valid only for small bandwidths, since the transfer function is actually multiplied by Q only on the peak.

We can calculate the equivalent angular rate noise using the sensitivity:

$$y_S = \frac{x_{D,0}}{\Delta\omega_{BW}}\Omega \quad \rightarrow \quad \Omega = \frac{y_S}{x_{D,0}}\Delta\omega_{BW} \quad \rightarrow \quad \sqrt{S_{\Omega n}} = \sqrt{S_{Yn}}\frac{\Delta\omega_{BW}}{x_{D,0}}$$

Therefore, we can write the expression of the Brownian-limited **N**oise **E**quivalent **R**ate **D**ensity, similarly to what we did for accelerometers (with NEAD in section 3.5):

$$NERD = \sqrt{S_{\Omega n}} = \frac{\sqrt{4k_BTb_S\frac{Q_S^2}{k_S^2}}}{x_{D,0}/\Delta\omega_{BW}} = \frac{\sqrt{4k_BTb_S\frac{Q_S^2}{k_S^2}\Delta\omega_{BW}^2}}{x_{D,0}} = \frac{\sqrt{4k_BTb_S\frac{Q_S^2}{k_S^2}\frac{\omega_0^2}{4Q_S^2}}}{x_{D,0}} =$$

$$= \frac{\sqrt{k_BTb_S\frac{\omega_0^2}{k_S^2}}}{x_{D,0}} = \frac{1}{x_{D,0}}\sqrt{k_BTb_S\frac{1}{\omega_0^2m_S^2}} = \frac{1}{x_{D,0}\omega_0m_S}\sqrt{k_BTb_S}$$

$$NERD = \frac{180}{\pi}\frac{1}{x_{D,0}\omega_0m_S}\sqrt{k_BTb_S} \qquad \left[\frac{dps}{\sqrt{Hz}}\right]$$

5.9.2 Electronic noise contributions

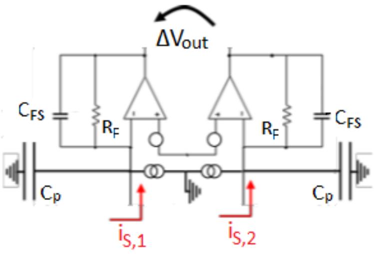
Also for the electronic noise, we should consider in principle both the sense and drive systems.

- **Sense System**
  Noise contributions are brought to the output in terms of voltage, and then scaled through the sensitivity $\left(\frac{\Delta V_{out}}{\Omega} = 2\,V_{DC}\frac{C_s}{C_{FS}}\frac{1}{g}\frac{x_{D,0}}{\Delta\omega_{BW}}\right)$ into equivalent input rate:
  - **feedback resistors:** current noise → output voltage → equivalent rate (in $dps/\sqrt{Hz}$):

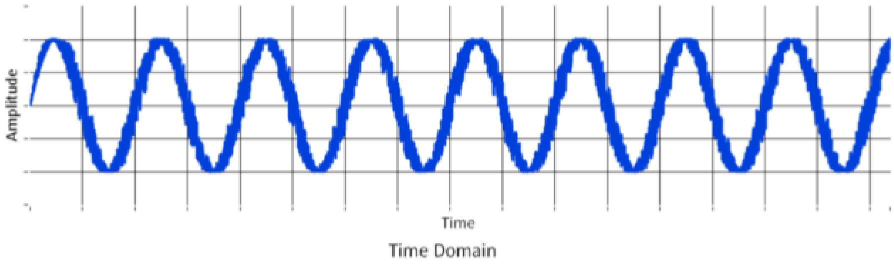
$$\sqrt{S_{\Omega n,R_F}} = \frac{\sqrt{2\cdot\frac{4k_bT}{R_F}\left(\frac{1}{\omega_0C_{FS}}\right)^2}}{\frac{\Delta V_{out}}{\Omega}}\frac{180}{\pi} = \sqrt{\frac{4k_bT}{2R_F}}\frac{g}{C_SV_{DC}\omega_0}\frac{\Delta\omega_{BW}}{x_{D,0}}\frac{180}{\pi}$$

  - **operational amplifier noise:** output voltage noise → equivalent rate (in $dps/\sqrt{Hz}$):

$$\sqrt{S_{\Omega n,CA}} = \frac{\sqrt{2\cdot S_{n,op}}\left(1+\frac{C_P}{C_{FS}}\right)}{\frac{\Delta V_{\text{out}}}{\Omega}}\frac{180}{\pi} = \sqrt{\frac{S_{n,op}}{2}}\left(1+\frac{C_P}{C_{FS}}\right)\frac{C_{FS}}{C_S}\frac{g}{V_{DC}}\frac{\Delta\omega_{BW}}{x_D}\frac{180}{\pi}$$



- **Drive Oscillator electronics**
  There is as well also a noise contribution from the drive oscillator electronics. In general, noise in oscillator is described by the **phase noise theory**, where noise in harmonic signals is seen as a combination (50% each) of amplitude noise and phase noise.



The Basic model for noisy harmonic signals with offset is:

$$u(t) = A_{\text{err}} + A[1+\alpha(t)]\cos\left(\omega_0 \cdot t + \varphi_{\text{err}} + \varphi_{\text{noise}}\right)$$

Where:

- A: average AC amplitude;
- $\alpha(t)$: fractional amplitude fluctuations (amplitude noise);
- A_{err}: offset;
- $\omega_0 \cdot t$: instantaneous phase (average frequency ω_0);
- ϕ_{noise}: phase fluctuations (phase noise);
- ϕ_{err}: phase offset.

In an oscillator, if amplitude noise and offset are cancelled by saturation of the square wave, the model (for the 1st harmonic) simplifies to:

$$u(t) = A\ cos(\omega_0 t + \phi_{err} + \phi_{noise})$$

Some notes: the effects of drive noise on the output white noise are not negligible and are even hard to predict as they depend, as we will see in par. 5.13, on the so-called **quadrature error**. The effects of phase offset drift (coupled to quadrature error) are a critical point for some applications like indoor inertial navigation;

Finally combining all noise contributions we conclude that to improve the gyro noise performance:

$$\sqrt{S_{\Omega n,tot}} \approx \frac{180}{\pi}\frac{1}{x_{D,0}}\sqrt{\left(\sqrt{2\cdot\frac{4k_BT}{R_F}}\frac{g\,\Delta\omega_{BW}}{2C_SV_{DC}\omega_0}\right)^2+\left(\sqrt{\frac{S_{n,op}}{2}}\frac{C_P}{C_S}\frac{g\,\Delta\omega_{BW}}{V_{DC}}\right)^2+\left(\frac{1}{\omega_0m_S}\sqrt{k_BTb_S}\right)^2}$$

As we can see:

- it is **undoubtedly useful** to increase $x_{D,0}$ (and, if possible, ω_0);
- if the **electronics noise** (first two addends) **dominates** (quite common situation):
  - try increasing the feedback resistance value (use off-MOS in integrated implementations);
  - try to lower the parasitic capacitance (smart routing of interconnections and pads etc...);
  - use the minimum gap value (this will however increase b_S because of squeeze film damping effect);
  - increase V_{DC} as much as you can (not for free: you pay in power dissipation).
- if the **thermo-mechanical noise dominates**:
  - increase m_s (not for free: either thicken your process – very good option – or pay in area);
  - decrease b_S (not for free: you lose your maximum sensing bandwidth).

5.10 Issues in mode-match gyroscopes

A first issue in mode-matched operation is the relatively limited bandwidth, its relationship with the damping coefficient and thus with the thermo-mechanical noise:

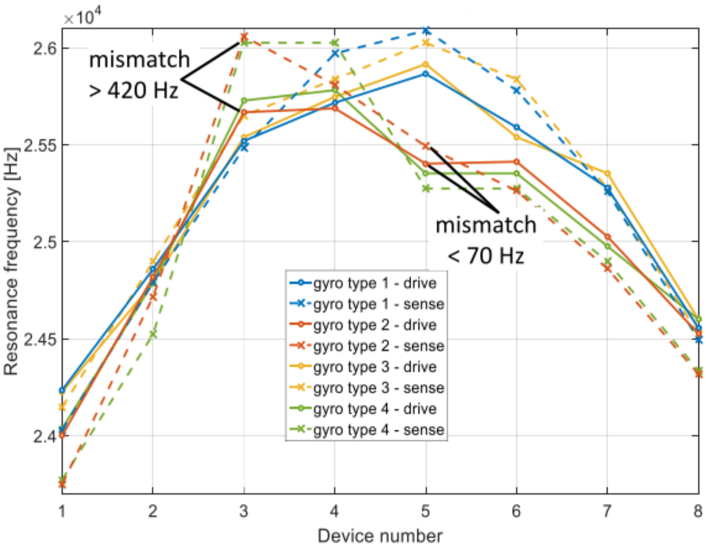
$$\Delta f_{BW} = \frac{f_S}{2Q_S} = \frac{b_S}{4\pi m_S}$$

$$NERD = \frac{1}{x_d\omega_dm_S}\sqrt{k_bTb_s}$$

There is thus a **marked trade-off** between best achievable noise density (assuming a negligible electronic noise) and maximum sensing bandwidth. This trade-off passes through the damping coefficient.

So far we assumed that the drive and sense resonance frequencies are perfectly matched. In practice, even with very good design and process, the frequencies will not be matched! E.g. due to process spread, typical f_s values can range within $f_d \pm 600\ Hz$ $(\pm 3\sigma)$, around a frequency of $20\ kHz$.

Modes distribution on a Si wafer. Non-uniformities are mostly due to local differences in DRIE of springs (cubic spring dependence on w). Therefore, non-uniformities affect similarly the two modes of the same gyro, as their springs lie close one another. However, a residual (f_s-f_d) difference still remains, varying between few 10s Hz to few 100s Hz.



We learned how frequencies can be tuned via electrostatic force ("*tuning plates*"). However, even using tuning plates on the sense mode to match the frequency at a given T, there is an issue associated with the temperature behavior of the **Young's modulus** in Silicon:

- the mechanical stiffness k goes with the Young's modulus E:

$$k = \beta\, E \qquad\qquad \frac{dk}{k} \propto \frac{dE}{E}$$

- the resonance frequency goes with the square root of the stiffness:

$$f_0 = \frac{1}{2\pi}\sqrt{\frac{k}{m}} \qquad\qquad \frac{df_0}{f_0} = \frac{1}{2}\frac{dk}{k}$$

- the temperature coefficient of frequency (**TCf**) in Silicon is therefore:[23]

$$TCf = \frac{\frac{df_0}{f_0}}{\Delta T} = \frac{1}{2} \cdot \frac{\frac{dE}{E}}{\Delta T} = -30\; ppm/K$$

Moreover, even if we use tuning to match the frequency at one T ($f_S - f_D = 0$), there will be a **variation of the split** of the frequencies, larger than $\pm 1\; Hz$ ($|f_S - f_D| > 0$), if we consider operation on the whole ΔT range. Indeed, the electrostatic stiffness that we use to tune does not follow the above law with respect to T (k_{elec} is basically independent of T changes).

---

[23] ppm: parts per million

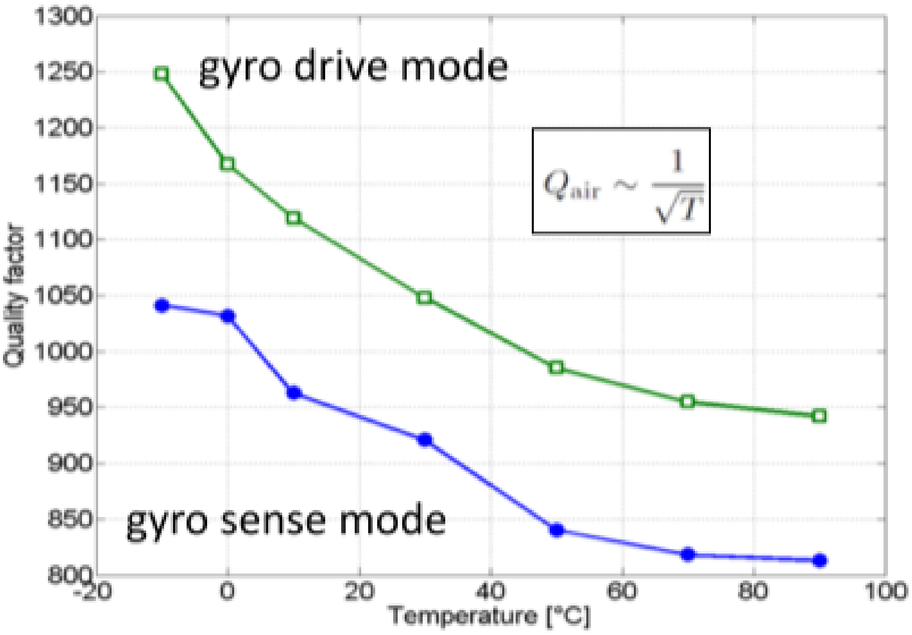
Example

Let's assume $f_D = 20\ kHz$ and $f_S = 20.6\ kHz$ and consider the max range to be guaranteed in automotive applications: -45°C to 125°C and so $\Delta T = 170K = \pm 85K$

$$df_D = -30\frac{ppm}{K} \cdot f_D \cdot \pm\Delta T = \pm 51 \text{ Hz}$$
$$df_S = -30\frac{ppm}{K} \cdot f_S \cdot \pm\Delta T = \pm 52.53 \text{ Hz}$$
$$d(\Delta f) = -30\frac{ppm}{K} \cdot \Delta f \cdot \pm\Delta T = \pm 1.53 \text{ Hz}$$

We already observed, when discussing the drive mode circuits, that the **drive Q** changes with T (due to b_D changes), and so does the sensitivity, which goes linearly with the drive displacement. Similiarly, the sense quality factor Q_S is affected by absolute temperature changes, following the same law ($\propto 1/\sqrt{T}$). The variations of b_S determine simultaneous changes in the maximum sensing bandwidth value and in the sensitivity!

$$\frac{y_{S,0}}{\Omega} = \frac{x_{D,0}}{\Delta\omega_{BW}(T)} = \frac{x_{D,0}}{b_S(T)/\left(2m_S\right)} = \frac{x_{D,0}}{\omega_0/\left(2Q_S(T)\right)} \propto Q_S(T)$$



Such percentage Q_S changes directly translates in bandwidth and sensitivity changes: bandwidth decreases at low temperatures and sensitivity decreases at high temperatures.
We do not want that just one device works in one particular operating condition. We want that all devices work within all (**guaranteed**) operating conditions!

5.11 Mode-split operation

Let us analyze the transfer function assuming that the drive mode is on-purpose designed **before** the sense mode ($\omega_D \ll \omega_S$). Since the x-axis displacement frequency is set by the drive mode resonance frequency, we calculate the modulus of the transfer function of the sense mode evaluated at the drive mode frequency:

$$T_{YF}(s) = \frac{Y(s)}{F_{Cor}(s)} = \frac{1/m_S}{\left(s^2 + \frac{b_S}{m_S}s + \frac{k_S}{m_S}\right)}$$

Let's consider $\Delta\omega_{MS} = \omega_S - \omega_D$, $\Delta\omega_{MS} \ll \omega_S, \omega_D$ and $\Delta\omega_{MS} \gg \Delta\omega_{BW}$ $(Q_S \gg 1)$:

$$|T_{YF}\left(j\omega_D\right)| = \left|\frac{1/m_S}{\left(-\omega_D^2 + j\frac{\omega_D\omega_S}{Q_S} + \omega_S^2\right)}\right| = \frac{1/m_S}{\sqrt{\left(\omega_S^2 - \omega_D^2\right)^2 + \left(\frac{\omega_D\omega_S}{Q_S}\right)^2}} =$$

$$= \frac{1/m_s}{\sqrt{\left(\omega_S^2 - \omega_S^2 + 2\omega_S\Delta\omega_{MS} - \Delta\omega_{MS}^2\right)^2 + \left(\frac{\omega_D\omega_S}{Q_S}\right)^2}} = \frac{1/m_S}{\sqrt{(2\omega_S\Delta\omega_{MS} - \cancel{\Delta\omega_{MS}^2})^2 + \left(\frac{\omega_S^2 - \cancel{\Delta\omega_{MS}\,\omega_S}}{Q_S}\right)^2}} =$$

$$\frac{1/m_S}{\sqrt{4\omega_S^2\Delta\omega_{MS}^2 + \frac{\omega_S^4}{Q_S^2}}} = \frac{1/m_S}{2\omega_S\sqrt{\Delta\omega_{MS}^2 + \cancel{\Delta\omega_{BW}^2}}} = \frac{1}{k_S}\frac{\omega_S}{2\Delta\omega_{MS}} = \frac{1}{k_S}Q_{eff}$$

So we could define the **effective Quality Factor** (**independent from T**):

$$Q_{eff} = \frac{\omega_S}{2\ \Delta\omega_{MS}} = \frac{f_S}{2\ \Delta f_{MS}}$$

Mode-split operation sacrifices sensitivity for bandwidth. GBPW is still constant. Let us see if there are advantages after this sacrifice:

- The drive mode is identical to mode-matched operation:

$$F_{elec} = \frac{\varepsilon_0 h N_{CF}}{g} 2V_{DC}v_a \sin\left(\omega_0 t\right) = F_{elec,0}\sin\left(\omega_D t\right)$$

$$x_{D,0} = F_{elec,0}\frac{Q_d}{k_d} \qquad\qquad v_{D,0} = F_{elec,0}\frac{Q_d}{k_d}\omega_D$$

- For the sense mode we will use the Q_{eff} instead of Q_S (**T independence!**)

$$F_{\text{Cor, }0} = 2m_S \cdot v_{D,0} \cdot \Omega$$

$$y_{S,0} = F_{\text{Cor, }0}\frac{Q_{eff}}{k_S} = 2m_S \cdot \dot{x} \cdot \Omega\frac{Q_{eff}}{k_S} = 2m_s x_{D,0}\omega_D\frac{Q_{eff}}{k_S}\Omega$$

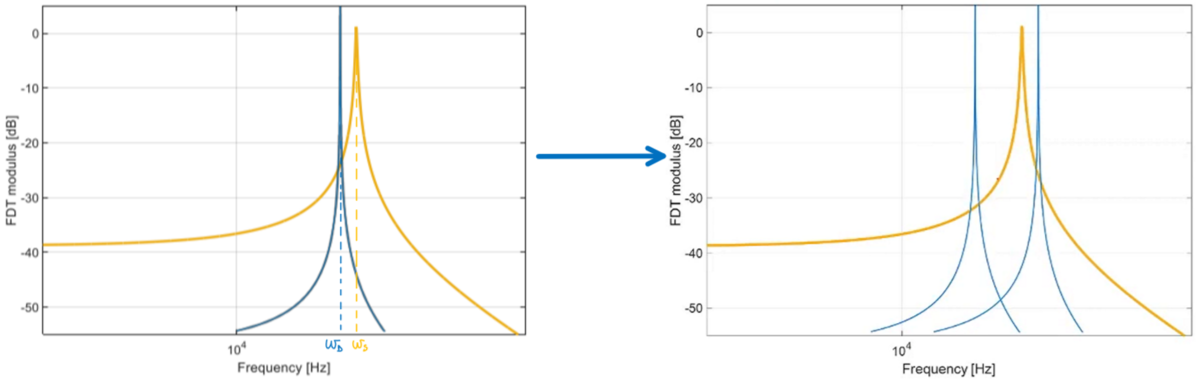
The sensitivity will therefore be:

$$\frac{y_{S,0}}{\Omega} = 2x_{D,0} \cdot \omega_D \cdot m_S\frac{Q_{eff}}{k_S} = x_{D,0}\frac{2Q_{eff}\omega_D}{\omega_S^2} =$$

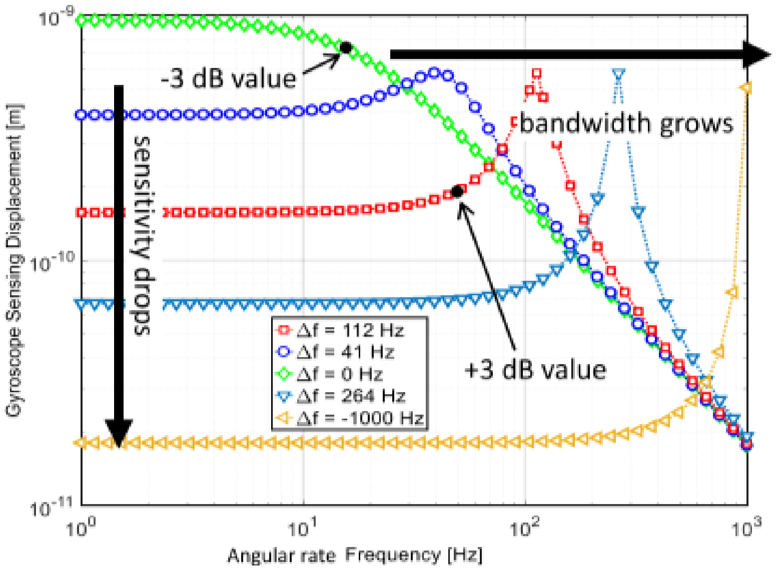
$$= x_{D,0}\frac{2\omega_D}{\omega_S^2}\frac{\omega_S}{2\Delta\omega_{MS}} = \frac{\omega_D}{\omega_S}\frac{x_{D,0}}{\Delta\omega_{MS}} \approx \frac{x_{D,0}}{\Delta\omega_{MS}}$$

5.11.1 Bandwidth in mode-split operation

We know that, because of the drive modulation, the AC components of the Coriolis force lie at $\omega_{D0} \pm \omega_\Omega$:



When ω_Ω increases, the sum of these values initially remains constant. At larger increases of ω_Ω, the component $\omega_{D0} + \omega_\Omega$ approaches the sense mode peak and generates an increase in the overall gyroscope sensitivity.
We are used to define -3dB bandwidth as would be the case for an overdamped axel or for a mode-matched gyro (by the way, shown by the light green curve). Instead, for such an underdamped overall response, we define the useful bandwidth as the ±3dB value (the smallest sets the band).



The overall system bandwidth will be given by the combination of the shown MEMS mechanical bandwidth, and an electronic filtering bandwidth in the sense chain that will try to flatten the peak and further extend the overall bandwidth.

5.12 Mode-Split gyroscope noise

We have already seen that the thermomechanical (**Brownian**) noise expression in terms of force power density (a function of the damping coefficient b of the considered mode) is:

$$S_{F,n} = 4\,k_B\,T\,b \qquad [N^2/Hz]$$

Let's analyze the two contributes:

- **Drive frame noise**
  Nothing changes with respect to mode-matched operation. We already observed that this noise contribution in terms of drive displacement (few pm_{rms}) is negligible when compared to typical drive amplitudes (few µm).

- **Sense frame noise**
  We have now a **different gain**, and in general a different operation mode. We thus need to check the validity of the expressions derived so far for the Brownian noise of the sense mode.
  The calculation of thermomechanical input-referred noise for the sense frame follows the same steps as for mode-matched operation, and comes to the same results.

MODE-MATCHED

$$\sqrt{S_{\Omega n}} = \sqrt{\frac{S_{Yn}}{\left(\frac{y_S}{\Omega}\right)^2}} = \sqrt{\frac{4k_BTb_S\left(\frac{Q_S}{k_S}\right)^2}{\left(\frac{x_{D,0}}{\Delta\omega_{BW}}\right)^2}}$$

$$\Delta\omega_{BW} = \frac{\omega_S}{2Q_S}$$

MODE-SPLIT

$$\sqrt{S_{\Omega n}} = \sqrt{\frac{S_{Yn}}{\left(\frac{y_S}{\Omega}\right)^2}} = \sqrt{\frac{4k_BTb_S\left(\frac{Q_{eff}}{k_S}\right)^2}{\left(\frac{x_{D,0}}{\Delta\omega_{MS}}\right)^2}}$$

$$\Delta\omega_{MS} = \frac{\omega_S}{2Q_{eff}}$$

$$\boxed{NERD = \frac{180}{\pi}\frac{1}{x_{D,0}\,m_S\,\omega_S}\sqrt{k_BT\,b_S}}$$

This is an expected result because Brownian noise is motion noise, therefore it is amplified as much as motion induced by Coriolis force. This time the amplification is by Q_{eff} instead of Q_S, but (apart from this) all the expressions remain identical. NERD equation does not change!
In the formulas above the white thermomechanical noise spectral density is amplified by the entire sense-mode transfer function (also at other frequencies). However, as later we will demodulate around the drive mode frequency and low-pass filter, we will effectively get relevant noise only from the region around the drive frequency.

One further, important remark is that the effective quality factor Q_{eff} is not a function of the damping coefficient. Therefore, the bandwidth (assumed as 1/4 to 1/2 of the mode split Δf_{MS}) is itself not related to the value of the damping coefficient!

$$\Delta\omega_{MS} = \omega_S - \omega_D = \frac{\omega_S}{2Q_{eff}} \qquad \Delta f_{MS} = f_S - f_D = \frac{f_S}{2Q_{eff}}$$

Therefore, it is possible to lower the thermomechanical noise (the NERD) by acting on b_s, without affecting the bandwidth, thus eliminating the trade-off that we had in mode-matched operation since we no longer necessarily need to keep a low value of the sense mode Q factor.

$$NERD = \frac{180}{\pi}\frac{1}{x_{D,0} m_S \omega_S}\sqrt{k_B\, T\, b_s}$$

Is there any disadvantage in mode-split operation?
The other noise contributions, from the electronics, are the same when calculated at the circuit output, but when calculated as equivalent input rate, they are now divided by a sensitivity which is lower by a factor $\Delta\omega_{BW}/\Delta\omega_{MS}$. We should design low-noise electronics, which requires high power dissipation.

$$\sqrt{S_{\Omega n,tot}} = \frac{180}{\pi}\frac{1}{x_{D,0}}\sqrt{\left(\sqrt{2\cdot\frac{4k_bT}{R_F}}\,\frac{g\,\Delta\omega_{MS}}{2C_S V_{DC}\omega_D}\right)^2 + \left(\sqrt{\frac{S_{n,op}}{2}}\,\frac{C_P}{C_S}\,\frac{g\,\Delta\omega_{MS}}{V_{DC}}\right)^2 + \left(\frac{1}{\omega_S m_S}\sqrt{k_B T b_S}\right)^2}$$

In conclusion, the 99% of the gyroscopes used in our cars or mobile phones actually **operate in mode-split** conditions, providing a positive impact on scale-factor stability. However, there are still challenges for offset stability, and this will be the topic of the following paragraphs.
Trials to track the sense frequency changes with temperature (so to keep mode-matching) were shown in the literature but they did not prove to be enough effective at low power consumption. Other techniques like feed-backs on the sense mode (similar to what we saw in accelerometers) are themselves too power hungry: in general, mode-matching can be used only for high-performance (low noise but high consumption), high-cost application fields.

5.13 Quadrature error issue

Let us start with a numerical example. Assuming a mode-split gyroscope for consumer application, we consider these typical values:

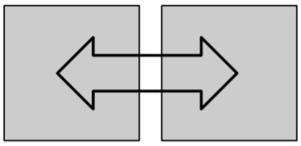
- The full-scale is about $2000°/s$;
- The resolution is about $100\ mdps$;
- Then bandwidth is about $1\ kHz$.

Assuming a drive mode displacement x controlled at $x_{REF} = 5\mu m$, we have a sense mode displacement y of about:

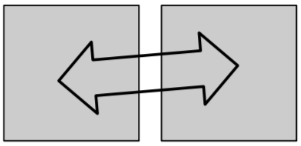
- $30\ nm$ at the full-scale range ($2000°/s$);
- $1.5\ pm$ at the resolution ($1.7 mrad/s$).

The drive mode is moving by 5 µm, i.e. about **3˙300˙000 times** more than the minimum sense mode motion to detect!!!

The gigantic displacement inequality can cause some issues if the drive motion is not perfectly orthogonal to the ideal sense mode direction. In this case, the sense mode would see a y-direction displacement at ω_D that is not caused by an angular rate.



top view: ideal in-plane drive motion only along the x-axis



top view: in-plane motion not only along the x-axis

This error, which is in a first approximation a DC term independent of Ω, may represent a **huge offset** and its compensation and/or drift represents a severe issue for gyros.
Note that this offset is in quadrature with respect to the Coriolis force, as it is proportional to x and not to $\dot{x}$. We model it with an equivalent quadrature force F_q:

$$|F_q| \propto x \qquad\qquad |F_{Cor}| \propto \dot{x}$$

Possible **sources of misalignment** in the drive mode are:

- **inhomogeneity** in the **stiffness** of the different springs because of local etching (DRIE) differences (in-plane) and because of the so-called skew-angle issue (in cross-section);
- **inhomogeneity** in the comb-drive **gaps** because of local etching (DRIE) differences (generally minor impact);
- bad design, design imperfections or masks misalignment;

Since the quadrature error results in an offset at the output, we can model it as an equivalent input-referred angular rate $\boldsymbol{B_q}$. We thus write the output voltage before demodulation as:

$$\Delta V_{\text{out}} = 2\frac{V_{DC}}{C_{FS}}\frac{C_S}{g}\frac{x_{D,0}}{\Delta\omega_{MS}}\left[\boldsymbol{\Omega}\cos\left(\omega_D t\right) + \boldsymbol{B_q}\sin\left(\omega_D t\right)\right] = S\left[\boldsymbol{\Omega}\cos\left(\omega_D t\right) + \boldsymbol{B_q}\sin\left(\omega_D t\right)\right]$$

An ideal demodulation (synchronous) consists in:

1. multiplying the signal above by $cos(\omega_D\ t)$;
2. filtering at the bandwidth BW.

Assuming Ω in DC and a LPF gain by a factor 2 that compensates the demodulation loss:

$$\begin{aligned} V_{\text{dem}} &= S\left[\Omega\cos\left(\omega_D t\right) + B_q\sin\left(\omega_D t\right)\right]\cdot\cos\left(\omega_D t\right) * LPF = \\ &= \frac{S}{2}\left\{\Omega\left[\cos(0) + \cos\left(2\omega_D t\right)\right] + \mathrm{B_q}\left[\sin(0) + \sin\left(2\omega_D t\right)\right]\right\} * LPF = \\ &= G_{LPF}\frac{S}{2}\left\{\Omega\cos(0) + \mathrm{B_q}\sin(0)\right\} = S\cdot\Omega = 2\frac{V_{DC}}{C_{FS}}\frac{C_S}{g}\frac{x_{D,0}}{\Delta\omega_{MS}}\cdot\Omega \end{aligned}$$

Quadrature error is **completely erased**. In this ideal condition, the care that one should take in the design is just the need for an extended electronic full-scale (for the supply before the demodulation stage) defined by:

$$V_{DD,min} = S\,\sqrt{\Omega^2 + B_q^2}\,G_{INA}$$

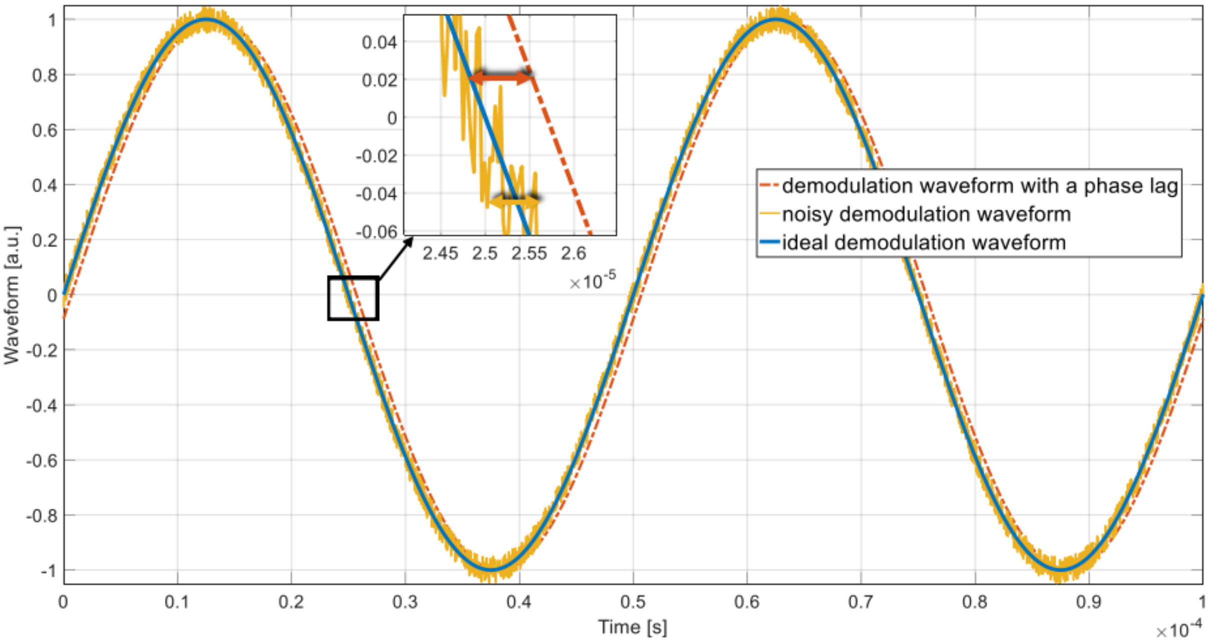
Unfortunately, in real systems the demodulation will unavoidably have a **phase error** and a **phase noise** (both small), as modelled previously (section 5.9.2). Therefore, the sense output signal:

$$\Delta V_{\text{out}} = 2\frac{V_{DC}}{C_{FS}}\frac{C_S}{g}\frac{x_{D,0}}{\Delta\omega_{MS}}\left[\boldsymbol{\Omega}\cos(\omega_D t) + \boldsymbol{B_q}\sin(\omega_D t)\right]$$

will not be multiplied simply by $cos(\omega_D\ t)$, but by $cos(\omega_D\ t + \phi_{err} + \phi_n)$.
The signal after demodulation and filtering is therefore:

$$\begin{aligned} V_{\text{dem}} &= S\left[\Omega\cos(\omega_D t) + B_q\sin(\omega_D t)\right]\cdot cos(\omega_D\ t + \phi_{err} + \phi_n) * LPF \\ &\approx S\left[\Omega\cos(\varphi_{\text{err}} + \varphi_n) + B_q\sin(\varphi_{\text{err}} + \varphi_n)\right] \\ &\approx S\left[\Omega + B_q\cdot\varphi_{\text{err}} + B_q\cdot\varphi_n\right] \\ &\approx \text{S}\cdot\Omega + \text{S}\cdot B_q\cdot\varphi_{\text{err}} + \text{S}\cdot B_q\cdot\varphi_n \end{aligned}$$

The figure below is an example of ideal and real demodulation waveform:



Input referred angular rate $\boldsymbol{B_q}$. To estimate the magnitude of $\boldsymbol{B_q}$, we add to the model a further cross-axis term due to the cross-axis stiffness. This term represents the quadrature force F_q in the y direction due to a displacement occurring in the x direction: $|F_q| = k_{ds}\, x$.
The motion equation for the gyroscope coupled modes becomes:

$$\begin{cases} m_s\ddot{\boldsymbol{y}} + b_s\dot{\boldsymbol{y}} + k_s\boldsymbol{y} - k_{ds}\boldsymbol{x} = -2m_s\Omega\dot{\boldsymbol{x}} \\ (m_D + m_S)\,\ddot{\boldsymbol{x}} + b_D\dot{\boldsymbol{x}} + k_d\boldsymbol{x} - \cancel{k_{ds}\boldsymbol{y}} = F_{elec} - \cancel{2\,(m_s + m_D)\,\Omega\dot{\boldsymbol{y}}} \end{cases} \qquad \begin{matrix} y \ll x \\ \dot{y} \ll \dot{x} \\ k_{ds} \ll k_d \end{matrix}$$

$$s^2 m_s Y(s) + s b_s Y(s) + k_s Y(s) = -2m_s\Omega s X(s) + k_{ds} X(s) = -\left(2m_s\Omega s - k_{ds}\right) X(s)$$

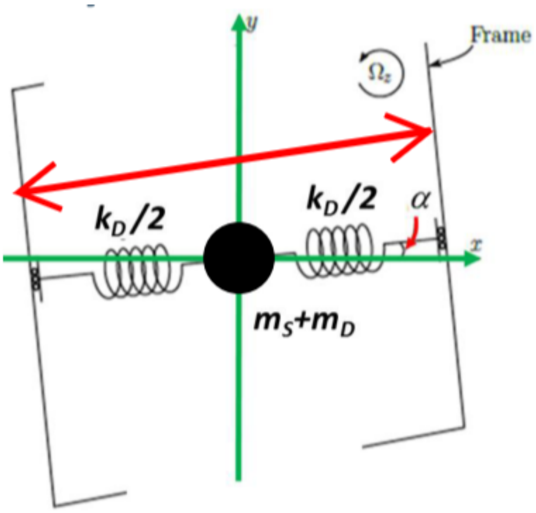
For the sense mode, the **offset term** can be written in terms of equivalent input-referred angular rate B_q , phase shifted by 90°(so by j in terms of Laplace Transform). $(s = j\omega_D \xrightarrow{90° \to \cdot j} -\omega_D)$

$$k_{ds} = 2m_s B_q \omega_D \qquad \Longrightarrow \qquad B_q = \frac{k_{ds}}{2m_S\omega_D}$$

So we could rewrite the expression:

$$s^2 m_s Y(s) + s b_s Y(s) + k_s Y(s) = -2m_s X(s)\,\omega_D\,(j\Omega + B_q)$$

Let us now determine k_{ds} so that we can determine B_q as well. We have assumed that the movement is not perfectly parallel to the x-axis, as showed below:



The y-direction displacement will be:

$$\frac{y}{x} = tan(\alpha) \cdot Q_{eff} \approx \alpha \cdot Q_{eff}$$

The corresponding force in the y-direction is related to y through k_{ds}:

$$y = \frac{F_q}{k_s} \cdot Q_{eff} = \frac{k_{ds}x}{k_s} \cdot Q_{eff} \quad \rightarrow \quad \frac{y}{x} = \frac{k_{ds}}{k_s} \cdot Q_{eff} \quad \rightarrow \quad k_{ds} = \alpha k_s$$

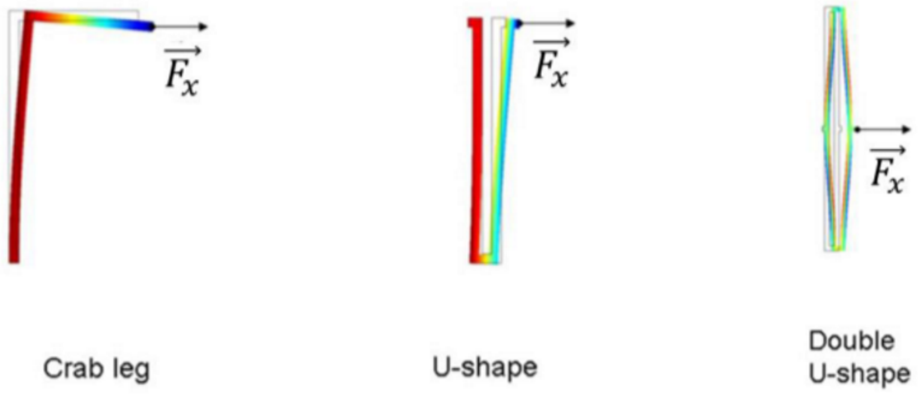
The model is self-consistent: with $k_{ds} = 0$ we have no displacement along x and we use the usual equations. We can thus find an expression for B_q as a function of the non-ideality α:

$$B_q = \frac{k_{ds}}{2m_S\omega_D} \approx \frac{\alpha k_S}{2m_S\omega_D} = \frac{\alpha\omega_S^2}{2\omega_D} \approx \frac{\alpha}{2}\omega_S \approx \frac{\alpha}{2}\omega_D$$

Now we can also understand why it is safe to keep the gyroscopes modes above the audio bandwidth (i.e. at about 20 kHz), but not higher than that!

5.13.1 Asymmetric stiffness effect in Z-axis gyroscopes

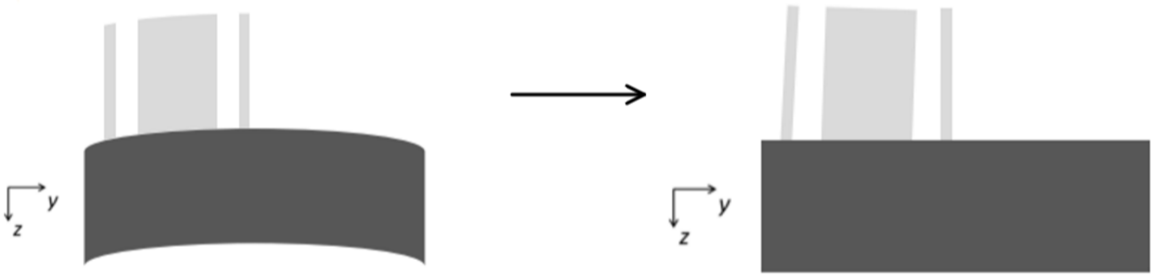
One critical parameter for quadrature in gyroscopes is the width of spring folds (and its non-uniformities) as the total in-plane stiffness is proportional to the cubic width (as seen in paragraph 3.8). There are springs that ideally do not deflect in the y direction under x-direction forces.



Usually, we have 4 symmetric springs per frame. Non-uniformities in springs width give rise to non-null k_{ds}, but with very low angle α (e.g. 10^{-3} ° to 10^{-2} °).

5.13.2 Skew-angle effect in X- and Y-axis gyroscope

This effect occurs in gyros for in-plane rate detection, where usually the sense frame has a low out-of-plane stiffness. Process non-uniformities make the **etching** non orthogonal to the substrate at wafer edges. This causes spring cross-sections as shown below.



An attempt to bend the beam in the x-direction results in a force with a z-axis component. The corresponding quadrature formula is similar to Z-axis devices, yet the value of a is typically much larger (e.g. 0.05° to 0.1°)

$$B_q = \frac{k_{ds}}{2\, m_S\, \omega_D} \approx \frac{\alpha}{2}\omega_S$$

This is the **largest source of quadrature in MEMS gyroscopes**.

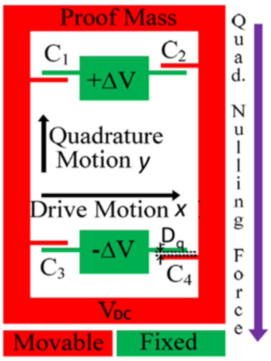
5.14 Quadrature error minimization

Let's see how we can minimize the quadrature error, described by the following formula:

$$B_q = \frac{k_{ds}}{2\, m_S\, \omega_D} = \frac{\alpha}{2}\omega_S$$

- **Minimize $\boldsymbol{k_{ds}}$** (so the angle α) by:
  - improving the process uniformity across the wafer;
  - increasing the beam width w (a larger w gives a lower relative weight of process non-uniformities dw, as already verified);
  - choosing springs with the lowest cross-axis term (folded or double-U);
  - choosing only gyroscope from wafer center.[24]
- **Decrease** the gyroscope **resonance**, but this is limited by acoustic disturbances (occurring at up to $>$ 20 kHz) and presence of vibrations (up to 1 kHz in consumer, 10 kHz in automotive, 50kHz in military applications). The frequency should be safely above!
- **Increase** the **inertial mass** (at constant frequency), but this is limited by the maximum area.

The best would be to compensate the quadrature error at its origin by acting on the device dynamics. This can be done, as usual, via electromechanical forces. We look for a force along the y-direction, proportional to x_D, and with selectable amplitude and sign. A pair of electrodes arranged as in green and biased as shown below, can be used to this purpose.



---

[24] Non-uniformities in wafers increase radially, hence the center of the wafer is the most uniform part.

$$F_1 = \frac{\varepsilon_0 \cdot h \cdot (L_{ov} + x)}{2D_Q^2} (V_{DC} - \Delta V)^2 \qquad F_2 = -\frac{\varepsilon_0 \cdot h \cdot (L_{ov} - x)}{2D_Q^2} (V_{DC} - \Delta V)^2$$

$$F_3 = -\frac{\varepsilon_0 \cdot h \cdot (L_{ov} + x)}{2D_Q^2} (V_{DC} + \Delta V)^2 \qquad F_4 = \frac{\varepsilon_0 \cdot h \cdot (L_{ov} - x)}{2D_Q^2} (V_{DC} + \Delta V)^2$$

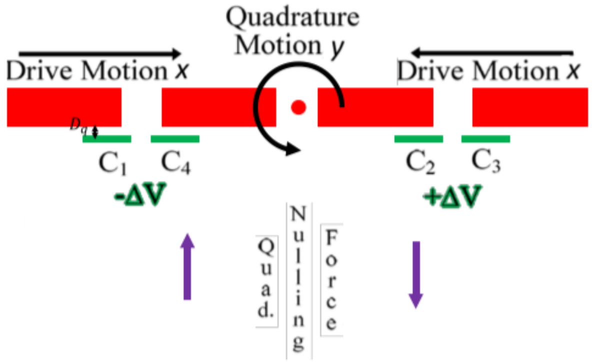
$$\Longrightarrow F_{QC} = -4\,\frac{\varepsilon_0 \cdot h}{D_Q^2}\, V_{DC}\, \Delta V\, x$$

The obtained force is:

- proportional to the drive motion amplitude x_D;
- orthogonal to the drive direction;
- with a modulus that depends of the values of V_{DC}, ΔV and the number of plates N_{QC};
- with a sign which can be set by choosing the sign of ΔV.

The force grows with the number of used electrode pairs, with more electrodes, the same quadrature compensation capability is achieved at lower voltages (but larger area).
With regard to the out-of-plane structure a similar effect can be obtained by using vertical plates with changing overlap with respect to the drive motion.

$$F_{QC} = -4\,\frac{\varepsilon_0 \cdot h}{D_Q^2}\, V_{DC}\, \Delta V\, x$$



6 MAGNETOMETER

To properly perform navigation based on inertial sensors (e.g. the MEMS gyroscopes and accelerometers studied so far), a mandatory element is the knowledge of the initial orientation with respect to Earth surface. Though not, strictly speaking, an inertial sensor, a sensor of the Earth magnetic field can provide this information. Therefore, magnetic field sensors are often associated to gyros and accelerometers in the so-called "9-axis" **Inertial Measurement Units (IMUs)**.

The applications of magnetometers are many:

- Current sensing;
- Electronic compass;
- Vehicle detection;
- Automotive features;
- Bio magnetic signals.

6.1 The Lorentz Force

The law describing the force on a charge moving in an electromagnetic field of vector components $\vec{E}$ and $\vec{B}$ respectively is known to be:

$$\vec{F} = q\,(\vec{E} + \vec{v} \times \vec{B})$$

Where:

- the first component is related to the **electrostatic force** on a charge inside an electric field and it is e.g. responsible of current density $\vec{J}$ flowing through a material of conductivity σ (generalized Ohm's law):
$$\vec{J} = \sigma\,\vec{E} = q\,n\,\mu_n\,\vec{E}$$
- the second component is known as **Lorentz force** and occurs in a direction orthogonal to the plane including the charged particle velocity and the magnetic field vector:
$$\vec{F}_{Lor} = q\,\vec{v} \times \vec{B}$$

If we now consider a conductive wire biased at its ends by different voltage levels, we have electrons flowing, following the electric field direction: the direction of the wire length.

In presence of a magnetic field with a component orthogonal to the wire length, we can calculate the net Lorentz force on the wire, formed by the individual Lorentz forces acting on all the N_{el} charges:

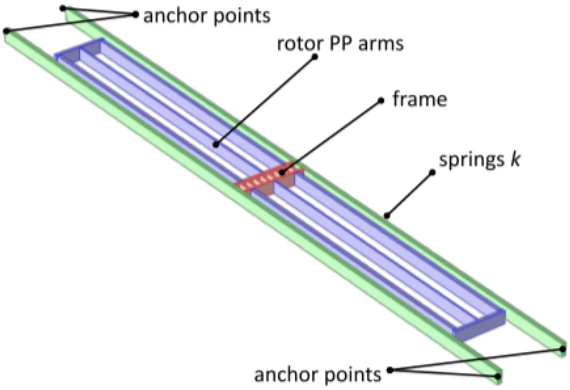
$$\vec{F}_{Lor} = N_{el}\,q\,\vec{v} \times \vec{B}$$

$$\vec{F}_{Lor} = \frac{N_{el}}{A\,L} A\,L\,q\,\vec{v} \times \vec{B} = n\,A\,L\,q\,\vec{v} \times \vec{B} = A\,L\,q\,n\,\mu_n\,\vec{E} \times \vec{B} = L\,A\,\sigma\,\vec{E}\,\times \vec{B} = L\,A\,\vec{J} \times \vec{B}$$

$$\Longrightarrow\ \vec{F}_{For} = L\,\vec{i} \times \vec{B}$$

6.2 General architecture

In its simplest form, the Lorentz-force based MEMS magnetometer is formed by: current-carrying springs of length L (possibly long), as small as possible frame, and rotor arms for capacitive sensing. The design usually sets the operation frequency (resonance) out of the audio and vibration bandwidth (> 20 kHz), like in gyroscopes.



6.2.1 Resonance operation

We first assume to operate the device with a **current at resonance**. This is reasonably motivated by the very low intensity of the Lorentz force, so that at least motion (and thus the displacement x) will be amplified through the quality factor.

$$x = \frac{Q}{2k} i \cdot B \cdot L$$

(The factor 2 at the denominator accounts for the fact that the force is <u>distributed</u> along the springs and not concentrated on the frame: in such case the stiffness doubles!)
The displacement is, as usual, read out through a differential capacitance variation ΔC:

$$\Delta C = 2C_0 \frac{x}{g} = \frac{2\varepsilon_0 NA}{g^2} x = \frac{\varepsilon_0 NA}{g^2} \frac{Q}{k} F_{Lor} = \frac{\varepsilon_0 NA}{g^2} \frac{Q}{k} iBL$$

We now write the expression of the capacitance variation per unit magnetic field, by making explicit the dependence on the damping coefficient b:

$$\frac{\Delta C}{B} = \frac{\varepsilon_0 NA}{g^2} \frac{Q}{k} iL = \frac{\varepsilon_0 NA}{g^2} \frac{\cancel{k}}{\cancel{k}\omega_0 b} iL = \frac{\varepsilon_0 iL}{\omega_0 g^2} \frac{NA}{b}$$

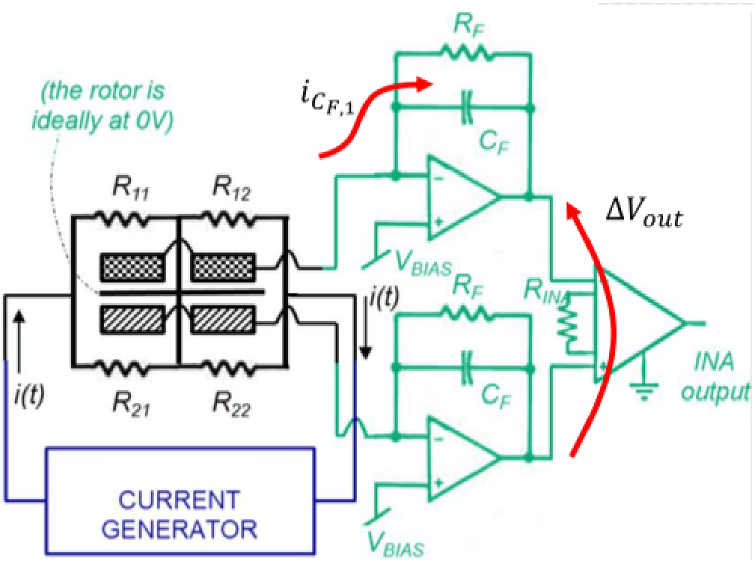
As we will see, in PP cells the dominant contribution (squeezed-film damping) is proportional (somewhat intuitively) to PP length and number. We thus re-write damping through a term per unit area (<u>h is fixed</u>):

$$\frac{\Delta C}{B} = \frac{\varepsilon_0 iL}{\omega_0 g^2} \frac{NA}{b_{\text{area}}\, 2NA} = \frac{\varepsilon_0 iL}{2\omega_0 g^2 b_{\text{area}}}$$

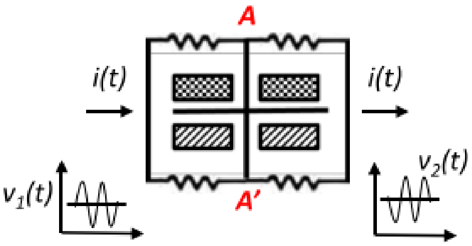
which verifies that the sensitivity is independent on the number of parallel plates.

6.3 Readout Circuit

The **readout circuit** is similar to the sensing of gyros. However note that the frame cannot be arbitrarily biased, as its voltage is determined by the current flowing in the springs.



If we model each spring as a pair of identical resistances ($R_{1,1} = R_{1,2}, R_{2,1} = R_{2,2}$), and if we inject a current through a differential voltage, we can reasonably assume that the voltage on the springs central points (AA') is always **constant** and **null**.



We can thus bias the stators[25] to have a current flowing through the feedback capacitor. The differential output voltage (at the INA input) can be calculated as for the gyroscope readout:

$$i_{C_{F,i}} = V_{BIAS}\frac{dC_{S,i}}{dt} = V_{BIAS}\frac{dC_{S,i}}{dx}\frac{dx}{dt} = V_{BIAS}\frac{dC_{S,i}}{dx}x\omega_0$$

$$\Delta V_{out} = 2\frac{i_{C_{F,i}}}{\omega_0 C_F} = \frac{V_{BIAS}\Delta C\cancel{\omega_0}}{\cancel{\omega_0}C_F} \quad \longrightarrow \quad \frac{\Delta V_{out}}{\Delta C} = \frac{V_{BIAS}}{C_F}$$

$$\frac{\Delta V_{out}}{B} = \frac{\Delta C}{B}\frac{\Delta V_{out}}{\Delta \mathrm{C}} = \frac{\varepsilon_0 i L}{2\,\omega_0\, g^2\, b_{\mathrm{area}}}\frac{V_{BIAS}}{C_F}$$

---

[25] Stators are biased at V_{BIAS} because of the negative feedback of the CAs that have a terminal at V_{BIAS}

6.4 Resonant operation issues

Bandwidth

Like for gyroscopes, the information on the value of the magnetic field is modulated around the drive frequency (which in this case is the Lorentz current frequency). The maximum sensing bandwidth is again given by the half width at half maximum (i.e. -3 dB) of the resonance peak:

$$\Delta\omega_{BW} = \frac{\omega_0}{2\,Q} = \frac{b}{2\,m} \qquad\qquad \Delta f_{BW} = \frac{f_0}{2\,Q} = \frac{b}{4\pi m}$$

Brownian noise equivalent magnetic field density

To calculate the Brownian noise we start as usual from the force noise density and we turn it to displacement and then into magnetic field to get **N**oise **E**quivalent **M**agnetic field **D**ensity:

$$S_{Fn} = 4\,k_B\,T\,b \qquad\qquad \sqrt{S_{Xn}} = \frac{i\,L\,Q}{2k}\,\sqrt{S_{Bn}}$$

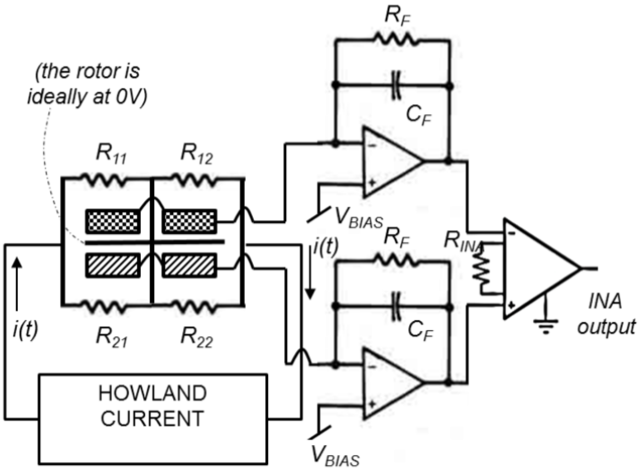
$$\Longrightarrow \sqrt{S_{Bn}} = \frac{\sqrt{4\,k_B\,T\,b\left(\frac{Q}{k}\right)^2}}{\frac{i\,L}{2k}Q} = \frac{4}{iL}\sqrt{k_BTb} = \textbf{NEMD}$$

The challenge is in improving the minimum detectable field:

- increasing the length (pay in area, and chip size and cost);
- increasing the injected current (pay in power consumption!);
- decreasing the damping coefficient (e.g. by decreasing the pressure).

6.5 Electronic noise

Like in the case of gyros, electronic noise is given by two major contributions: the feedback resistance and the OpAmp noise.



$$S_{\mathrm{B}n,R_F} = \frac{\sqrt{2\cdot\frac{4k_bT}{R_F}\frac{1}{\omega_0 C_F}}}{\frac{\Delta V_{out}}{\mathrm{B}}} = \frac{\sqrt{\frac{8k_bT}{R_F}}2g^2b_{area}}{\varepsilon_0 iLV_{BIAS}}$$

$$\begin{aligned}S_{\mathrm{B}n,CA} &= \frac{\sqrt{2\cdot S_{n,op}}\left(1+\frac{C_P}{C_F}\right)}{\frac{\Delta V_{out}}{\mathrm{B}}} = \\ &= \sqrt{2\cdot S_{n,op}}\left(1+\frac{C_P}{C_F}\right)\frac{2\omega_0 g^2 b_{area} C_F}{\varepsilon_0 iLV_{BIAS}} = \\ &\approx \sqrt{2\cdot S_{n,op}}C_P\frac{2\omega_0 g^2 b_{area}}{\varepsilon_0 iLV_{BIAS}}\end{aligned}$$

There is a difference with respect to the gyro case: here the constant voltage V_{BIAS} at the stators cannot be as high as it was for the rotor in gyros (up to 10 V), because the OpAmps operate between the power supply!

We can put together the found expressions to write the overall input-referred noise in terms of magnetic field density:

$$\begin{aligned}\sqrt{S_{\mathrm{B}n,\text{ tot}}} &= \sqrt{\left(\sqrt{2\cdot S_{n,op}}C_P\frac{2\omega_0 g^2 b_{\text{area}}}{\varepsilon_0 iLV_{BIAS}}\right)^2 + \left(\frac{\sqrt{\frac{8k_bT}{R_F}}2g^2b_{\text{area}}}{\varepsilon_0 iLV_{BIAS}}\right)^2 + \left(\frac{4}{i\cdot L}\sqrt{k_BTb}\right)^2} = \\ &= \frac{1}{iL}\sqrt{\left(\sqrt{2\cdot S_{n,op}}C_P\frac{2\omega_0 g^2 b_{\text{area}}}{\varepsilon_0 V_{BIAS}}\right)^2 + \left(\frac{\sqrt{\frac{8k_bT}{R_F}}2g^2b_{\text{area}}}{\varepsilon_0 V_{BIAS}}\right)^2 + \left(4\sqrt{k_BT\,b}\right)^2}\end{aligned}$$

- If the **electronic noise dominates**: we can act on any parameter related to the **sensitivity** (gap, bias voltage) or we should minimize **parasitics** (C_P) and maximize the feedback **resistance** R_F;
- If **device noise dominates**: we don't have so many options!

6.6 Advanced architectures

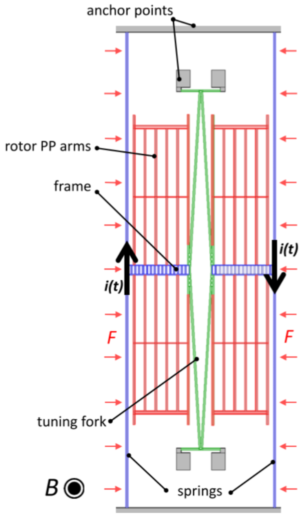
The consideration that magnetometers are often coupled to inertial MEMS (gyros, axels + pressure sensors) to form multi-axis, multi-parameter inertial measurement and indoor localization units, drives the research towards the realization of magnetic sensors in MEMS processes: a single technology for an entire 9- or 10-axis IMU!

Several advantages of the "MEMS" approach for magnetometers push the research towards solutions to simultaneously improve:

- robustness;
- noise density vs bandwidth performance;
- area.

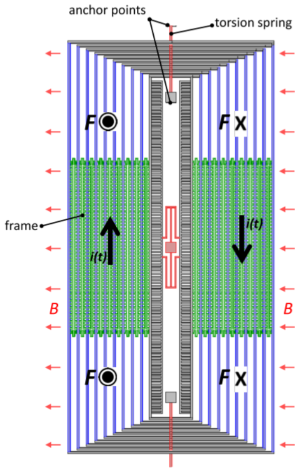
6.6.1 Differential architecture of out-of-plane field

A differential architecture can be obtained if the device is split into two halves, with the current circulating in opposite directions. The Lorentz force will itself have different directions on the two device halves. Instead, the accelerations, which have the same effect on the two halves, will be seen as a common mode and automatically cancelled by the capacitive readout.



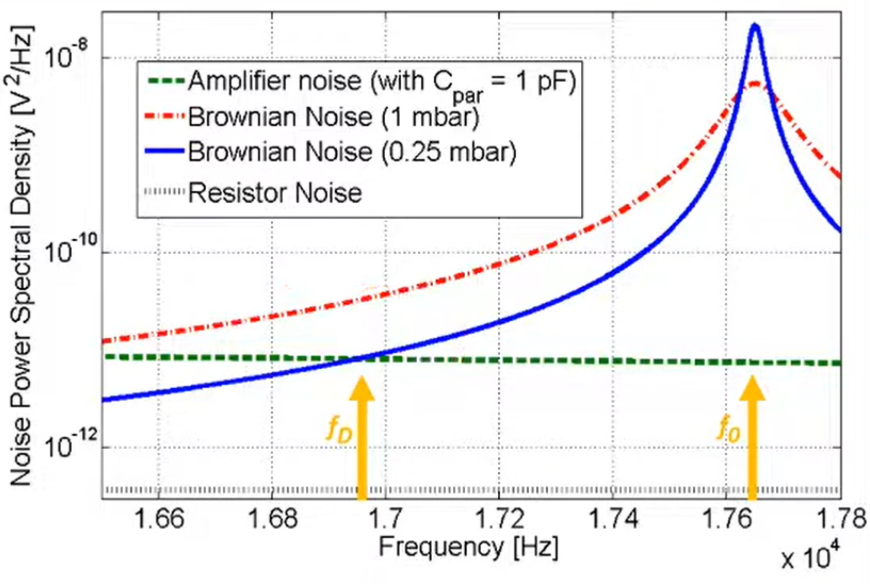
6.6.2 Differential architecture for in-plane field

Devices moving in the OOP direction (i.e. sensitive to in-plane fields) correctly show a differential readout and reject accelerations provided that they are balanced (in terms of gravity) with respect to the rotation axis, and the current flows in opposite directions.



6.7 Off-resonance (mode-split) operation

Unlike axels and gyros, mags suffer from a non-negligible thermomechanical noise. Lowering the damping coefficient would imply a consistent reduction in the maximum achievable bandwidth. Mimicking the behavior of mode-split gyroscopes, our lab first proposed the operation of MEMS magnetometers with a frequency split between driving current and resonance frequency, with the aim of solving the **bandwidth-resolution trade-off**.
The figure summarizes this concept:



The sensitivity can be calculated as in the case of resonant operation, provided the quality factor is substituted by the gain at a distance $\Delta f = f_0 - f_D$, which we called **effective quality factor** (a.k.a. Leeson's effect):

$$Q_{eff} = \frac{f_0}{2\,\Delta f} = \frac{\omega_0}{2\,\Delta\omega}$$

For a tuning-fork structure, we can evaluate the scale-factor expression for just half of the device and then simply multiply by a factor 2:

$$k_{1/2} = k_{1-beam} + k_{1/2\,TF} \qquad\qquad x = B\,i\,L\,\frac{Q_{eff}}{2\,k_{1/2}}$$

$$\Delta C = 2\frac{C_0}{g}x = \frac{C_0}{g}B\,i\,L\frac{Q_{eff}}{k_{1/2}} = \frac{C_0}{g\,k_{1/2}}B\,i\,L\frac{f_0}{2\,\Delta f}$$

$$\Longrightarrow \frac{\Delta V_{out}}{B} = \frac{V_{BIAS}}{C_F}\frac{C_0}{g}\frac{i\,L}{k_{1/2}}\frac{f_0}{2\,\Delta f} = \frac{V_{DC}}{C_F}\frac{\varepsilon_0\,A\,N_{PP}}{g^2}\frac{i\,L}{k_{1/2}}\frac{f_0}{2\,\Delta f}$$

Design guidelines for high sensitivity:

- This time the sensitivity is dependent on the number of PP! Indeed, Q_{eff} is independent from b!
- At a given f_0, the sensitivity grows with a lower k (and a lower m).
- The "gain-bandwidth" product is constant (the lower Δf, the larger the sensitivity).
- Usual PP capacitive dependence on the inverse of the squared gap.
- Obvious dependence on the Lorentz force intensity (i and L).

6.8 Noise in off-resonance operation

For what matters the **achievable SNR**, all the expressions found for resonant operation are still valid, provided that we substitute the quality factor Q with the effective quality factor Q_{eff}. The -3 dB bandwidth value Δf_{eln} is now given by an electronic LPF.

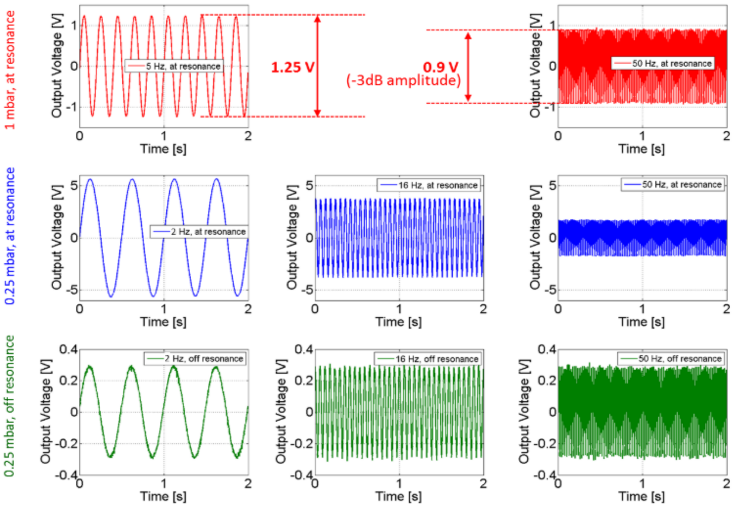
$$\frac{x}{B} = \frac{i\,L}{2\,k_{1/2}} Q_{eff}$$

$$\sqrt{S_{Xn}} = \sqrt{4\,k_B\,T\,b\,\left(\frac{Q_{eff}}{k_{1/2}}\right)^2}$$

So the NEMD doesn't change:

$$NEMD = \sqrt{S_{Bn}} = \frac{\sqrt{S_{Xn}}}{x/B} = \frac{4}{i\,L}\sqrt{k_B\,T\,b}$$

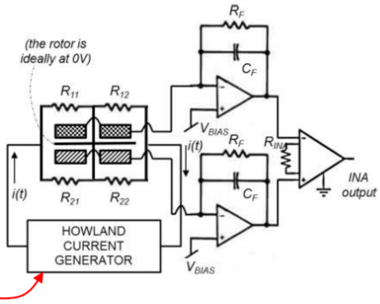
This means that b can be lowered (by decreasing package pressure) to improve the NEMD, while the maximum sensing BW is chosen through an electronic LPF with a cutoff at $\approx \Delta f/2$ to $\Delta f/3$.



If the electronic filter is made programmable, the sensor can achieve high-performance or high-bandwidth in a dual-mode operation:

- low LPF pole → low overall rms noise, small BW;
- high LPF pole → large sensing BW, worse rms noise performance.

The working principle, as assumed so far, implies the injection of a current exactly at the desired frequency f_D.



who provides the frequency reference to this block??!!

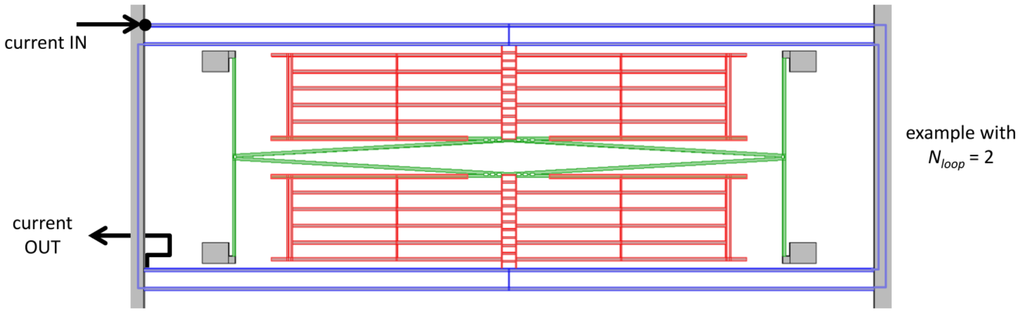
This could be achieved with an electronic oscillator (e.g. a ring oscillator) in the ASIC. However, this cannot solve the problem, because the MEMS frequency drifts with temperature and thus the ring oscillator frequency would match the desired one only at a certain T.
A possible solution is to use an oscillator at the chosen frequency split through a MEMS resonator put in the same package of the magnetometer. Indeed, in case of temperature changes, the resonator and the mag will drift (almost) in the same way, without consistent changes in Δf and sensitivity. In this way, we basically have the same advantages that exist between mode-matched and mode-split gyroscopes!

6.9 Multi-loop architectures

Like in the case of gyros and axels, electronic noise is due to two major contributions: the feedback resistance and the OpAmp noise. Since the bandwidth-resolution trade-off is solved, the damping can be decreased to achieve a better SNR if the thermomechanical noise dominates. However, as we decreased the device gain (and thus the overall sensitivity $\frac{\Delta V_{out}}{B}$), **electronic noise will likely dominate**, so that advantages of off-resonance mode will be somewhat reduced:

$$\sqrt{S_{Bn}} = \frac{4}{iL}\sqrt{k_B T b} \qquad \sqrt{S_{Bn,R_F}} = \frac{\sqrt{2 \cdot \frac{4k_bT}{R_F}\frac{1}{\omega_0 C_F}}}{\frac{\Delta V_{out}}{B}} \qquad \sqrt{S_{Bn,CA}} = \frac{\sqrt{2 \cdot S_{n,op}}\left(1+\frac{C_P}{C_F}\right)}{\frac{\Delta V_{out}}{B}}$$

We thus need to find out a way to re-boost the device gain (remember that the Lorentz force is natively very tiny).
Since we cannot act on the length (area constraints) and on the current intensity (power dissipation constraints), the only way to boost the sensitivity is to act on the Lorentz force. To achieve an increase in Lorentz force, we can reuse the current by making it re-circulate several times (N_{loop}), thus obtaining a N_{loop} factor boost in sensitivity!



New expressions of sensitivity and NEMD become:

$$\frac{\Delta V_{out}}{B} = \frac{V_{DC}}{C_F}\frac{\varepsilon_0 \, A \, N_{PP}}{g^2 \, k_{1/2}} i \, N_{loop} \, L Q_{eff}$$

$$NEMD = \frac{4}{N_{loop} \, i \, L}\sqrt{k_B \, T \, b}$$

$$\sqrt{S_{\mathrm{B}n,\ \mathrm{tot}}} = \sqrt{\left(\sqrt{2 \cdot S_{n,op}} C_P \frac{2\omega_0 g^2 b_{\mathrm{area}}}{\varepsilon_0 i L V_{BIAS} \, N_{loop}}\right)^2 + \left(\frac{\sqrt{\frac{8k_bT}{R_F}} 2g^2 b_{\mathrm{area}}}{\varepsilon_0 i L V_{BIAS} \, N_{loop}}\right)^2 + \left(\frac{4}{N_{loop} \, i \cdot L}\sqrt{k_B T b}\right)^2} =$$

$$= \frac{1}{N_{loop} \, i \, L}\sqrt{\left(\sqrt{2 \cdot S_{n,op}} C_P \frac{2\omega_0 g^2 b_{\mathrm{area}}}{\varepsilon_0 V_{BIAS}}\right)^2 + \left(\frac{\sqrt{\frac{8k_bT}{R_F}} 2g^2 b_{\mathrm{area}}}{\varepsilon_0 V_{BIAS}}\right)^2 + \left(4\sqrt{k_B T \, b}\right)^2}$$

Adding recirculating loops **decreases all input-referred noise contributions**, as it directly increases the Lorentz force!

6.10 System optimization for power consumption

Given an available current consumption i_{tot}, how much would you give to the sensor (i_{MEMS}), and how much to the electronics ($N_{MOS} \, i_{MOS} + i_{OSC}$), to **optimize noise**?
As a first approximation, the current needed by the oscillator to sustain the resonator does not influence noise (we assume it as a constant term,i_{OSC}). The current consumed by the sense CAs, on the contrary, determines to a large extent the electronic noise S_{Vn}:

$$S_{Vn,MOS} = \frac{\gamma \, 4 \, k_B \, T}{g_m} = \frac{\gamma \, 4 \, k_B \, T}{2\sqrt{k_n \, i_{MOS}}}$$

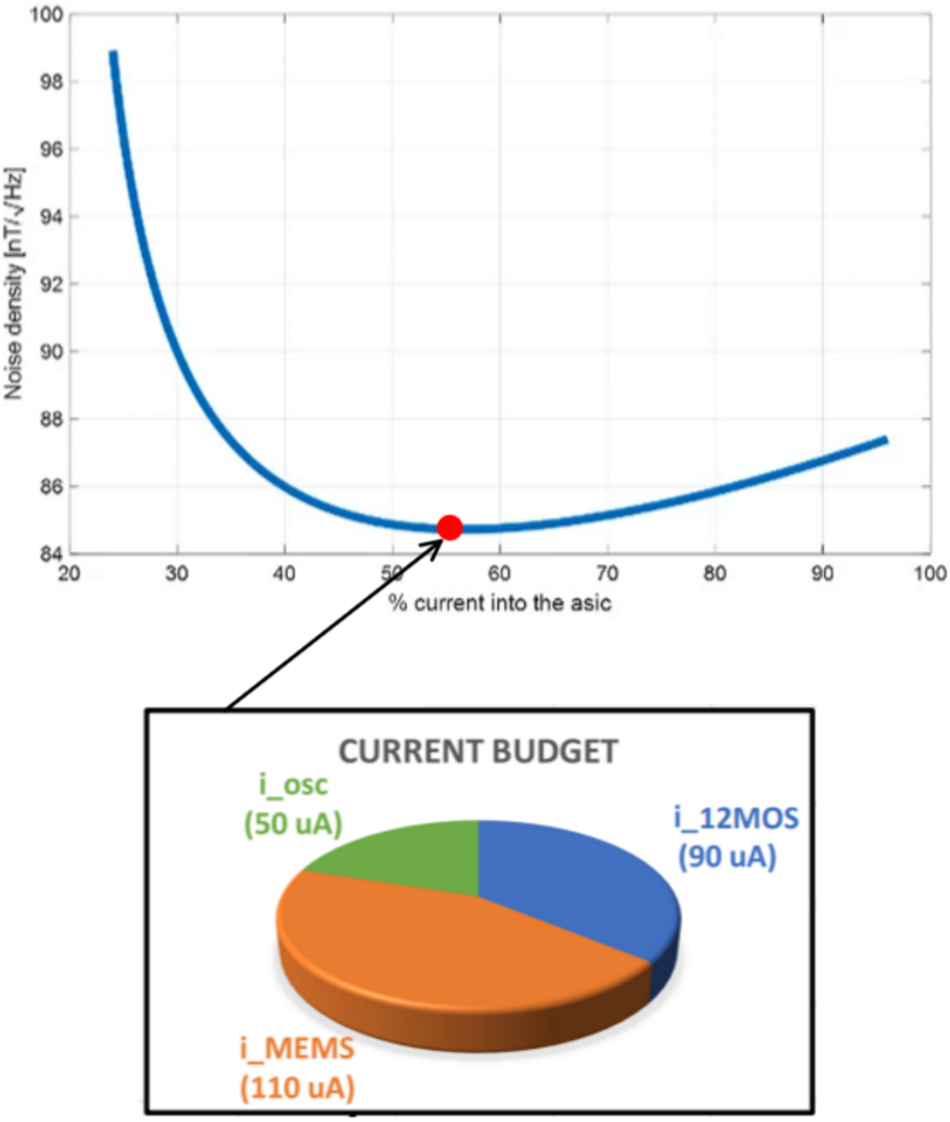
A typical differential operational amplifier has a pair of input transistors, each with its own voltage noise component. The amplifier noise that we modeled so far, which we named S_{Vn} corresponds thus to twice the noise of a single transistor. We neglect noise of transistors of the following stages, which is usually made negligible.

$$S_{Vn} = 2\frac{\gamma 4 k_B T}{g_m} = \frac{\gamma 4 k_B T}{\sqrt{k_n i_{MOS}}}$$

Note that, if you have two amplifiers for differential sensing, then noise power spectral density should be multiplied by another factor 2.
We now write the overall noise equation as a function of the current injected in the MEMS (i_{MEMS}) and that of the input pair (i_{MOS}).

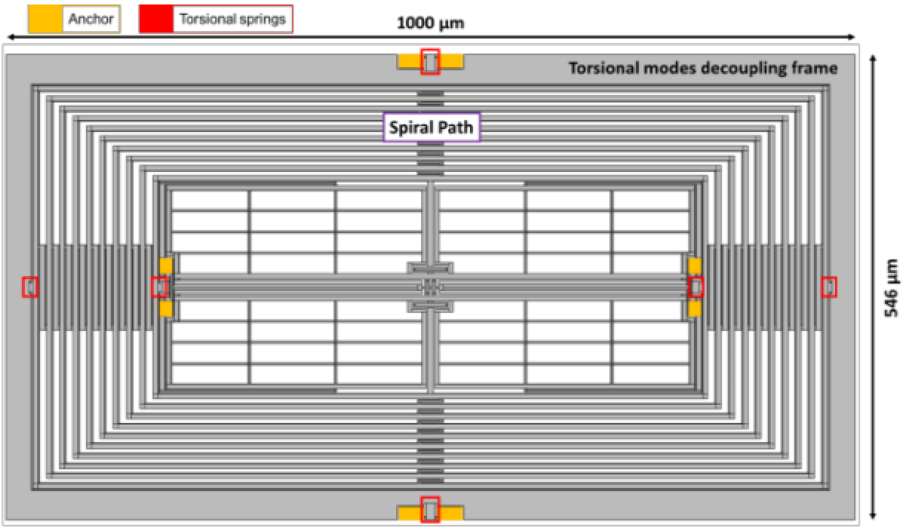
$$\sqrt{S_{\text{Bn},tot}} = \frac{1}{N_{\text{loop}}\,\boldsymbol{i}_{mems}L}\sqrt{\left(\sqrt{2\frac{\gamma 4k_BT}{\sqrt{k_n\boldsymbol{i}_{MOS}}}}\frac{C_P}{V_{BIAS}}\frac{g^2k_{1/2}2\Delta\omega}{\epsilon_0AN_{PP}\omega_0}\right)^2+\left(\sqrt{\frac{8k_bT}{R_F}}\frac{g^2k_{1/2}2\Delta\omega}{2\epsilon_0AN_{PP}V_{BIAS}\omega_0^2}\right)^2+\left(4\sqrt{k_BTb}\right)^2}$$

As an example, assuming a current of 50 μA to sustain the oscillator (no AGC needed), the optimum partitioning between MEMS current and overall ASIC current results to be well balanced (about 50%).



6.11 Monolithic multi-loop architectures

The solution presented thus far, still takes up an overall area of about $2 \times 2.5mm^2$. This is to be compared to $1 \times 1mm^2$ areas of non-MEMS solutions. Therefore, we need to reduce the devices' area approximately by a factor 5.
To solve this issue, we can try to use the same 10-loop path to excite three different sensing axis in a monolithic magnetic field sensor, instead of designing three times 10 recirculation loops.
This is the case of a **monolithic 3-axis magnetometer**, where a single, 10-loop path is used, with three sensing modes in a single monolithic structure.



The sensor is currently still under performance evaluation. Noise as low as $85nT/\sqrt{Hz}$ was achieved and the goal is to couple it with an accelerometer to form a **6-DOF "accel+mag" compass unit**.

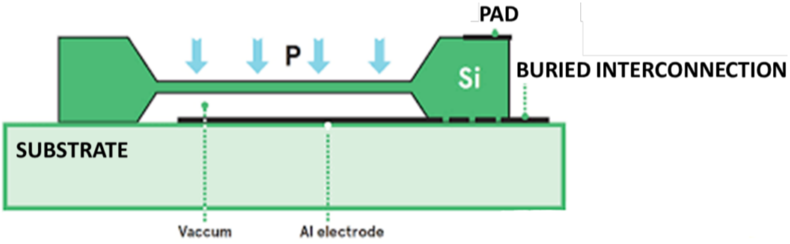
Some design rules. For this kind of magnetometer it is important to take into account the following design rules:

- **Resonant mode should be different** for every axis sensing, to avoid cross-coupling;
- Deformation due to one mode **shouldn't excite other modes** (cross-talk):
  - Avoiding spurious capacitance variations is paramount since we operate with such low forces!

6.12 Other capacitive MEMS sensors

6.12.1 Pressure sensors

The general concept is to use a **hermetic-sealing membrane** between a known pressure level (typically much lower than ambient pressure) and the quasi-DC ambient absolute pressure value. The membrane can be obtained from an epitaxial growth and a further, selective, thinning and filling of the release holes. The difference in pressure makes the suspended membrane **deflect** as a function of the outside pressure. This deflection generates a capacitive change (single ended, in this example), which is readout through capacitive sensing interfaces.

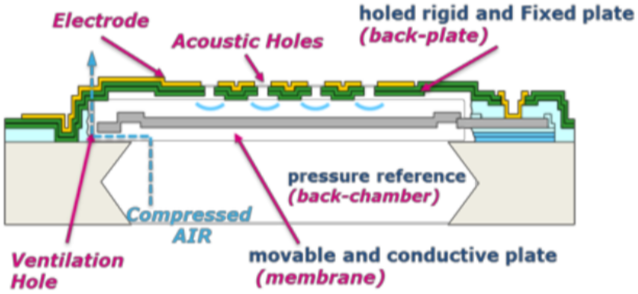


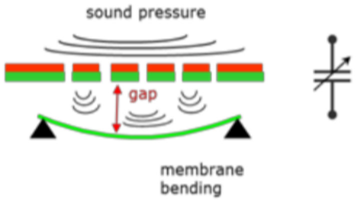
The sensitivity will be:

$$S = \frac{\Delta V_{out}}{P} = \frac{V_{DD}}{C_F}\frac{C_0}{g}\frac{dy}{dP} = \frac{V_{DD}}{C_F}\frac{C_0}{g}\frac{A}{2k}$$

6.12.2 Microphones

Acoustic waves are pressure waves: a microphone needs thus to sense the AC pressure, usually from 20 Hz to 20 kHz, while leaving out the DC value of pressure. To let this happen, a fixed membrane features specific holes (acoustic holes) in such a way that there is no DC pressure difference between outside and the two chambers. However, for AC waves, there is an effective AC pressure difference between the back chamber and the front chamber, which makes the membrane deflect.





7 CHARACTERIZATION

Characterization is a very relevant field for the MEMS industry and research. Learning about MEMS characterization methods and parameters is thus appropriate for:

- electromechanical characterization (Q and f_0 measurements);
- sensitivity, linearity and bandwidth characterization (as already discussed);
- noise and stability characterization.

This section's aim is to introduce a general method to characterize noise and stability performance of a sensor, specifically (but not only) useful in long-term operation when signals need to be integrated. This is the case of position/angle recovery from acceleration/velocity sensors.
The method is called (**root**) **Allan Variance**. The goal is to understand the reasons to use such a method/graph and to learn how to compare the performance of different sensors of the same type when looking at their root Allan variance graphs.

7.1 Case study: unassisted navigation

Navigation has dramatically evolved during the last years thanks to accurate mapping services and satellites system (GPS) triangulation. However, GPS-assisted navigation has some issues:

- GPS is **power hungry**;
- GPS will not work indoor, in skyscrapers areas, in tunnels, underwater;
- GPS has relatively poor accuracy;
- GPS has high development, maintenance and dismantling costs and pollutes space.

A possible, promising alternative, is the use of inertial sensors. Indeed inertial sensors measure linear and rotational motion and could be thus used to reconstruct a trajectory (dead reckoning) on a map, without any additional information from external nodes (GPS, Wi-Fi nodes, etc. . .), except for initial positioning. However, the sensors seen so far require to integrate the measured signals for positioning applications:

- **position**: reconstruction of x,y and z

$$x(t) = x_0 + \int_0^t v(t)dt = x_0 + \int_0^t \left[v_0 + \left(\int_0^t a(t)dt \right) \right] dt$$

- **angle**: reconstruction of θ, δ and γ

$$\theta = \theta_0 + \int_0^t \Omega(t)dt$$

An **INS** (**inertial navigation system**) can combine GPS signals for absolute «coarse» positioning and an IMU for highly-accurate «local» positioning.

Let us assume (reasonably) a constant offset associated to the noisy sensor output. The equations can be written as:

$$x(t) = x_0 + \int_0^t \left(v_0 + \int_0^t a_S(t) + a_{OS} + a_{noise}(t)dt \right) dt$$

$$\theta = \theta_0 + \int_0^t \Omega_{\mathrm{S}}(t) + \Omega_{OS} + \Omega_{\text{noise}}\ (t)dt$$

Offset quickly results in positioning errors while time is passing:

- e.g. a position error of 71 m after an integration for 2 minutes (120 s) of a **1 mĝ** offset;
- e.g. an angle error of 60° after an integration for 2 minutes of a **0.5 dps** offset;

Offset is not the only non-ideality affecting the sensor output. We also have unavoidable noise, with the different physical origins that we studied deeply during the course (electronic, thermomechanical, quantization). Noise, as a statistical quantity, cannot be compensated during a calibration phase, so we need to take into account its effect.

$$x(t) = x_0 + \int_0^t \left(\int_0^t a_S(t) + a_{noise}(t)dt \right) + v_0 dt$$

$$\theta = \theta_0 + \int_0^t \Omega_{\mathrm{S}}(t) + \Omega_{\text{noise}}\ (t)dt$$

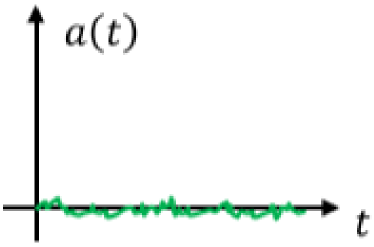


Figure 45: Zero mean noise

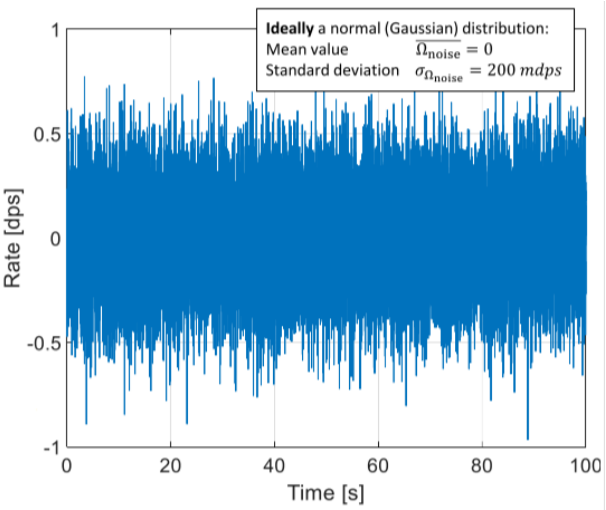
The goal becomes the study of consequences deriving from integration of (nominally) **"zero-mean" noise**!

7.2 Case study: noise from a gyroscope

To study the problem of noise integration, we consider the simplest situation below:

- 3-axis IMU for in-plane navigation (just x/y axels and z gyro);
- no errors from ideal x/y accelerometers;
- integration of angular rate from a noisy gyroscope (one integration only);
- no angular rate signal and no offset ($\Omega = 0$) either we are stationary, or moving by linear motion only.

An **ideal** (noiseless) gyroscope in this situation would give a constant, null, navigation angle $\theta = 0$ (with respect to the initial angle θ_0, e.g. measured with the magnetometer).
Let us see which output is produced by a real, noisy, gyroscope. For the sake of simplicity, we initially assume white noise only. The figure represents 100 seconds of a noisy gyroscope output with nominal standard deviation (input-referred as an angular rate) of 200 $mdps_{rms}$.

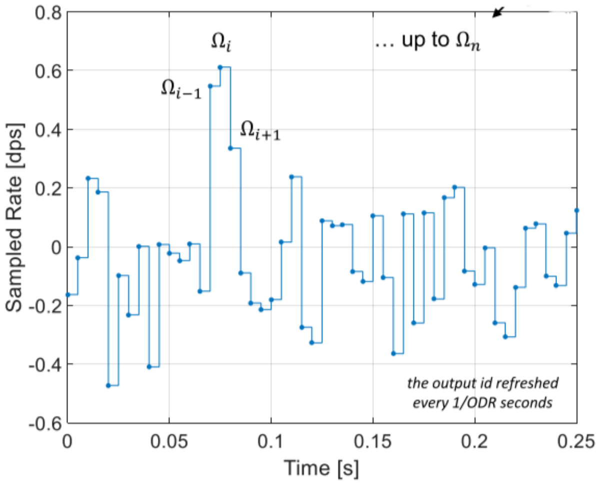


The output data rate (**ODR**) (i.e. the frequency of output refresh) is of 200 $Hz = 200\ Sample/s$. According to the sampling theorem, this corresponds to a maximum $BW \leq 100Hz$. The total number of samples equals the ODR (200 Hz) multiplied by the observation time (100 s) so: 20000 samples. At any point in time, the angle (which should be ideally 0) can be estimated through integration from the rate values up to the actual point:

$$\theta = \theta_0 + \int_0^t \Omega_{\mathrm{S}}(t) + \Omega_{\mathrm{noise}}(t)dt$$

$$\Omega_{\mathrm{S}} = 0 \qquad \theta_0 = 0 \qquad \theta = \int_0^t \Omega_{\mathrm{noise}}\ (t)dt$$

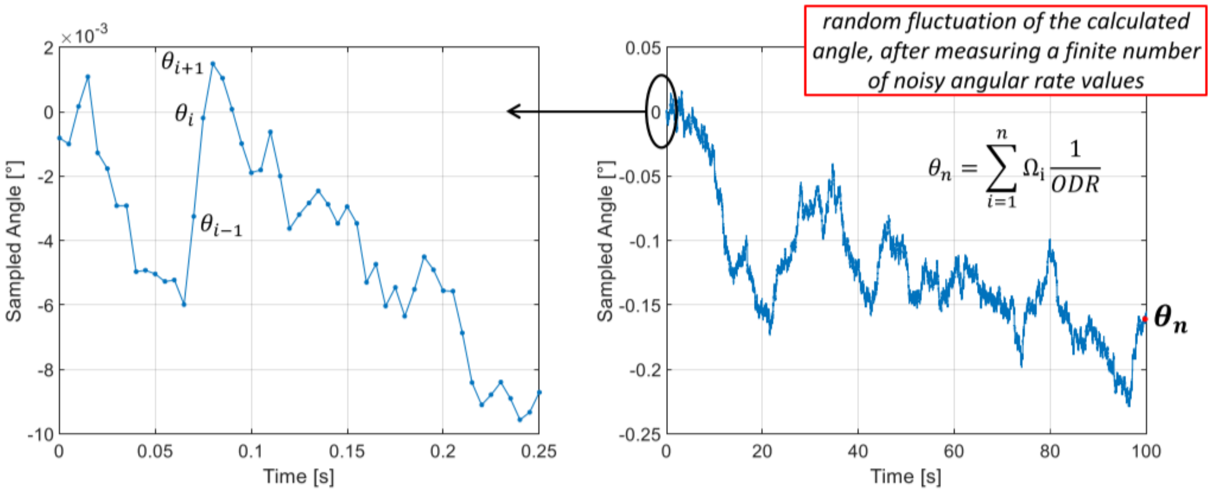
We zoom on the figure for the first sampled time steps:



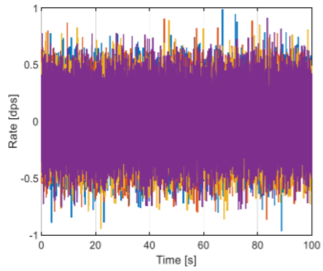
As the output is updated at discrete points in time, we write the above integral as a sum:

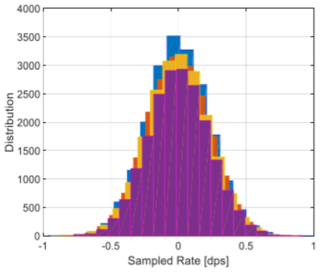
$$\theta_n = \sum_{i=1}^{n} \Omega_{\mathrm{i}} \Delta t_i = \sum_{i=1}^{n} \Omega_{\mathrm{i}} \frac{1}{ODR}$$

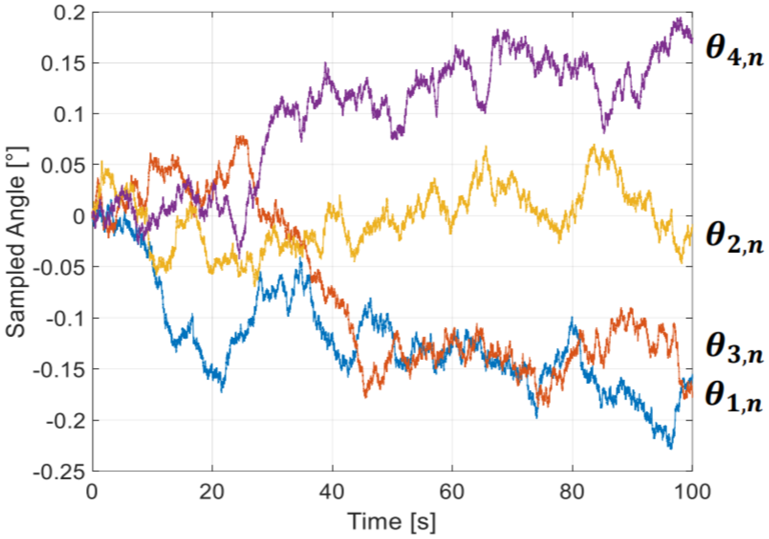
We define "**angle random walk**" the random behavior that the angle undergoes as a consequence of the noise integration described in the previous page. An example of the result is given, for just one random noise distribution lasting 100 s.



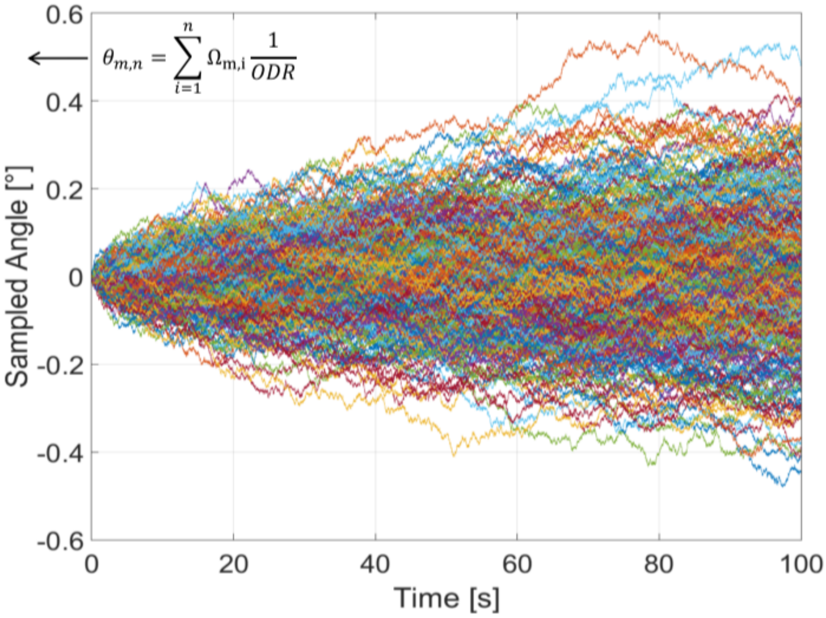
Since noise is a statistical quantity, if we repeat multiple times the same experiment with different random white noise, even though the noise probability distribution is similar in terms of mean value and standard deviation, we obtain different angles after the same time interval.







Increasing the number of experiments to m (a large number), we create a statistics of obtainable final angles at a given observation time $t = \frac{n}{ODR}$.



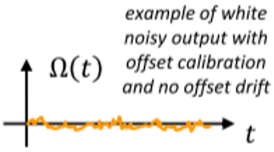
In summary:

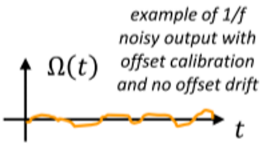
- a "zero-mean" noisy rate has non-null mean value after a finite time slot integration;
- this generates a non-null final angle;
- repeating the experiment m-times, we find a statistics of final angle random walk.

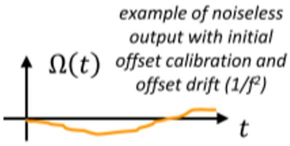
7.3 Allan Variance

We look for a single technique that simultaneously identifies effects of **offset drift** and **noise** (angle random walk, ARW) on the sensor performance.
You can see offset drift and noise in a similar way:

- **white noise** causes rapid random variation of the output;
- **1/f noise** causes slower random output variations;
- **offset drift** causes very slow variation of the output (this variation depends e.g. on temperature or relative humidity in a deterministic manner; however if offset variations vs T/RH are not calibrated, the behavior can be assumed as random – we do not know how T/RH are changing) → $1/f^2$.



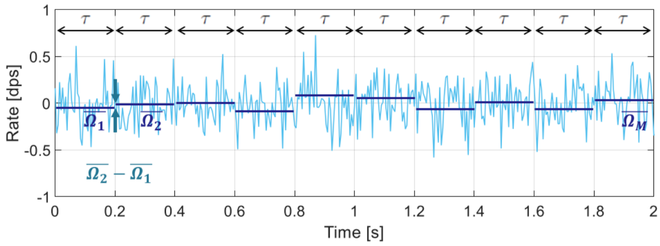




The stability over time of the output signal (affected by noise and offset drift only) is calculated as the minimum average change in consecutive gyroscope rate measurements when analyzed under null angular rate over varying sampling times. The **Allan Variance** (**AV**) is defined to capture this effect. Indeed, stability in an AV plot is qualified by the minimum rate and its corresponding observation time. We can also give a practical definition of Allan Variance:

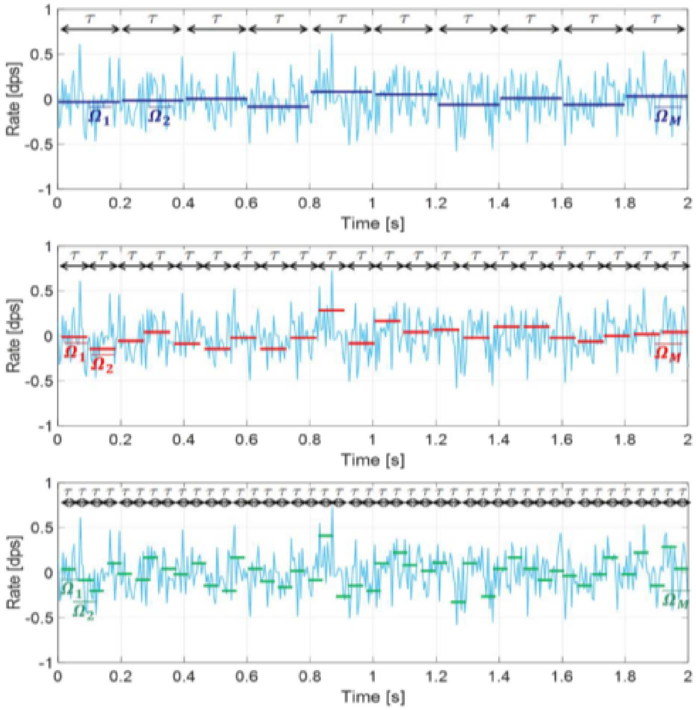
- Acquire a signal for an overall time length t_{TOT}.
- Split the acquisition interval in M slots, each of a same duration τ.
- Calculate the average value $\overline{\Omega_k}$ of the signal along each slot.
- Calculate the changes in consecutive measurements (to make variations over time explicit).
- We define the Allan Variance for a specific observation interval τ the quantity:

$$\sigma^2_{AV,\Omega}(\tau) = \frac{1}{2(M-1)} \sum_{k=1}^{M-1} \left(\overline{\Omega_{k+1}} - \overline{\Omega_k}\right)^2$$

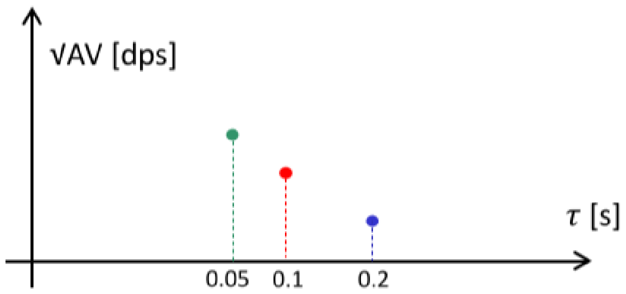


Let's see how we construct a root AV graph:

- Set τ on the x axis and $\sqrt{\sigma^2_{AV,\Omega}(t)}$ on the y axis.
- Calculate a first value by splitting the total time with a first value of τ. Report the point on the graph.
- Then, choose a different interval τ. Report the second point on the graph. And so on for other τ.
- Take care to have: $\tau_{max} = \frac{1}{10}\, t_{TOT}$ and $\tau_{min} = 10\, t_{sample}$

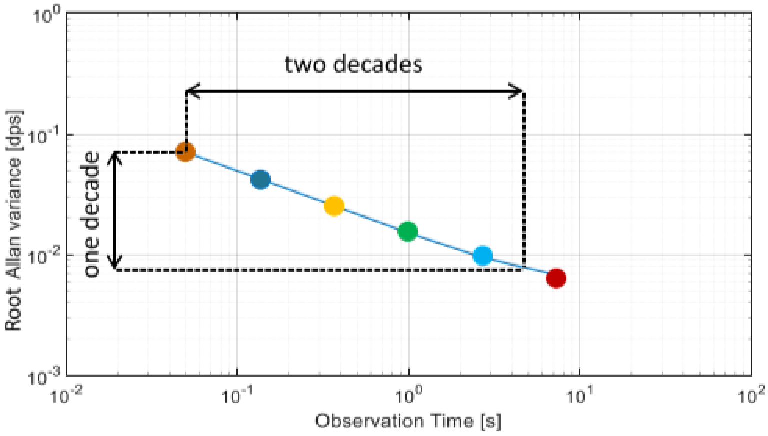


If we plot the root of the AV as a function of τ:



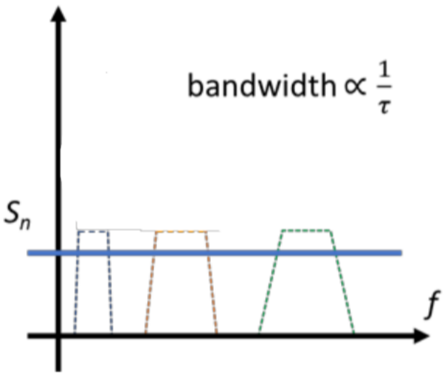
7.3.1 White Noise

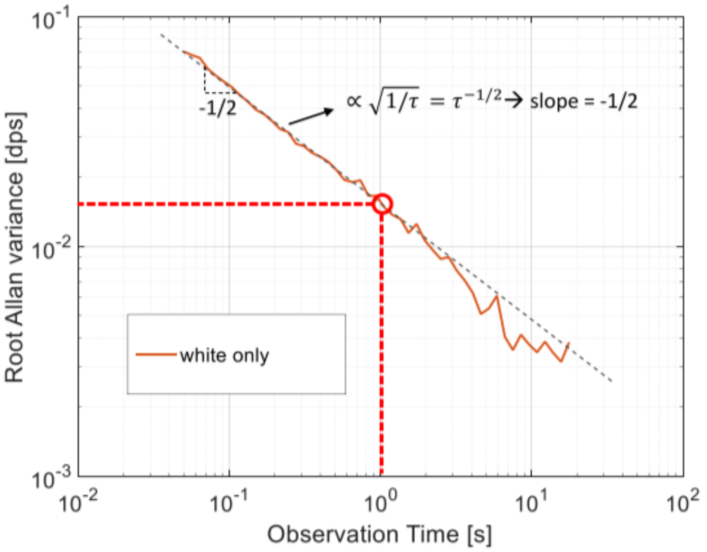
For **white noise only**, one finds that the root Allan Variance graph decreases with a slope equal to $-1/2$ in a log log scale ($\frac{1}{\sqrt{\tau}}$). This is expected, as white noise gets averaged with the square root of the n of samples as well as with the averaging time τ.



Therefore, for a **white power spectral density**, observing for a longer time (i.e. a lower BW, $\Delta f_{BW} = \frac{1}{\tau}$) means getting a lower rms value. There is indeed a perfect equivalence between the white noise density and the slope coefficient of the root Allen Variance (RAV):

$$\sigma^2_{AV,\Omega}(\tau) = 2\int_0^{\infty} S_{\Omega,W}\frac{\sin^4(\pi\tau f)}{(\pi\tau f)^2}df = \frac{S_{\Omega,W}}{2\tau}$$





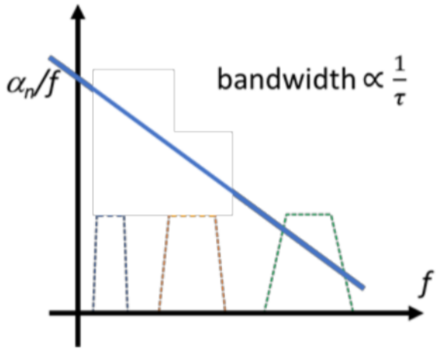
If we want to evaluate the white noise from an Allan variance plot:

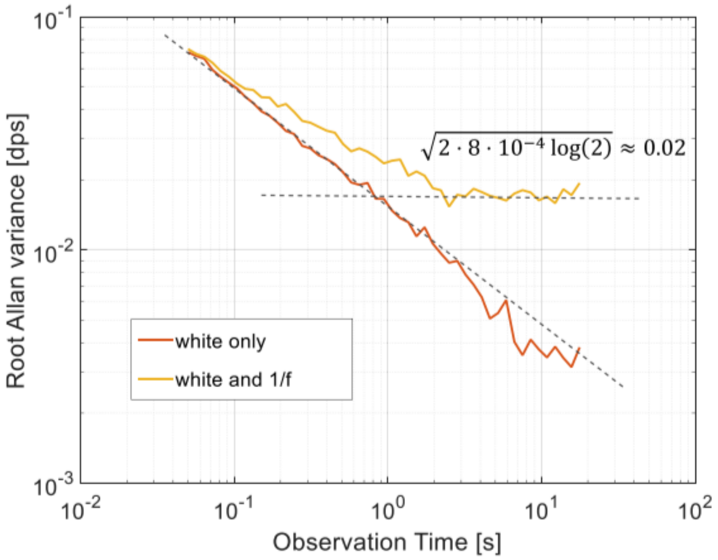
$$\sigma^2_{AV,\Omega}(\tau) = \frac{S_{\Omega,\mathrm{W}}}{2\tau} \qquad \sigma_{AV,\Omega}(\tau) = \sqrt{\frac{S_{\Omega,\mathrm{W}}}{2\tau}}$$

7.3.2 1/f noise

For a 1/f power spectral density, integrating for a longer time means filtering with a smaller bandwidth a larger noise value. In other words, increasing τ means increasing the number of averages but with more and more 1/f noise. As a result, the AV value remains constant.

$$\sigma^2_{AV,\Omega}(\tau) = 2\int_0^{\infty} \frac{\alpha_n}{f}\frac{\sin^4(\pi\tau f)}{(\pi\tau f)^2}df = 2\alpha_n \log(2)$$





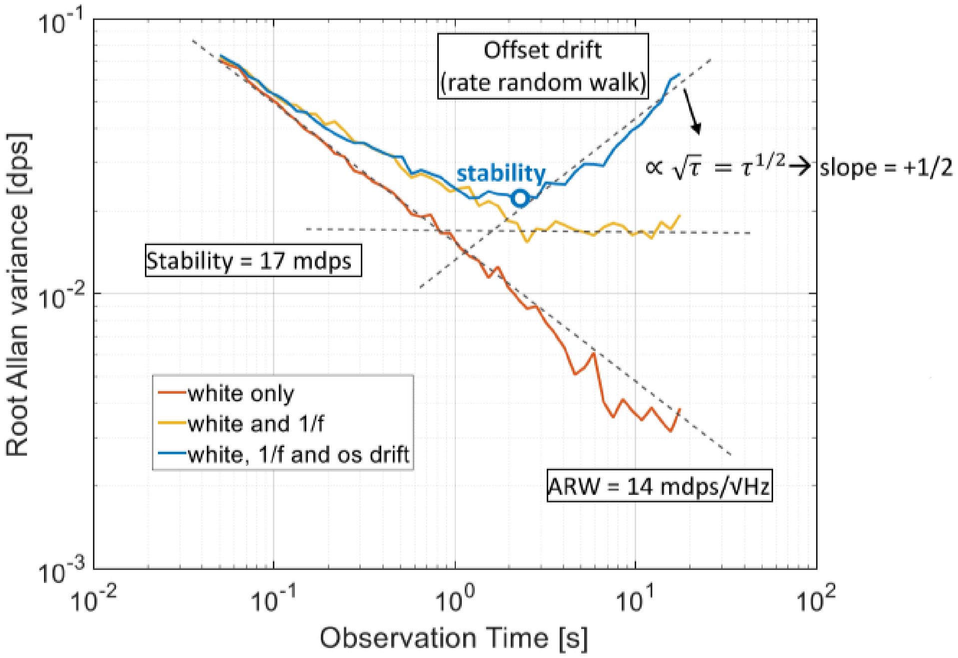
If we want to evaluate of 1/f noise coefficient from an Allan variance plot:

$$\sigma^2_{AV,\Omega}(\tau) = 2\alpha_n \log(2) \quad \sigma_{AV,\Omega}(\tau) = \sqrt{2\alpha_n \log(2)}$$

7.3.3 Offset drifts

Offset drifts are critical for navigation as they slowly but **unpredictably** affect the measurement for long observation times.

$$\sigma^2_{AV,\Omega}(\tau) = \beta_n \tau \qquad \sigma_{AV,\Omega}(\tau) = \sqrt{\beta_n \tau}$$



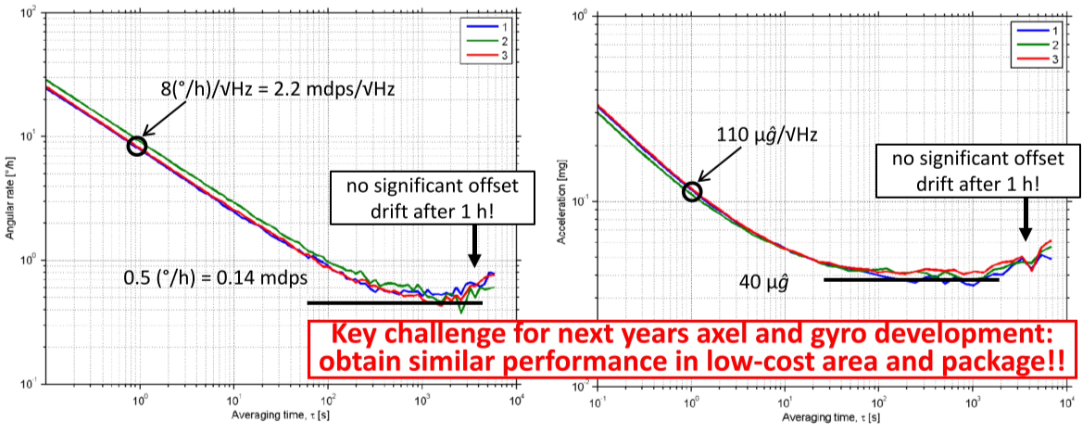
Best IMUs have nowadays ARW (angular random walk) in the order of few $mdps/\sqrt{Hz}$ or fractions thereof. Stability is usually in the order of fractions of mdps. RRW (*Rate Random Walk*) needs to be compensated possibly by design (stability vs T) but also via post-acquisition digital compensation.[26]

---

[26] Note that, since ARW refers to the angle and RRW refers to the rate, the former is the integral of the latter.

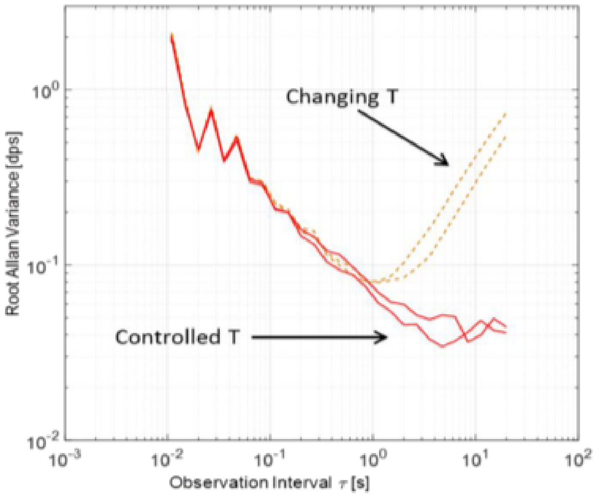
7.3.4 How to use an AV graph

If we look only for low noise in our device (e.g. no need to integrate), look for an AV with the **lowest ARW**. Whereas, if we look for high stability over a certain observation interval, look for the **lowest RAV at that specific observation time**! Note that this time may depend on the application and on the possibility of online offset recalibration.



AV is a method of representing root mean square (**RMS**) random drift error as a function of averaging time (IEEE Std 952-1977). It is relatively simple to compute, to interpret and to understand. The AV method can be used to determine the characteristics of the underlying random processes that give rise to the noise data:

- achievable noise density limit (white region) and ARW;
- minimum achievable rms noise (look at the minimum plateau);
- impact of offset drifts (RRW) due e.g. to temperature changes, vibrations and any other disturbances.



8 CMOS IMAGE SENSORS

Like MEMS, **CMOS image sensors** (**CIS**) are based on measuring the charge induced by a variation in the quantity to measure (light). Also these sensors are systems formed by the sensor itself, the electronics, and «auxiliary» multiphysics system components. If with MEMS, we have e.g. seen the role of packaging and the effects of the environment (thermal) and process, for CMOS sensors, we have a strong system-level interaction with optical elements (lenses, filters) which are part of the system, and data processing. One important point to remark is that: while for MEMS sensors the physical quantities we measured (acceleration, angular rate, magnetic field etc. . .) have a well-known definition, for CMOS image sensors, physical quantities to measure are: **light intensity** as perceived by human eyes and **light color**.

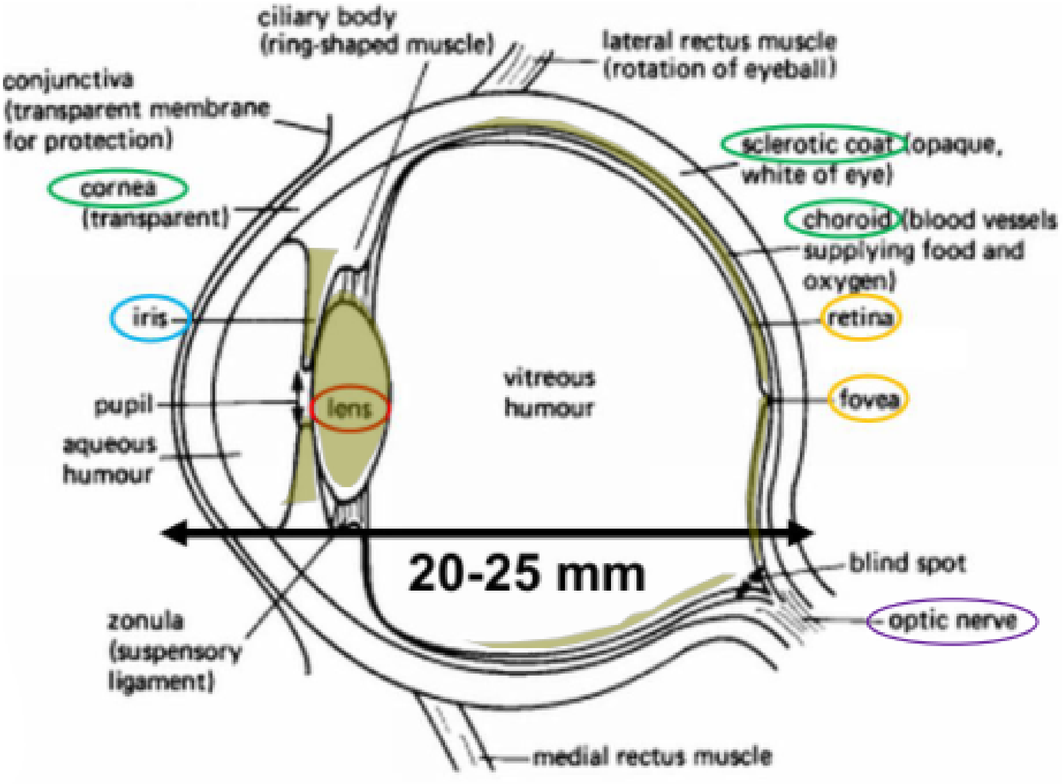
8.1 The human visual system

Why do we need to study the way the **human eye** records a scene?
On one side, because often it is a human eye that judges the result of an image acquisition. Color (physical) definition is itself based on human vision. On the other side, the human eye is a nice micro-system (µm-size receptors) and we learn and are inspired from it in designing digital imaging systems.
It is hence appropriate to develop a rudimental knowledge of human vision as a first step towards CMOS image sensors classes.

8.1.1 The human vision at glance

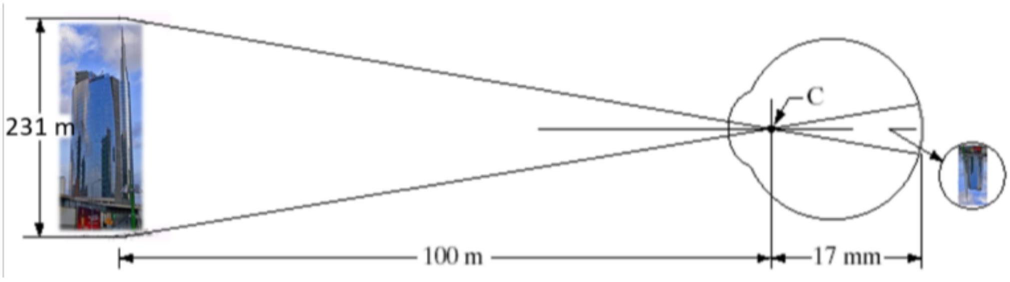
The five major system components of the human eye are:

- **Membranes:** transparent to let the light in and opaque to avoid backscatter within the optical globe.
- **Iris:** part of the choroid, it contracts or expands to control the amount of light entering the eye. The opening (**pupil**) varies in diameter between 1 and 8 mm, this gives roughly a 64:1 ratio of *"programmable"* brightness control.
- **Lens:** its major function is to **focus the image** on the retinal plane. Its shape is controlled by ciliary body muscles: it flattens to focus far objects and becomes thicker to focus near objects (it is basically a flexible device that adjusts focus on demand).
- **Retina:**: it covers the whole inner posterior portion of the eye, where light rays are imaged. The structured vision is obtained from photoreceptors spread on the retina (our "image sensor"). The receptors density is the highest in the central region (**Fovea**).
- **Optic nerve:** brings stimuli from the retina to the brain, where elaboration takes place.



The retina size ($\sim 25\ mm$ diameter) is not so different from sensor size of a full-frame camera ($24 \times 36\ mm^2$). However the **retina is curved** along the back surface of the eyeball, which gives advantages as the edges of the retina are all at (roughly) the same distance from the lens center. The retina has also more receptors ($\sim 80 - 150\ M$) than a common high-end camera ($12 \div 40\ M$) and we have two of such image sensors. Receptors generate electric stimuli (state transition of *retinene* molecule). Such "pixels" are also better distributed: they are denser in the central region (few millions in the 1.5 mm-diameter fovea) where the attention of the visual system is concentrated. Color as well is better discriminated in the central rather than in the peripheral vision.

The distance between the lens center and the retina (eye focal length) can vary thanks to ciliary muscles from approximately 17 mm to 14 mm. From a focusing distance of $\gtrsim 4\ m$, the lens exhibits its lowest refractive power (**most flattened**). With the dimensions in the figure below, we can easily find that the size of the tower on the retina is de-magnified to about 39 mm.



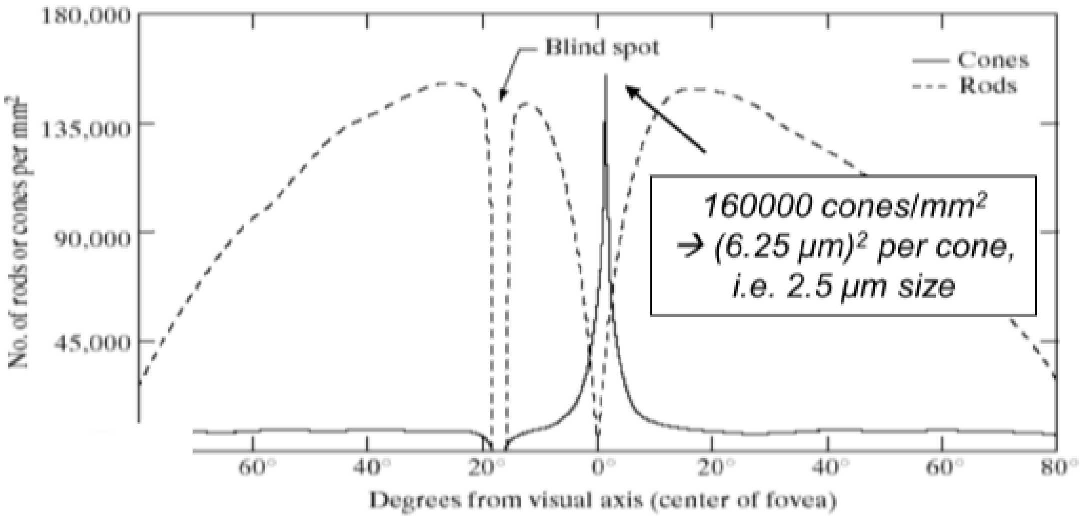
8.1.2 The photoreceptors

Photoreceptors are the *«photodiodes»* of our visual system. Pattern vision (recognition of structures in a scene) is allowed by the distribution of such discrete light receptors on the retinal surface (3D vision is given by 2 eyes). The retina features two types of receptors:

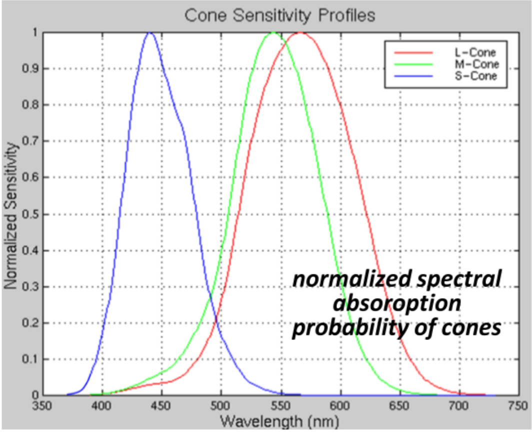
- **Cones: Photopic (bright) vision**
  Three kinds of cones, with different spectral response, allow color discrimination (by the brain) during intense illumination. The cones in each eye number between 5 and 7 million, are mostly located in the central retinal portion (fovea) and each one of them is connected to its own nerve end.

- **Rods: Scotopic (dark) vision**
  One spectral type of photoreceptor which is active at low-light levels: therefore we cannot discriminate colors in low-light to dark conditions. Rods number around 100 million and they are located in the peripheral region of the eye. Finally, more than one rod is connected to a single nerve end (concept of binning).

In the distribution below we can note a blind spot per eye (corresponding to the optical nerve termination). This is however in the peripheral area and it is compensated by stereoscopic vision and brain elaboration. To check your poor rods density in the central part of the retina try to look at a faint star in the night sky.



On-average values: 65.3% of the cones contain **L** (*“red”*), 33.1% contain **M** (*"green"*), 1.6% contain **S** (*"blue“*) photo-pigments.
The cones are actually named after their sensitivity (probability of absorption) in the short (S), medium (M) or long (L) wavelength range. The S pigment has a peak response at about 445 nm. It is insensitive (i.e. almost zero probability of light absorption) to wavelength longer than about 520 nm. The other two types of pigments (M and L) are maximally sensitive to 535 nm and 575 nm respectively. Both respond over almost the whole visual range.



Putting together the information on the spectral absorption probability of each cone and their typical numerical distribution in the eye, one gets an overall responses from all cones (or a portion with $N_{cones} \gg 1$), that looks like the ones depicted below.
It is generally assumed that the S-cone peak is 20 times lower than the M-cone peak (33.1%/1.6%), and that the L-cone peak is twice the M-cone peak (65.3%/33.1%).
The weighted sum of cones spectral sensitivities gives the overall sensitivity to photons during daylight vision for the human eye – named **photopic curve V(λ)**:

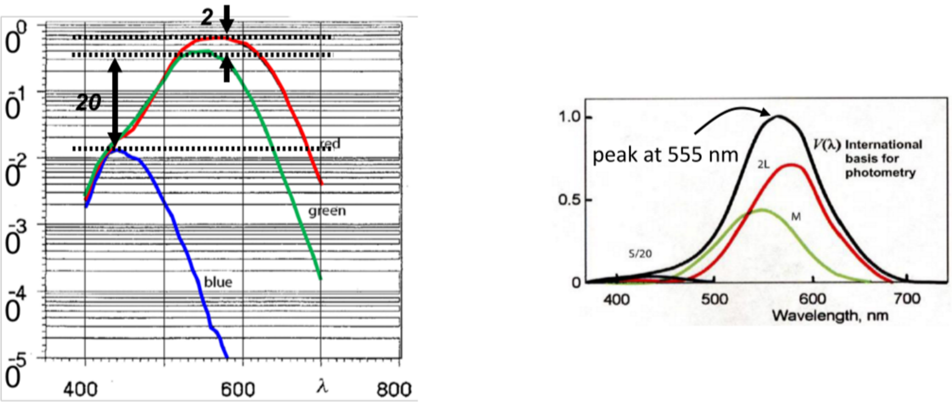


Figure 48: **Right:** spectral absorption probability **Left:** Photopic curve

How is color perceived by the brain?
Color results from the interaction between radiation spectrum, the cones, the nervous system and the brain. Cones have different probability of absorbing photons at a certain wavelength, correlated to their spectral distribution. The signal from each cone is related to the number of absorbed photons, but not to their wavelength (just as in a quantum photodetector): *1 absorbed photon = 1 nerve stimulus*. A single photoreceptor does not give information on the spectral distribution of the radiation. The "color" sensation is obtained from the combination of signals measured by the three cone types: **L**, **M** and **S**. This indicates that color (as perceived by a human visual system) is a three-dimensional quantity. It is thus interesting to know the way these signals are processed by the brain, to later in the course define suitable color spaces.

8.2 Digital Imaging pipeline

We can thus write a model for the three stimuli generated by a certain light power spectral density s(λ) impinging on a human eye:

$$\boldsymbol{S} = \int_{400}^{700} s(\lambda) \cdot r(\lambda) \cdot \cancel{t(\lambda)} \cdot \boldsymbol{s}_{\text{cone}} (\lambda) d\lambda \sim \int_{400}^{700} s(\lambda) r(\lambda) \boldsymbol{s}_{\text{cone}} (\lambda) d\lambda$$

$$\boldsymbol{M} = \int_{400}^{700} s(\lambda) \cdot r(\lambda) \cdot \cancel{t(\lambda)} \cdot \boldsymbol{m}_{\text{cone}} (\lambda) d\lambda \sim \int_{400}^{700} s(\lambda) r(\lambda) \boldsymbol{m}_{\text{cone}} (\lambda) d\lambda$$

$$\boldsymbol{L} = \int_{400}^{700} s(\lambda) \cdot r(\lambda) \cdot \cancel{t(\lambda)} \cdot \boldsymbol{l}_{\text{cone}} (\lambda) d\lambda \sim \int_{400}^{700} s(\lambda) r(\lambda) \boldsymbol{l}_{\text{cone}} (\lambda) d\lambda$$

Let us try to build a camera following the same principle. We will need, or we will rely on:

- a focusing system (lenses);
- microsensors that imitate the photoreceptors (photodiodes);
- 3 types of spectral responses (color filters implementing the camera color space);
- post processing of acquired information (electronics).

The color coordinates of an object (e.g. its blue coordinate below) are the result of integral functions that include the illuminant power spectral density.

$$\boldsymbol{B} = \int_{400}^{700} \overset{\text{illuminant}}{s(\lambda)} \cdot \underset{\text{object reflectance}}{r(\lambda)} \cdot \overset{\text{optics transmittance (typically ~ even on the visible range)}}{\cancel{t(\lambda)}} \cdot \underset{\text{blue pixel response}}{\boldsymbol{b}(\boldsymbol{\lambda})} d\lambda \sim \int_{400}^{700} s(\lambda) r(\lambda) \boldsymbol{b}(\boldsymbol{\lambda}) d\lambda$$

The light reflected from an object depends as much on the source illuminant s(λ) as it does on the object reflectance r(λ). Therefore, there is no meaning in giving color coordinates without specifying the illuminant. Indeed, the RGB color coordinates depend on the light source and on the specific imaging system: for the same camera, the same object has different coordinates under different sources. So, apparently, if we shine different light spectra on the same object we should perceive different colors. However the human vision does adapt! Indeed, you can see a **distinct white** patch in the color chart of each of the images below, rendered as under different illuminant despite colors corresponding to these whites are totally different!
A photodiode does not have such an automatic adaptation and we will see that it is mandatory to take into account this for a correct rendering of images in a post processing phase!

8.3 Geometric Optics

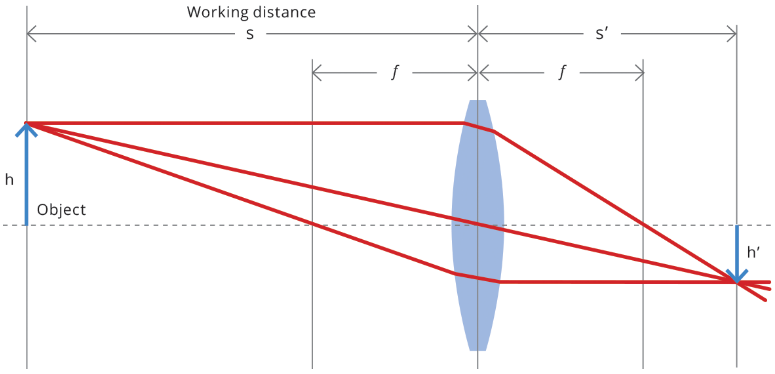
Optical elements are usually spherical **lenses**. To describe their behaviour, we do some initial approximations:

- light is described by **rays**;
- rays travel in **straight lines** until they are deflected by an optical element;
- deflection of rays is described through the Snell's law;
- **rays are paraxial**: their angle with the optical axis is small and they lie close to the optical axis (rays make a small incident angle with respect to the optical surface).

Snell's law describes reflection and refraction at the boundary between media with different refractive index:

- incident, reflected and refracted rays belong to a single plane;
- the incident angle α is the same as the reflected one;
- the refraction angle β is related to α through the Snell's law: $\mathbf{n_1 sin}(\alpha) = \mathbf{n_2 sin}(\beta)$

The general equation of a thin lens in paraxial approximation relates the distance of the object (s_1) and of the image (s_2) from the lens center.



We can thus relate object and image distance from the lens to geometrical (s_0: lens half width on optical axis) and optical (n: lens refraction index) parameters:

$$\left(\frac{1}{s_1}+\frac{n}{s_0}\right)=(n-1)\frac{1}{R_1} \qquad \left(\frac{1}{s_2}-\frac{n}{s_0}\right)=(n-1)\frac{-1}{R_2}$$

$$\left(\frac{1}{s_1}+\frac{1}{s_2}\right)=(n-1)\left(\frac{1}{R_1}-\frac{1}{R_2}\right)$$

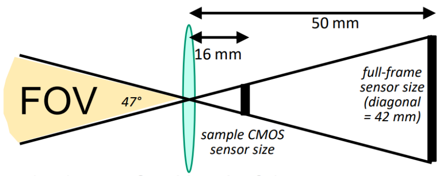
The equation shown below, called **thin lens equation**, is obtained by defining the **focal length f** as a single parameter that depends on the lens geometry (R_1, R_2) and material (n) through the **lensmaker's equation**:

$$\frac{1}{s_1}+\frac{1}{s_2}=\frac{1}{f} \qquad \text{with} \qquad \frac{1}{f}=(n-1)\left[\frac{1}{R_1}-\frac{1}{R_2}\right]$$

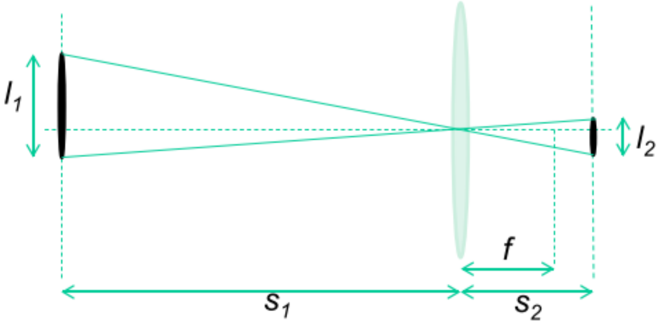
The thin lens equation relates the object and image distance from the lens (s_1 and s_2) to the focal length f of the lens. In typical digital imaging systems $s_1 \gg s_2$ and images are formed close to the focal length ($s_2 \sim f$).
A typical focal length shall match the field of view (**FOV**) of normal attention of human eyes (fovea). Thus, for full-frame format, this FOV ($\approx 50°$) was matched by $\sim 50\ mm$ focal lengths. Focal lengths lower than this were assumed as **wide-angle lenses** (e.g. 28 mm), whereas focal lengths larger than this were assumed **teleobjectives** (e.g. 300 mm). When digital photography era begun, the different size of sensors (depending on manufacturer and application) made the true focal length meaningless to customers (if sensor size was not specificied). So, digital system are generally characterized by an *«equivalent»* (to the full frame) focal length, as shown in the example.

By definition, the **magnification factor m** represents the ratio of a dimension l_2 in the image with respect to the dimension l_1 in the object:

$$m = \frac{l_2}{l_1}$$



*Example: the **true focal lengh** of the sensor is 16 mm. However, given this distance, the **field of view** of the sensor forms an angle that, **using a full-frame** sensor, would correspond to a **50 mm focal length**. The system is thus **sold as a 50 mm (equivalent) camera**.*



Through simple geometry, one can verify that the magnification factor is equal to the ratio of the distances from the lens:

$$m = \frac{s_2}{s_1} \sim \frac{f}{s_1}$$

Obviously, area is magnified by a factor m^2. This is useful to calculate the area (of the captured scene) from which the photons captured by each pixel belong. In imaging systems, it is always $m \ll 1$.
As we cannot use a variable-shape lens, more lenses with variable distance can be combined to obtain a variable focal length:

- Adding a positive (convex) lens decreases the focal length.
- Adding a negative (concave) lens increases the focal length.
- Moving the relative distance between lenses changes the focal length.

Systems of multiple lenses can be approximated through a lens of equivalent focal length. We thus hold this assumption and proceed assuming that a lens system is always described through a single f value.

8.4 Aberrations

Aberrations are phenomena that cause images to be **imperfect** replicas of objects: points do not appear as points, planes do not appear as planes and the image and the object do not have the same shape. Since aberrations worsen the resolution, they will help us to decide which is the minimum pixel size that is needed. Aberrations become evident when two phenomena are considered:

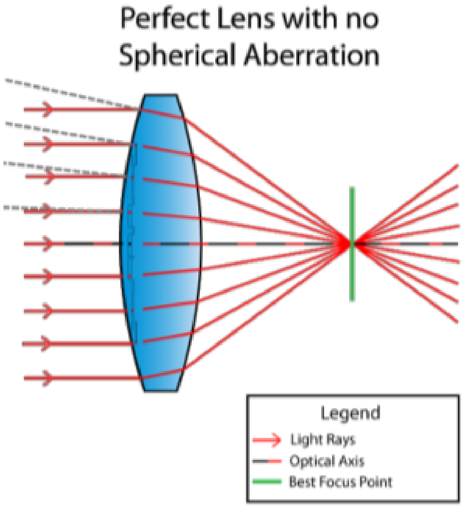
- errors due to first order **paraxial approximation**: (known as 3rd order or spherical aberrations, as they appear as soon as the 3rd-order term is considered in a spherical lens).

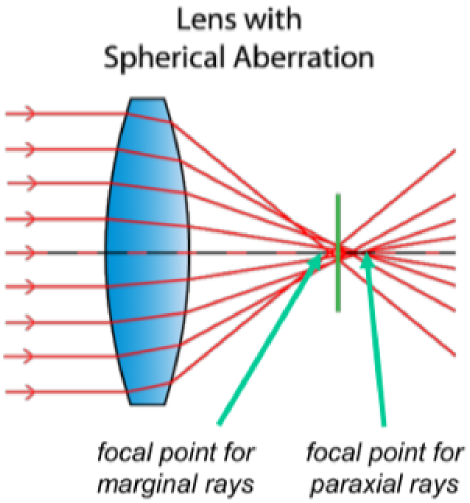
$$sin(\alpha) = \alpha - \frac{\alpha^3}{3!} + \frac{\alpha^5}{5!} - \frac{\alpha^7}{7!} + \cdots$$

- the fact that radiation is not monochromatic and the refractive index (e.g. of glass) is a function of the wavelength (known as **chromatic aberrations**): $\mathbf{n_1} = \mathbf{n_1}(\lambda)$

Aberrations can be mitigated: via **hardware** (lens shape and material), in **operation** (closing the aperture) or via **software** (their math origin is known).

The **Spherical aberrations** occur as spherical surfaces are not suited to make "ideal" lenses, though easy and cheap. Consider light coming from a point like object at infinite distance, and an ideal spherical lens. Aberrations are due to increased refraction of rays when striking a spherical lens near its edge, compared to those striking close to the center. Marginal and paraxial rays have thus different focus points (at large angles α, refraction power is larger). The point of best focus with the smallest **"disk of least confusion"** is illustrated as the thick green line.





Spherical aberration is most pronounced when the diaphragm of the lens is wide open (maximum aperture). Reducing the lens aperture even by a single or two stops dramatically reduces spherical aberrations, because aperture blades block the edges of the spherical lens (the resolution improves). As a secondary effect, it can be observed that if the aperture is closed and blocks the marginal rays, the best focus point shifts to the right.

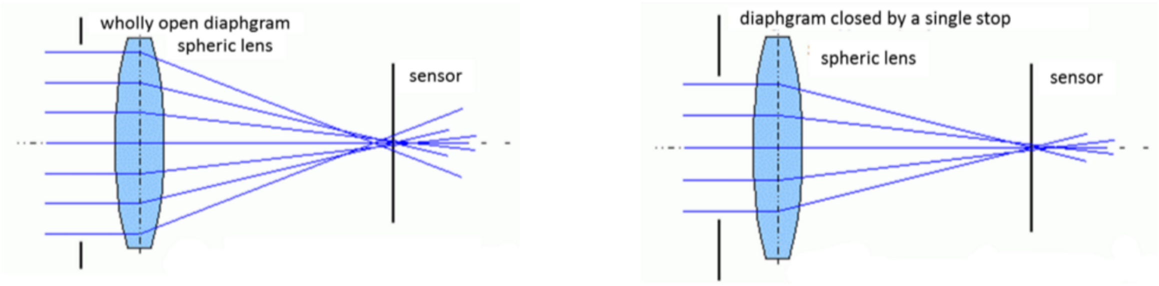
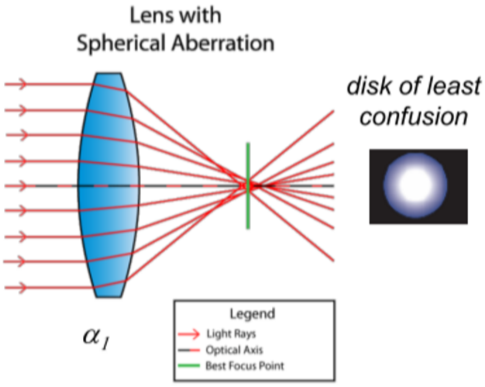
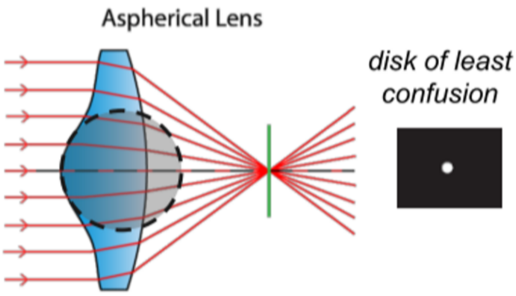


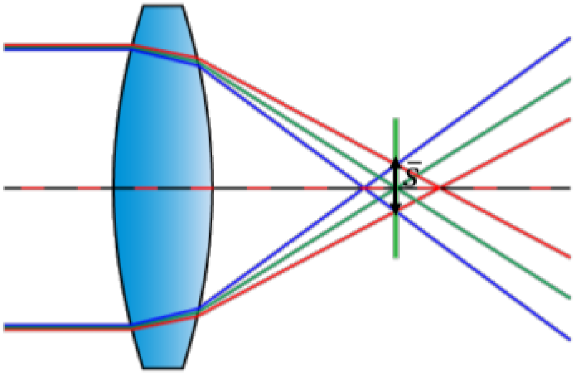
Figure 50: **Right:** Larger aperture **Left:** Reduced aperture

Spherical aberrations can be avoided using **aspheric lenses**. An aspheric lens is a lens whose surface profiles at the edges are not portions of a sphere or cylinder. It is typically more complicated to build and therefore more expensive. Plastic is preferable with respect to glass in this sense. The definition of the surface profile takes into account third-order terms and reduce by a large amount the effects of aberrations.





The longitudinal (or axial) **chromatic aberrations** is due to fact that the focal length is different for any different wavelength. In particular, the refractive index n is larger for short wavelengths than for long ones. Therefore the refractive power is higher for blue light which is focused closer to the lens than red light.



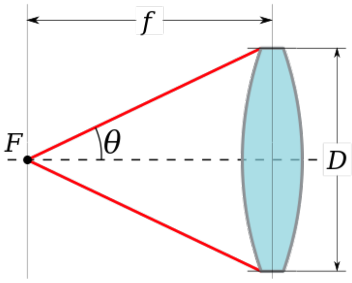
A non-monochromatic point-like object is not imaged as a point: resolution worsening!
Means used to correct chromatic aberrations are:

- use of low dispersion (LD) glasses (the refraction index changes poorly with wavelength);
- use of achromatic doublets: they are made of two lenses of different glass, with different concavity, index of refraction and dispersion. The first element is a convex lens with low dispersion and high refractive power. The other lens is concave with higher dispersion and lower refractive power. The result is a positive lens where the two dispersions are compensated.

8.5 F# number

The $F_{\#}$ **number** of a lens is defined as the ratio between the focal length and the lens diameter (or aperture) D:

$$F_{\#} = \frac{f}{D}$$



It therefore defines also the maximum aperture cone of the focused rays:

$$\theta = \arctan\left(\frac{D}{2f}\right) = \arctan\left(\frac{1}{2F_{\#}}\right)$$

The F number is an important parameter in the optics of a camera as it affects gathered light signal (and SNR) and spatial resolution.
Lenses of high-end cameras usually have adjustable aperture. A lens is characterized by the $F_{\#}$ corresponding to the widest allowed opening. A lens with a **smaller** $F_{\#}$ provides **brighter images** (more light in). Digital still cameras (DSC) lenses use a standard F-stop scale, which includes numbers corresponding to the sequence of powers of the square root of 2.

8.6 Diffraction

As we have propagating light (EM waves) entering an aperture, we need to consider diffraction. Its effects are most evident when the aperture size is of the same order of magnitude as the wavelength λ. Therefore, in image acquisition systems, where λ is typically 4 or 5 orders of magnitude smaller than apertures (e.g. 500nm vs 5cm), images are not notably distorted by diffraction, but it unavoidably affects the resolution.
In the **Huygens model**, every point of a wave-front is a source of secondary waves. Subsequent fronts are obtained as an envelop of those waves, thus the intensity distribution as a function of the angle can be obtained from multiple interferences of secondary waves. The first minimum falls at:

$$sin(\theta) \sim \theta = \frac{\lambda}{D} \qquad \text{for infinite plane aperture}$$

$$sin(\theta) \sim \theta = 1.22\frac{\lambda}{D} \qquad \text{for circular aperture}$$

The lens circular aperture diffracts light. We are interested in evaluating how a far, on-axis, a monochromatic point-like source imaged on the sensor can be.
The first Airy disk diameter turns out to be:

$$d_{\text{Airy}} \ = f * tan(2.44\frac{\lambda}{D}) \simeq 2.44\frac{\lambda}{D}f = 2.44 \cdot \lambda \cdot F\#$$

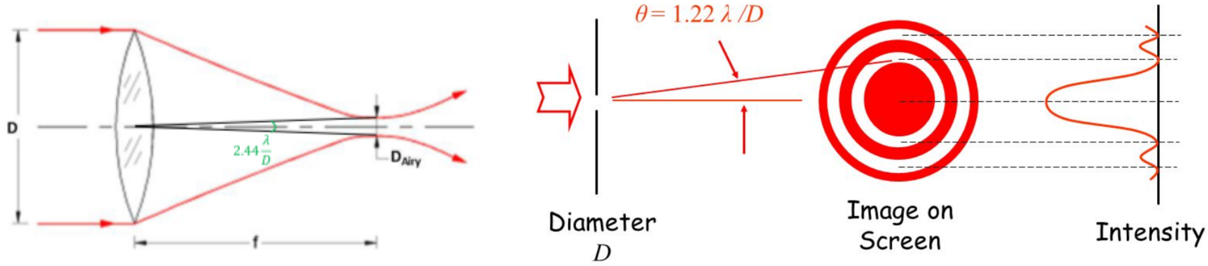


Figure 51: Airy disk formed by the diffraction of light impinging on a circular aperture.

Therefore, the higher is the F#, the bigger is the spot size and worse the resolution. As for a designer prospective, there is almost no use in designing pixels smaller than the spot size.

8.7 Pixel optics guidelines

1. **How much light or how many photons reach each pixel of the silicon sensor?**
   Calculation of the photons/s impinging on a single pixel:
   (a) start with source power spectral density in W/nm;
   (b) calculate how this is (on average) reflected by an object, per unit area of reflection and solid angle $[\frac{W}{nm \cdot m^2 \cdot sr}]$;
   (c) calculate the solid angle from which the lens is seen by a point in my object;
   $$\Omega_{lens} = \frac{\pi (D_{lens}/2)^2}{s_1^2}$$
   (d) calculate the area that, in the scene, corresponds to a pixel (through magnification);
   $$A_{scene} = \frac{A_{pixel}}{m^2} \approx \frac{A_{pixel}}{f^2} s_1^2$$
   (e) the **light impinging on the pixel area** is found by multiplying the quantity obtained at point (b), by the two formulas found at points (c) and (d).

   We will see that photon flux (so, the pixel photocurrent) is thus proportional to pixel area and to the $F_{\#}$ number squared inverse:
   $$\Phi_{pixel} \propto \frac{A_{pixel}}{F_{\#}^2}$$

2. **How spatial resolution is limited by the optics?**
   Minimum diffraction disk occurs at short λ and wide-open lens. However, for wide-open lenses, resolution is limited by aberrations. The $F_{\#}$ at which best optical resolution is achieved is typically two stops below the widest aperture, i.e. $F_{\#} = 4 - 5.6$. In this case diffraction-limited resolution has a spot:
   $$d_{Airy} = 2.44 \cdot 400nm \cdot 4 = 3.9\ \mu m$$
   Therefore, in order not to worsen the resolution with sensor spatial sampling, the pixel size should be ranging between 2 μm and 6 μm.

Mobile imaging vs pro cameras. A smaller sensor allows, in mobile photography, to obtain the same field of view of digital cameras. To fit within a small thickness, high refractive power is needed, leading to relatively large aberrations, hence lower resolution. Moreover, the focal length adjustment can be done by combining acquisition from multiple cameras.
However, for the same number of pixels, a smaller sensor size corresponds to a smaller pixel size (often smaller than optic limitations): the amount of light per pixel (for the same exposure time) is thus small and this will affect the pixel performance.
In general we conclude that native performance of mobile imaging are unavoidably lower than digital cameras. Such a difference is partially compensated by the heavy role of digital post processing of mobile images and simply by the fact that most images are viewed only through a relatively small mobile phone screen.

9 CMOS PASSIVE/ACTIVE PIXEL SENSORS

9.1 Photogeneration in the simplest PN junction

A solid-state image sensor is a semiconductor (**Silicon**) device that converts an image formed by a lens into electronic (digital) signals. The image is acquired through a matrix of elementary picture elements (**pixels**). This conversion is nearly linear and is governed by three main parameters:

1. overall quantum efficiency (from photon to current);
2. integration time (from current to charge);
3. conversion gain (from charge to voltage).

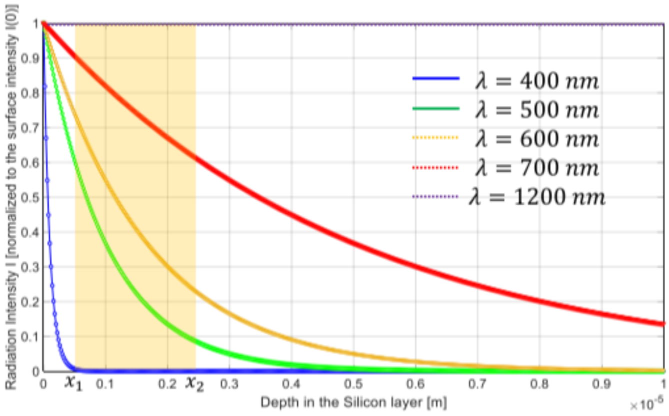
9.1.1 Photon absorption

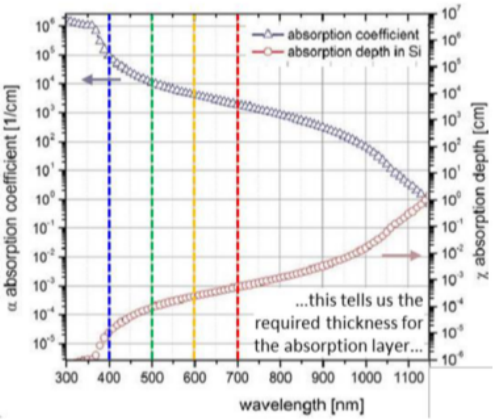
Everything starts with photon absorption that occurs (as always for light and particles) through an exponential decay formula. Therefore, the intensity of light propagating through a medium is described by:

$$I(x) = I(0)e^{-\alpha(\lambda)x}$$

The percentage of absorbed light in the portion of material between depth x_1 and depth x_2 will be:

$$\frac{I(0)e^{-\alpha(\lambda)x_1} - I(0)e^{-\alpha(\lambda)x_2}}{I(0)} = e^{-\alpha(\lambda)x_1} - e^{-\alpha(\lambda)x_2}$$





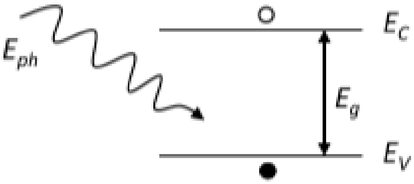
"Blue" photons are all absorbed within few hundred nm, "red" photons are absorbed within more than 10 µm. This tells us the required thickness of the active layer[27] (at least few µm).

---

[27] The active layer is the device zone in which we want electron-hole couplets generation through light absorption.

9.1.2 Charge generation

Absorption occurs when the energy of the incoming photon is **larger** than the energy gap. If this is the case, an electron-hole pair is formed for each absorbed photon, whatever its energy (quantum detector, like the eye).
With a Si gap $E_g = 1.12\ eV$, the corresponding cut-off wavelength turns out to be, using Plank's law, $\lambda_{cut-off} = \frac{hc}{E_g} = 1100\ nm$, which falls in the near infrared (NIR) range.



After charge is generated, we need to understand how it is collected. In the simplest case, the active area of the pixel is a reversely biased PN junction with typical doping values around $N_D \sim 10^{20}\ cm^{-3}$ for the N^+ implant and $N_A = 10^{15}\ cm^{-3}$ for uniform P-type epitaxial layer.
A few important parameters to understand the true operation of the image sensor are:

- **depletion layer width** (e.g. at a typical reverse voltage $V_R = 3.3\ V$):

$$x_{\text{depl}} = \sqrt{\frac{2\varepsilon_0\varepsilon_{Si}\left(V_R + V_{bi}\right)}{qN_A}} = 1.5\mu m$$

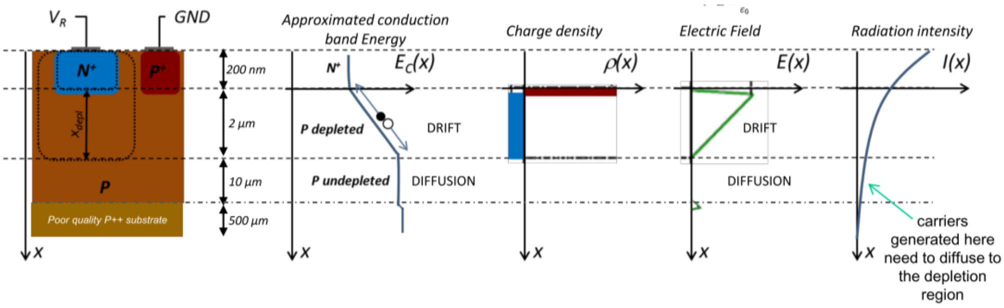
- **diffusion length** at $\tau_n \sim 1 - 5\ \mu s$, $D_n \sim 10\ cm^2/s$ (epitaxial layer):

$$x_{\text{diff}} = \sqrt{D_n\tau_n} > 10\mu m$$

- **diffusion length** at $\tau_p < 0.1\ ns$, $D_p \sim 3\ cm^2/s$ (surface layer):

$$x_{\text{diff}} = \sqrt{D_p\tau_p} < 50\ \text{nm}$$

It is therefore evident that charge collection occurs both by drift and by diffusion (in the lowly doped layer). The percentage of charge collected by diffusion increases with the wavelength.

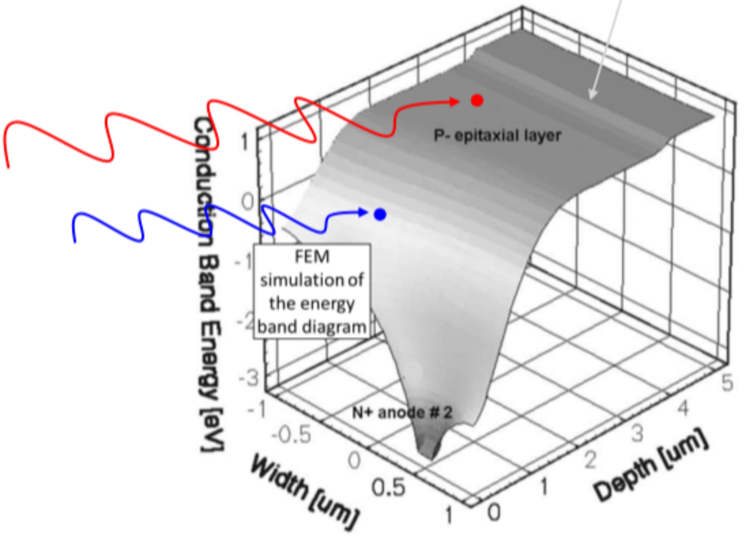


(note: figure not to scale. Typical implant depth is about 200 nm)

Apparently, we would like the largest possible τ_n, but this would create too large diffusion lengths and spatial resolution worsening due to "red" generated photoelectrons.
High energy (short wavelength) photons are collected mostly in the depleted region. **Collection** occurs in this case by drift: fast (high electric field) and with a well defined trajectory.
Low energy (long wavelength) photons are collected mostly by diffusion:

- initial random path, due to absence of carrier gradient and electric field;
- Once captured by the electric field, collection becomes similar to above;
- overall, longer collection time;
- charge collection at the electrodes per unit time (photocurrent i_{ph}).

note how the small barrier created by the P++ substrate prevents photogenerated electrons to be lost in the substrate



holes not shown

9.1.3 Quantum efficiency and responsivity

The ratio of collected electrons over incident photons is called **quantum efficiency**.

$$\boldsymbol{\eta(\lambda)} = \frac{\textit{collected electrons}}{\textit{incident photons}} = \frac{I_0 e^{-\alpha(\lambda)x_1} - I_0 e^{-\alpha(\lambda)x_2}}{I_0} T_{Si}(\lambda) \approx \left(\boldsymbol{e}^{-\alpha(\lambda)x_1} - \boldsymbol{e}^{-\alpha(\lambda)x_2}\right) \boldsymbol{T}_{Si}(\boldsymbol{\lambda})$$

Usually η is < 1 due to:

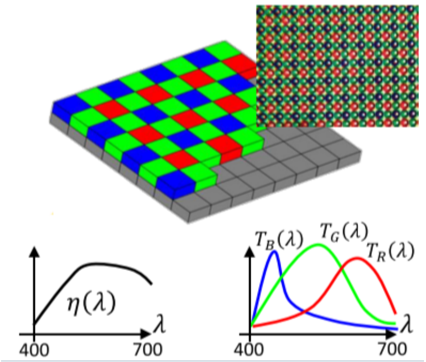
- surface recombination at short λ
- uncollected carriers at long λ
- trasmittance $T_{Si}(\lambda) < 1$ at Si-air interface.

A more commonly used parameter (because it's easily measurable) is the **responsivity**, which is defined as the ratio of the output current to the input optical power.

$$\boldsymbol{\mathcal{R}(\lambda)} = \frac{i_{ph}}{P_{in}} = \frac{q \cdot collected\ electrons\ per\ sec}{E_{ph} \cdot incident\ photons\ per\ sec} = \frac{q}{E_{ph}}\eta(\lambda) = \boldsymbol{\frac{q\lambda}{hc}\eta(\lambda)}$$

Si detectors are quantum detectors: every photon generates an electron-hole pair, regardless of the energy (or λ) of the absorbed radiation (energy in excess is transferred to lattice phonons). Thus, a matrix of identical photo-sites (like rods) cannot distinguish the color of the impinging radiation, and it is just a mono-chrome sensor. Human vision under medium/high light is based on 3 functions of the wavelength (cones responses). To retrieve color information, a camera should implement as well 3 different functions of the incoming wavelength. There are different approaches to this, the most common being the use of a mosaic of **Color Filter Arrays** (**CFA**), where µ-filters are deposited on top of the Silicon layer. The overall quantum efficiency of each pixel becomes the product of the Silicon quantum efficiency and the filter transmittance.

$$\boldsymbol{\eta}(\lambda) = \left(e^{-\alpha(\lambda)x_1} - e^{-\alpha(\lambda)x_2}\right)\boldsymbol{T}(\lambda)\boldsymbol{T}_{FIL}(\lambda)$$



9.2 Signal and noise

9.2.1 Signal generation

The photogenerated current is calculated as:

$$i_{ph}(\lambda) = \int_{Si} PSD(\lambda)\mathcal{R}(\lambda)d\lambda = \int_{Si} \frac{dP(\lambda)}{d\lambda}\mathcal{R}(\lambda)d\lambda \approx P(\bar{\lambda}) \cdot \mathcal{R}(\bar{\lambda})$$

From a radiation intensity I $[W/cm^2]$ (or a radiation power P $[W]$) or from a photon flux per unit area Φ_{ph} $[\frac{ph}{s \cdot cm^2}]$, the photogenerated current can be simply evaluated as:

$$\boldsymbol{i}_{ph}(\bar{\lambda}) = P \cdot \mathcal{R}(\bar{\lambda}) = I \cdot A_{\text{pixel}} \cdot \mathcal{R}(\bar{\lambda}) = I \cdot A_{\text{pixel}} \cdot \frac{q\bar{\lambda}}{hc}\eta(\bar{\lambda}) = \Phi_{ph} \cdot \boldsymbol{A}_{\text{pixel}} \cdot \boldsymbol{q} \cdot \boldsymbol{\eta}(\bar{\lambda})$$

Typical values for an average photon rate of 10^{16} $\frac{ph}{s \cdot m^2}$ are, e.g., for a compact camera pixel (2 mm) with a 60% quantum efficiency:

$$i_{ph}(\lambda) = 10^{16}\frac{\text{ph}}{\text{m}^2\ \text{s}} \cdot (2\mu\text{m})^2 \cdot 1.610^{-19}\text{C} \cdot 0.6 = 3.8\ fA$$

9.2.2 Dark current offset

The dark current is the photodetector current when no illumination is present. It both represents an offset and noise source. It is generated by various phenomena, but mainly ($T \lesssim 70°C$) by carrier thermal generation in the depletion region and at Si/SiO$_2$ interface. With previous values:

$$i_d = \frac{q\, n_i}{2\tau_n}\, x_{coll}\, A_{pixel} = 0.48\, fA$$

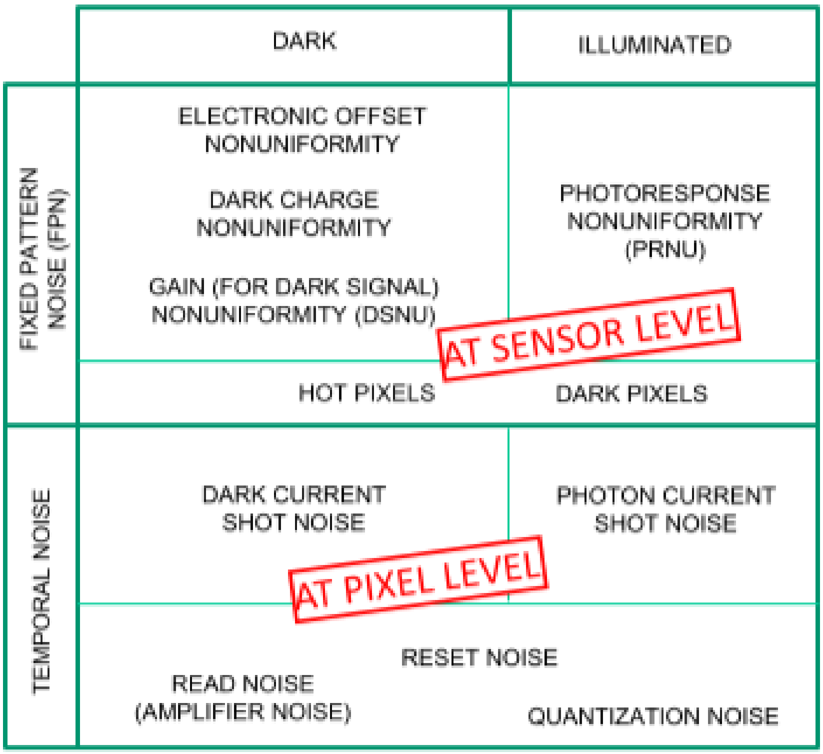
Dark current exponentially rises with T, roughly doubling every 7°C. The intrinsic concentration n_i is a very strong function of temperature. In ultra-low-noise (scientific) imaging systems the detector can be cooled to improve the SNR.

9.2.3 Noise contribution overview

Noise contributions of image sensors are split into two groups:

- noise fluctuating over time is referred to as random or **temporal noise** (shot, thermal, kTC are clear examples...);
- noise appearing in the final image as a result of different gain or offset paths of the different pixels (fluctuating over space) is referred to as spatial or **Fixed-Pattern noise** (**FPN**).

Random temporal noise sets the ultimate limit on signal to noise ratio, while deterministic FPN can be reduced with specific compensation strategies (seen in next paragraphs).
We can sum up these noises in a table:



	DARK	ILLUMINATED
FIXED PATTERN NOISE (FPN)	ELECTRONIC OFFSET NONUNIFORMITY DARK CHARGE NONUNIFORMITY GAIN (FOR DARK SIGNAL) NONUNIFORMITY (DSNU)	PHOTORESPONSE NONUNIFORMITY (PRNU)
	HOT PIXELS	DARK PIXELS
TEMPORAL NOISE	DARK CURRENT SHOT NOISE	PHOTON CURRENT SHOT NOISE
	READ NOISE (AMPLIFIER NOISE) RESET NOISE	QUANTIZATION NOISE

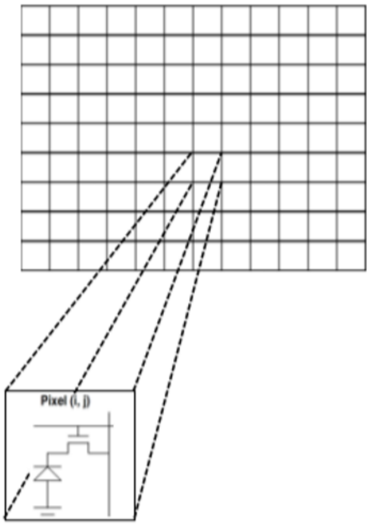
9.2.4 Passive-pixel vs Active-pixel sensor (PPS vs APS)

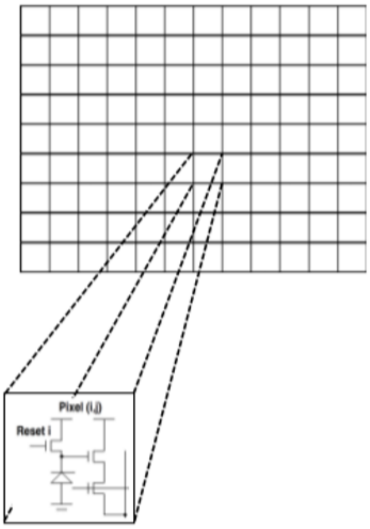
PPS

Each individual pixel is just formed by the **photosensitive element** (e.g. the pn photodiode) and a **selection transistor** (which acts as a switch). Photogenerated charge at pixel level should be transferred to a single charge-to-voltage conversion stage (charge amplifier) at the matrix output.

APS

Each individual pixel is formed by the **photosensitive element** (e.g. the pn photodiode) and a **compact charge to voltage conversion**. Photogenerated charge is thus directly converted into voltage at pixel level, with the minimum required number of transistors (3 or 4, as we will see).





Pros and cons of PPS:

- CCD (charge-coupled device)/PPS readout architecture requires minimal pixel overhead (almost no electronics), making it possible to design very small pixel sizes.
- In PPSs the charge transfer is passive and therefore does not introduce pixel to pixel variations known as fixed-pattern noise (FPN).
- In PPSs, charge transfer readout is serial with limited readout speed.
- CCDs/PPSs are fabricated in specialized technologies solely optimized for charge transfer. The disadvantage of using such specialized technologies is the inability to integrate other camera functions (processing) on the same chip.
- Voltages needed for charge transfer are rather high (e.g. -8/15 V) for fast drift transport.
- A larger pixel area directly requires higher voltage or longer time to hold an acceptable efficiency in transferring the charge, making it difficult to reduce consumption at an acceptable frame rate.

Pros and cons of APS:

- The charge to voltage conversion at pixel level enables a simple voltage "scanning" mechanism, high-speed readout and window-of-interest operations.
- It however introduces gain/offset non-uniformity (FPN, which needs to be compensated).
- CMOS image sensors are mostly fabricated in standard technologies and thus can be readily integrated with other processing and control circuits, enabling analog and digital processing (e.g. FPN compensation!) in the same chip.
- Voltages needed for CMOS operation are usually in the order of 2.5 V to 5 V.
- In CMOS APS, a larger pixel area does not require larger voltages.
- Because of a CMOS sensor X–Y address scanning scheme, only selected pixels consume power at any given time.

9.3 Further considerations on APS

The light and dark signals generated within a pixel are given by **photo and dark currents** lasting a certain amount of time, thus by photo-generated charges:

$$Q_{ph} = \int_0^{t_{int}} i_{ph}(t)\, dt = i_{ph}\, t_{int} \qquad\qquad Q_d = i_d\, t_{int}$$

Note: the image does not change during acquisition, so the photocurrent is constant for every pixel

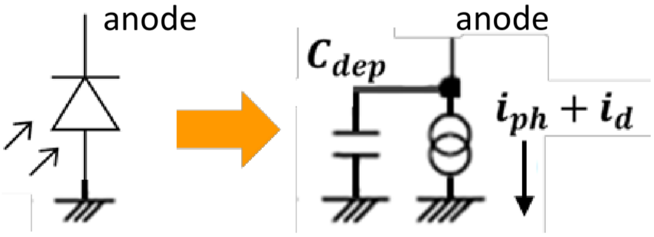
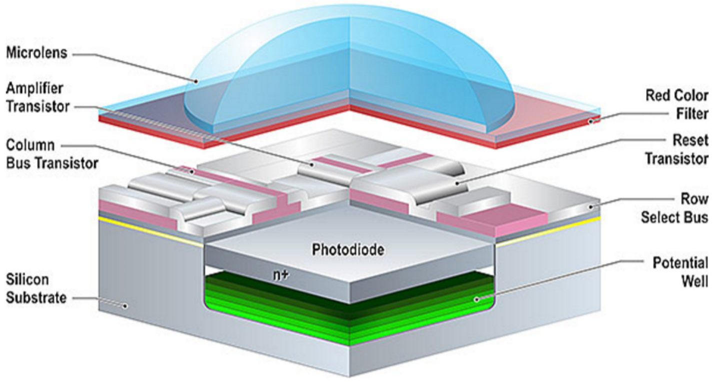


Figure 52: **Right:** diode **Left:** Small signal model

Unlike for passive pixel sensors (PPS), in CMOS active pixel sensors (APS), charge-to-voltage conversion and buffering is operated at pixel level, by a dedicated small (and smart) low-power circuitry (usually a few transistors).

The photodiode is only one part of the active pixel, which includes as well transistors, electrical interconnections, pixel-level µ-lenses and color filters. The mostly adopted CFA (color filter array) configuration is the **GRGB pattern** (G is used twice as it better emulates the photopic curve than B/R, and lets more light in).

Note that we consider the **N region as anode** for the photodiode because of reverse bias.



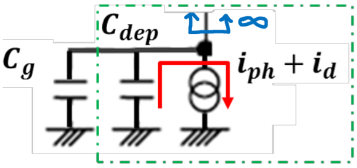
Additional filters (not shown) are placed on top of the whole sensor to:

- **cut-off UV/IR light** (RGB filters let some light pass in the IR range);
- avoid aliasing by "blurring" images to ensure a spatial frequency of the optical signal lower than $f_{\text{Nyq}} = \frac{1}{2 \cdot d_{pixel}}$. This is not needed if diffraction and aberrations already blur the image enough, compared to pixel size (small pixels are used in mobile applications).

We call **Fill Factor** the ratio of the pixel active area over the whole pixel area, including electronics, interconnections and dead Si area. The FF for PPS can be as high as >90%, whereas typical fill factor for the simplest active pixel sensor (APS with 3 transistors) and front-side illumination is in the order of 35% to 45%. The fill factor in APS can be increased by:

- **increasing** photodiode **area** (and thus the overall pixel size);
- **minimizing the electronics** per pixel, by sharing part of the same electronics among different pixels;
- **using microlenses** to focus radiation at pixel level on the active area.

If the anode sees only high impedances, the photocurrent i_{ph} and the dark current i_d are **integrated** over a capacitance that sums the diode depletion capacitance $C_{dep} = \frac{\epsilon_0 \epsilon_{Si} A_{pd}}{x_{dep}}$ and the gate capacitance of one transistor, C_g.



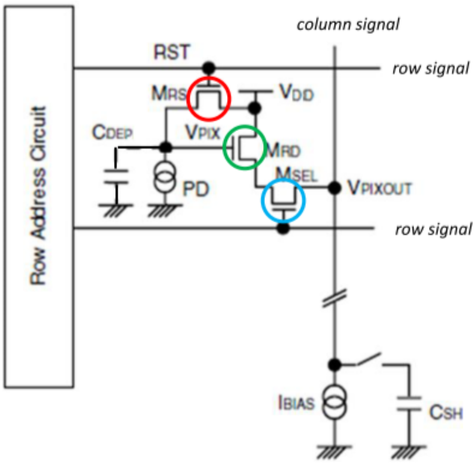
This kind of integration, which occurs directly on the anode capacitance, is called **direct integration**. Instead, if it is done with a Charge Amplifier we have the **indirect integration**. Assuming to first reset the photodiode reverse bias, the voltage across the capacitance decreases during integration. This charge-to-voltage conversion gain, measured in µV/electron, is governed just by one parameter (the **integration capacitance**) and is almost linear:

$$G = \frac{\Delta V_{out}}{N_{el}} = \frac{Q_{ph}}{(C_{dep} + C_g)\, N_{el}} = \frac{q N_{el}}{(C_{dep} + C_g)\, N_{el}} = \frac{q}{(C_{dep} + C_g)}$$

9.4 Three-Transistor topology (APS3T)

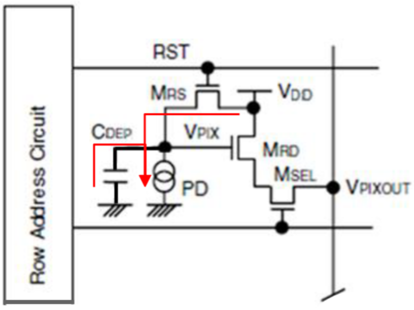
According to what have said, we can get direct integration and fast scan if we guarantee an high input impedance seen by the anode and a low output impedance towards the column bus. The simplest active pixel to readout the photocurrent within a large matrix is formed by the PN junction and three transistors:

- a transistor M_{RD} arranged in a **source follower** configuration to buffer the signal to the output node with low impedance;
- a **reset** transistor M_{RS} whose gate is pulsed on to reset (begin) the pixel operation before a new image acquisition is taken;
- a **row-selection** transistor M_{SEL} that latches the voltage to the column bus when the pixel is to be read (i.e. at the end of the "electronic" exposure).

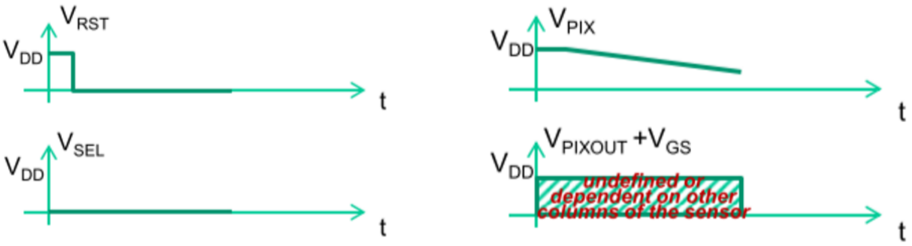


The signal formation could be divided in three parts:

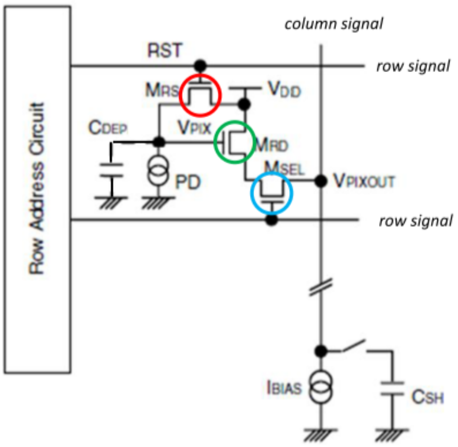
1. **Reset:**
   The voltage V_{PIX} across the diode's capacitance is reset to V_{DD} by closing M_{RS}. The selection transistor is off and the source follower is not biased (does not dissipate current). Any (dark or photo) current flowing through the photodiode contacts is provided by the supply.

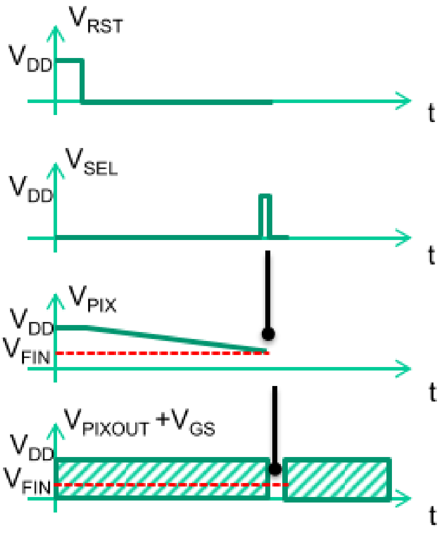


2. **Integration:**
   M_{RS} opens and disconnects the photodiode from the supply. Therefore the photodiode current flows across the capacitance (as we seen before in 9.3) and the voltage V_{PIX} decreases. The source follower is still disconnected and does not dissipate current.



3. **Readout:**
   M_{SEL} closes and connects the source of M_{RD} to the current generator. The source follower is therefore biased and buffers the voltage V_{PIX} at its gate to the output V_{PIXOUT}.





The final value at the end of the integration can be therefore calculated as:

$$V_{PIX}\left(t_{int}\right)=V_{DD}-\frac{\Delta Q}{\left(C_{dep}+C_{g}\right)}=V_{DD}-\frac{\left(i_{ph}+i_{d}\right)t_{int}}{\left(C_{dep}+C_{g}\right)}$$

$$V_{PIXOUT}\left(t_{int}\right)=V_{PIX}\left(t_{int}\right)-V_{GS_{M_{RD}}}$$

$$\Delta V_{PIXOUT}=\Delta V_{PIX}=\frac{\left(i_{ph}+i_{d}\right)t_{int}}{\left(C_{dep}+C_{g}\right)}=\frac{\left(\Phi_{ph}q\eta(\bar{\lambda})A_{pix}\mathrm{FF}+i_{d}\right)t_{int}}{\left(C_{dep}+C_{g}\right)}$$

$$\frac{\Delta V_{PIXOUT}}{\Delta\Phi_{ph}}=\frac{\left(q\eta(\bar{\lambda})A_{pix}\mathrm{FF}\right)t_{int}}{\left(C_{dep}+C_{g}\right)}=G\eta(\bar{\lambda})A_{pix}\mathrm{FF}t_{int}$$

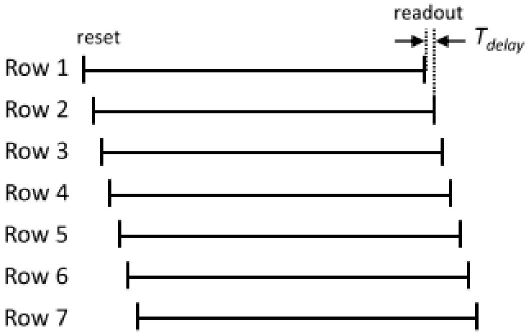
Summing up:

- In the reset phase the photo and dark current discharges to the supply without causing changes to the photodiode voltage.
- During direct integration the current is integrated on the node capacitance determining the voltage drop.
- The voltage value at the end of the integration time is sampled by the buffer.

According to the described operation, all the pixels of the same row have their reset transistor simultaneously active, and then, they have their selection transistor simultaneously active. Therefore, they integrate light during the same time interval (from reset end to readout). On the contrary, all the pixels of the same column have their readout transistor M_{RD} connected sequentially to a single column output are thus readout in a sequential manner.

9.5 Rolling shutter readout

However, the total exposure time of each pixel in the image should be the same. If integration ends at different time instants (impossibility to have simultaneous readout for different rows) then integration needs also to start at different time instants for the pixels of different rows.



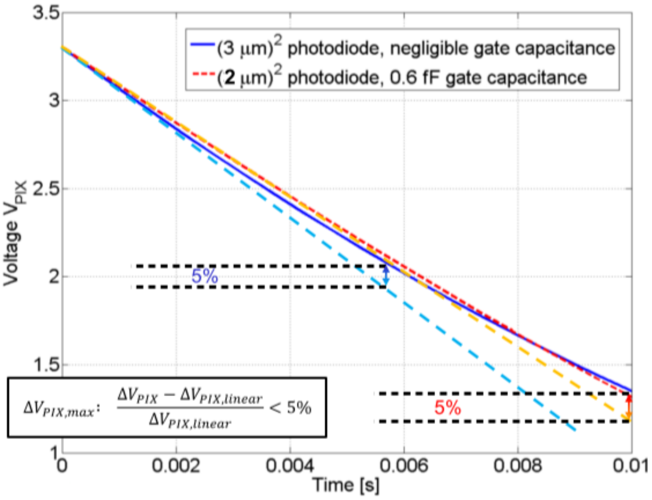
As a consequence, in order to have the same exposure time for all pixels, the reset signal that triggers the charge integration inside each pixel must be delivered with the same delay T_{delay}. This kind of operation is called "**rolling shutter**" mode. In other words, all the pixels acquire the image for the same amount of time but not all parts of the image are recorded at exactly the same time: this may produce noticeable distortion of fast-moving objects relative to the camera.

9.6 Linearity

The photodiode junction capacitance is a function of the reverse bias, giving a **non-linearity** in conversion gain G and sensitivity $\frac{\Delta V_{PIXOUT}}{\Phi_{ph}(\lambda)}$.

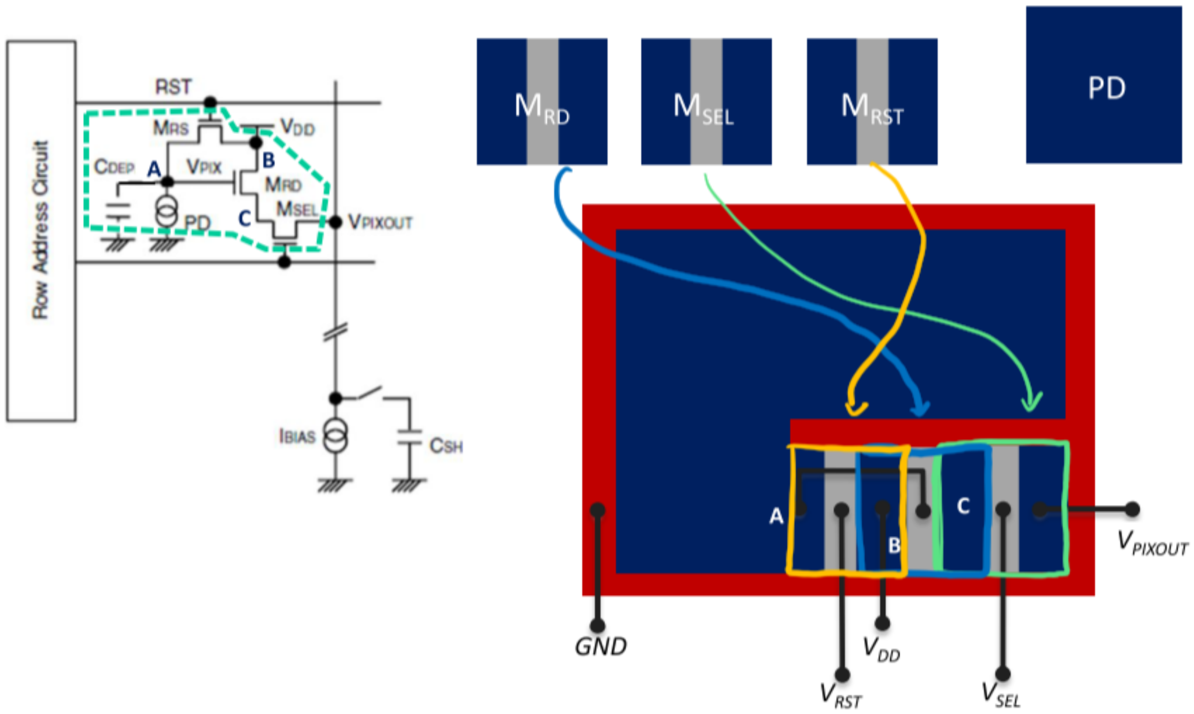
$$C_{dep}\left(V_{PIX}\right) = \frac{\varepsilon_0 \varepsilon_{Si} A_{pd}}{x_{dep}\left(V_{PIX}\right)} = \frac{\varepsilon_0 \varepsilon_{Si} A_{pd}}{\sqrt{\frac{2\varepsilon_0 \varepsilon_{Si}(V_{PIX}+V_{b.i.})}{qN_a}}}$$

The constant gate capacitance luckily mitigates the effect.



Usually, acceptable linearity errors are in the order of **5-7%** (the human eye itself is rather nonlinear! The increase in C_{dep} reduces gain and delays saturation).
As the voltage across the pixel decreases, the PN junction approaches the built-in condition. At the saturation condition: $\mathbf{\Delta V_{PIX,max}} \sim \mathbf{V_{DD}}$, the potential well of the diode is completely full of electrons. Any electrons in excess flows out of the well eventually creating blooming to neighbor pixels. The correct signal information is completely lost for saturated pixels.

Where possible the transistor implants are shared (Area↓, FF↑) as we can see in the figure below:



9.7 SNR of the 3T topology

To evaluate the signal to noise ratio for a given photo- and dark current i_{ph} and i_d, we need to take into account:

- current **shot noise**:
$$S_{i,shot} = 2q(i_{ph} + i_d) \qquad \left[\frac{A^2}{Hz}\right]$$
- **kTC noise** due to the periodical reset of the anode capacitance connected to the M_{RST} MOS "on" resistance:
$$S_{V_{R_{on}}} = 4\,k\,T\,R_{on}$$
- **other** ("readout") **noise contributions**: quantization noise that occurs in the ADC (usually at column level before multiplexing to the output), 1/f and thermal noise of the source follower (typically made negligible)

9.7.1 kTC noise origin

The on-resistance of the M_{RST} transistor is characterized by its own Johnson thermal noise. When the reset switch opens, one random voltage value within the noise power spectral density is "sampled" and frozen as starting voltage for the integration. The bandwidth is represented by the 1/(4RC) equivalent noise bandwidth.

$$\sigma^2_{V,kTC} = \frac{4kTR_{on}}{4R_{on}\left(C_{dep}+C_g\right)} = \frac{kT}{\left(C_{dep}+C_g\right)} \qquad \boldsymbol{\sigma^2_{q,kTC}} = \frac{kT\left(C_{dep}+C_g\right)^2}{\left(C_{dep}+C_g\right)} = \boldsymbol{kT}\left(C_{dep}+C_g\right)$$

This causes a variation for repeated measurement of identical signals.

9.7.2 Shot noise bandwidth

A readout performed as in the described approach (reset + integration) corresponds to a gated-integrator scheme, i.e. an ideal integration for a finite time t_{int}. The noise equivalent bandwidth of such a finite-time integrator is $BW = 1/(2t_{int})$. The shot noise can be therefore converted in terms of squared charge (to be compared with kTC) by integrating over the noise bandwidth and multiplying the squared current by the squared integration time:

$$\boldsymbol{\sigma^2_{q,shot}} = \sigma^2_{i,\text{ shot}}\, t^2_{\text{int}} = \frac{S^2_{i,\text{ shot}}}{2t_{\text{int}}} t^2_{\text{int}} = \boldsymbol{q\left(i_{ph}+i_d\right)t}_{\text{int}} \qquad \left[C^2\right]$$

9.7.3 Quantization noise

Assuming to have a certain number of bits N, we know that quantization noise in terms of voltage at the output is calculated as:

$$\sqrt{\sigma^2_{\text{quant}}} = \frac{LSB}{\sqrt{12}} = \frac{V_{DD}}{2^N\sqrt{12}}$$

As the conversion between charge and voltage is simply given by the sum of the integration capacitances, we can easily write the charge-referred squared quantization noise as:

$$\sigma^2_{q,quant} = \left(\frac{V_{DD}}{2^N\sqrt{12}}\left(C_g+C_{dep}\right)\right)^2$$

Quantization noise will be sized in such a way that its effects are just negligible compared to other noise sources. We will see how the optimum number of bits can be indeed determined as a function of noise and of the maximum measurable signal.

9.7.4 Signal to noise ratio

By considering the shot and kTC noise contributions only, the signal to noise ratio, expressed in dB, is by definition:

$$SNR = 20\cdot\log_{10}\left[\frac{i_{ph}\cdot t_{int}}{\sqrt{q\cdot\left(i_{ph}+i_d\right)\cdot t_{int}+kTC}}\right]$$

The SNR improves with all the parameters that determine a high photo-signal: long exposure time, large lens aperture, large pixel area and large fill-factor (use of µlenses). In particular it is also worthwhile to note that SNR decreases with the pixel size: two sensors with different size and the same number of pixels behave differently in terms of SNR.

9.7.5 Example

The human eye does not perceive differences in the SNR for SNR larger than approximately 30 dB. Even if the SNR is characteristic of the pixel, for the same image under different exposure time or lighting or processing conditions, one can define the average SNR. This allows making comparisons, even if not among different images. For this reason several images (SNR test charts) have become quite popular as testing references.

9.8 Dynamic range of the 3T topology

The **Dynamic Range** (**DR**) quantifies the ability of a sensor to reproduce scenes with wide variations in illumination. With limited DR, a camera loses the capability of reproducing at the same time details both in dark areas (due to **noise**) and in bright areas (due to **saturation**).

Figure 54: **Left:** Noise **Right:** Saturation

Possible techniques to improve the dynamic range:

- alternative pixel topologies;
- multiple images fusion (e.g. HDR option in your mobile phones);
- a combination of both.

The **maximum measurable signal** is the charge $i_{ph,max}\, t_{int}$ corresponding to linearity **saturation**. It is derived from the following condition:

$$dQ = C(V)dV \qquad (i_{ph,\max} + i_d)\, t_{int} = Q_{\max,\text{ well}} \; = \int_{V_{DD}}^{V_{\text{lin}}} \left[C_{dep}(V) + C_g\right] dV$$

We can do some approximations:

1. We assume nonlinearity to be negligible, and thus we consider the entire maximum voltage sweep (i.e. $V_{lin} = 0$).
2. We assume that the maximum signal current $i_{ph,max}$ is far larger than the dark current i_d (this usually occurs in a well-designed advanced photo systems).
3. We assume an average value $\overline{C_{\text{dep}}}$ for the depletion capacitance during integration.

So we obtain:

$$\boldsymbol{Q}_{\max} = i_{ph,\max} \cdot t_{int} \approx Q_{\max,\text{ well}} \; \approx \int_{V_{DD}}^{0} \left[C_{dep}(V) + C_g\right] dV \approx \left(\overline{\boldsymbol{C}_{\text{dep}}} + \boldsymbol{C}_g\right) \boldsymbol{V}_{DD}$$

We calculate the **minimum measurable number of electrons** (**NEQ**) by setting the SNR equal to 1 (dark current will be much larger than photocurrent) (we assume $\sigma_{\text{quant.}} = 0$):

$$i_{ph,\min} \cdot t_{int} = Q_{\min} = \sqrt{\cancel{qQ_{\min}} + q i_d t_{int} + kTC} \quad \xrightarrow[q i_{ph} t_{int} = qQ_{\min}]{i_d \gg i_{ph,\min}} \quad Q_{\min}^2 - (q i_d t_{int} + kTC) = 0$$

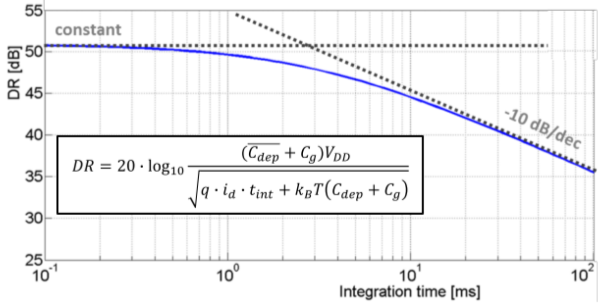
$$\boldsymbol{Q_{\min} = \sqrt{q i_d t_{int} + k_B T C}} \qquad\qquad \text{NEQ} = N_{el} = \frac{Q_{min}}{q} = \sqrt{\frac{i_d t_{int}}{q} + \frac{kTC}{q^2}}$$

With typical numbers we end out with a few electrons to few tens electrons of minimum measurable charge: e.g. $i_d = 0.05\ fA$, $C_{\text{TOT}} = 0.5\ fF$, $t_{int} = 2\ ms \implies NEQ_{\text{rms}} = 9$ electrons. We can verify that it takes only $N = 10\ bits$ at 3 V to get a lower quantization noise.

The **dynamic range** is, by definition, the ratio of the maximum to the minimum measurable signals (it can be calculated as charge ratio, current ratio, whatsoever-quantity ratio):

$$DR = 20 \cdot \log_{10} \frac{Q_{\max}}{Q_{\min}} = 20 \cdot \log_{10} \frac{\left(\overline{C_{dep}} + C_g\right) V_{DD}}{\sqrt{q \cdot i_d \cdot t_{int} + k_B T \left(C_{dep,V_{dd}} + C_g\right)}}$$

As anticipated, the DR quantifies the ability of a sensor to image scenes with wide variations in illumination: all signals between the maximum (brightest) and minimum (darkest), thus all details in between them, can be correctly represented. Dynamic range is a characteristic parameter of a sensor: is independent of the signal, **grows with the pixel area** and is **dependent** on the **integration time**.



Indeed, DR decreases as integration time increases due to the adverse effects of the dark current. The maximum DR is therefore quoted at short integration times and depends only on kTC noise.

Max DR increases: increasing **well capacity** (area and/or supply) or decreasing **read (kTC) noise**.

How to choose the optimum integration time?

- If the scene has a relatively **uniform** illumination:
  there is no need for a large dynamic range and it is wise to increase the integration time to obtain a good SNR (up to 30dB), especially under poor average illumination. However, is important is to use the maximum integration time that avoids saturation (of the brightest pixel). Moreover, avoiding too large integration times will help limiting motion blur.

- If the scene has **wide variation** in illumination:
  we need a large dynamic range and thus is better to keep a lower integration time to reach a higher DR. As a consequence, the SNR, especially for the darkest pixels, will be somewhat sacrificed. A sensor with a low dark current helps, in this case, to reach the maximum DR also at longer integration times.

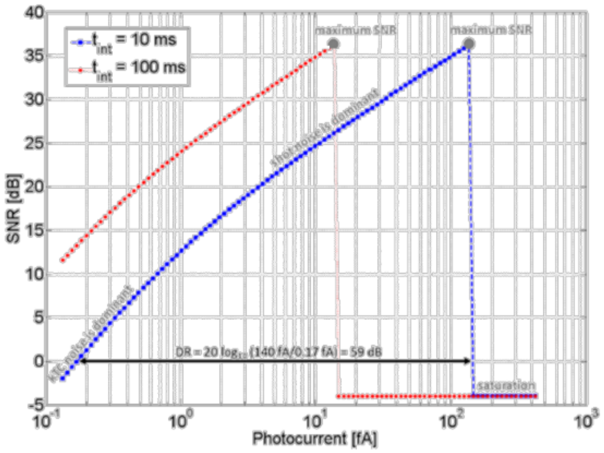
The dynamic range does not correspond to the maximum signal to noise ratio a pixel can reach (the DR inherently refers to more than one pixel). The difference is indeed in that the maximum SNR, should take into account as well the photocurrent shot noise.

$$SNR_{\max} = 20 \cdot \log_{10} \frac{Q_{\max}}{\text{noise corresponding to } Q_{\max}} \sim 20 \cdot \log_{10} \frac{i_{ph,\max} \cdot t_{\text{int}}}{\sqrt{q \cdot i_{ph,\max} \cdot t_{\text{int}}}} =$$

$$= 20 \cdot \log_{10} \frac{Q_{\max}}{\sqrt{q Q_{\max}}} = 20 \cdot \log_{10} \sqrt{\frac{Q_{\max}}{q}} = 20 \cdot \log_{10} \sqrt{N_{el,\max}}$$

The expression of the maximum SNR (always lower than the DR) is thus the typical expression of a Poisson process (mean value N_{el} equals the variance). The maximum SNR is simply dependent on the number of electrons that the photodiode potential well can host.
E.g. to reach $SNR_{max} = 40\ dB$, the well should accommodate $10^4 e^-$.
To conclude this part, it is interesting to plot a graph of the **SNR versus** the **photocurrent**.

$$SNR = 20 \cdot \log_{10} \left[\frac{i_{ph} \cdot t_{int}}{\sqrt{q\,(i_{ph} + i_d) \cdot t_{int} + k_B T C}} \right]$$



As we can see:

- it grows linearly with i_{ph} at low i_{ph} value;
- it increases with the $\sqrt{i_{ph}}$ at high signals;
- it drops suddenly to zero after saturation (here the SNR loses sense: we can assume it as null, indeed no reliable information is given by saturated pixels).

9.9 Fixed pattern noise

Some FPN gain sources are due to the optical responses causing different photo-signals for nominally identical, evenly illuminated, pixels.
The overall pixel response to an input flux $\psi(\lambda)$ (in $[\frac{ph}{s \cdot m^2}]$) is indeed related to:

- nominal Si quantum efficiency;
- color filter transmittance;
- anti-reflection coating and µ-lens transmittance;
- implant depth;
- width of the depletion region;
- presence of defects.

$$i_{ph}(\lambda) = \phi(\lambda) \cdot \eta(\lambda) \cdot q \cdot A_{\text{pixel}} \cdot FF \cdot T_{CFA}(\lambda) \cdot T_{OPT}(\lambda)$$

Other FPN sources lie in the pixel level electronics, and do affect the following pixel gains:

- value of the diode's capacitance and of the M_{RD} gate capacitance:

$$G = \frac{q}{(C_{dep} + C_g)}$$

- source follower gain (transconduttance g_m), select transistor and current source resistance:

$$G_{SF} = \frac{g_m R_{source}}{1 + g_m R_{source}}$$

So we have:

$$\Delta V_{\text{pixout}}(\lambda) = \frac{i_{ph}(\lambda) t_{\text{int}}}{C_{\text{dep}} + C_g} G_{SF} + \frac{i_d t_{\text{int}}}{C_{\text{dep}} + C_g} G_{SF} + \Delta V_{\text{os,eln}}$$

The sources of unwanted offset variations from pixel to pixel are source follower voltage (V_{gs} can vary by tens of mV) and dark current (usually the dominant term).

$$|V_{os,eln}| = \sqrt{\frac{i_{\text{bias}}}{\mu_n C_{ox} \frac{W}{L}}} + V_T \qquad V_{out,d} = \frac{i_d \cdot t_{int}}{(C_{dep} + C_g)}$$

9.9.1 Temporal noise vs Fixed Pattern Noise

Let's assume that we want to compensate fixed pattern noise due to discussed phenomena. Therefore, we are looking for a method to distinguish the FPN and the temporal noise from an output image. It is impossible to infer whether noise is temporal or spatial by just looking at a single FPN-uncorrected image. However, by repeating the same measurement several times, temporal noise reduces with averaging, leaving clearly visible spatial (FPN) noise only.
Techniques to calibrate and compensate the overall sensor's FPN rely on this kind of approach. Once calibrated, every pixel of the camera will have a correction value for offset and gain stored in the memory of the digital section. Unfortunately, the correction can't be perfect, and thus there will be a residual noise effect, even though compensation can reduce FPN percentage values by a factor ~ 10.

9.9.2 Dark Signal Non-Uniformity and Photo Response Non-Uniformity

We will now focus on two sources of fixed pattern noise, affecting respectively the **offset** and the **gain** of CMOS image sensors pixels:

- **Dark Signal Non-Uniformity (DSNU):**
  random variations in offset (the dark current in general dominates in determining the offset value of each pixel, and its fluctuations);
- **Photo Response Non-Uniformity (PRNU):**
  random variations in the spectral response and in the electronic gain of each pixel.

DSNU is the distribution of the dark output voltages among all the pixels in a sensor.

$$\Delta V_{pixout} = \left(\frac{i_d t_{\text{int}}}{C_{dep} + C_g} G_{SF} + \Delta V_{os,eln} \right)$$

To measure the DSNU, one takes a series of K consecutive photos in dark (so that no light enters the camera) for a selected exposure time, the following steps are then performed:

1. The calculation of the average of the series for each (m,n) pixel, to reject temporal noise;

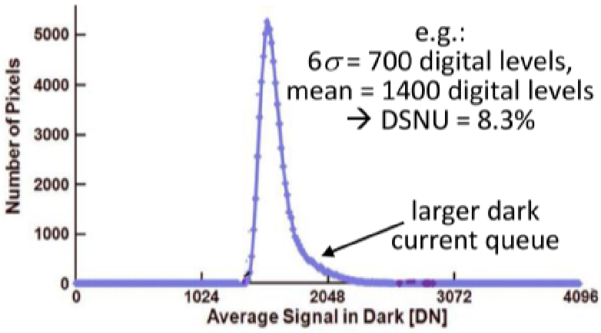
$$v_{out,d_{m,n}} = \frac{1}{K} \sum v_{out,d_{m,n,k}}$$

2. The calculation the standard deviation of all values across the sensor, to quantify DSNU;

$$\sigma_d^2 = \frac{1}{m \cdot n} \sum \left(v_{out,d_{m,n}} - \overline{v_{out,d_{m,n}}} \right)^2$$

3. Normalization to extrapolate % DSNU.

$$\sigma_{DSNU,\%} = \frac{\sigma_d}{\overline{v_{out,d_{m,n}}}} \cdot 100$$



Typical DSNU values obtained before compensation can be as large as **10-12%** (standard deviation to mean ratio). e.g. assuming a 0.2 fA dark current and 10% DSNU, dominated by its non-uniformity at a 10 ms integration time, we obtain 1.3 electrons rms.

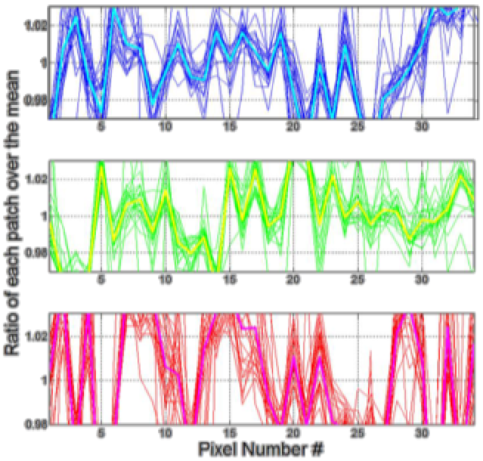
$$\sigma_{DSNU,q} = i_d t_{int} \frac{\sigma_{DSNU,\%}}{100}$$

The **PRNU** is measured as the deviation of the color coordinates of a pixel from the average, due to gain fluctuations calculated (again after averaging) on a suitable set of PSDs which evenly illuminate the sensor. The effects of this kind of variance are harder to predict than for DSNU, as the final coordinates result from the integration of the light spectrum multiplied by the overall photo-response across the wavelengths.

$$\Delta V_{\text{pixout },B} = \phi(\lambda)\frac{\eta(\lambda) q A_{\text{pixel}}\, FF \cdot T_{CFA,B}(\lambda) \cdot T_{OPT}(\lambda) t_{\text{int}}}{C_{\text{dep}} + C_g} G_{SF}$$

In other words: it might be that a defect which is evident at a certain λ is not relevant at another λ. This is why we need to make an average across a suitable set of representative spectral reflectances, illuminated by a representative source

Therefore an option to estimate the PRNU is to measure the differences in color coordinates ($\Delta V_{pixout,B}$, $\Delta V_{pixout,G}$, $\Delta V_{pixout,R}$) when all pixels capture the same color patch (and repeat it for different "typical" target colors). This is done usually with a color checker chart, illuminated by calibrated (known) illuminants. The evaluation method follows these passages:



1. Capture the 24 reference colors one at a time, making averages to remove temporal noise.
2. Evaluate the RGB coordinates (in [V]) for all the pixels (35 shown in the figure aside), for the first reference color (repeating it K times and taking the average).
3. Divide the results by the average coordinate for that color among all pixels. Draw the graph.
4. Repeat the same for all other 23 colors (more lines somewhat overlapping on the graph).
5. For each color channel, the average of the different curves at each pixel position can be interpreted as PRNU (shown in cyan, magenta, yellow).

$$\sigma_{PRNU} = i_{ph} t_{int} \frac{\sigma_{PRNU,\%}}{100}$$

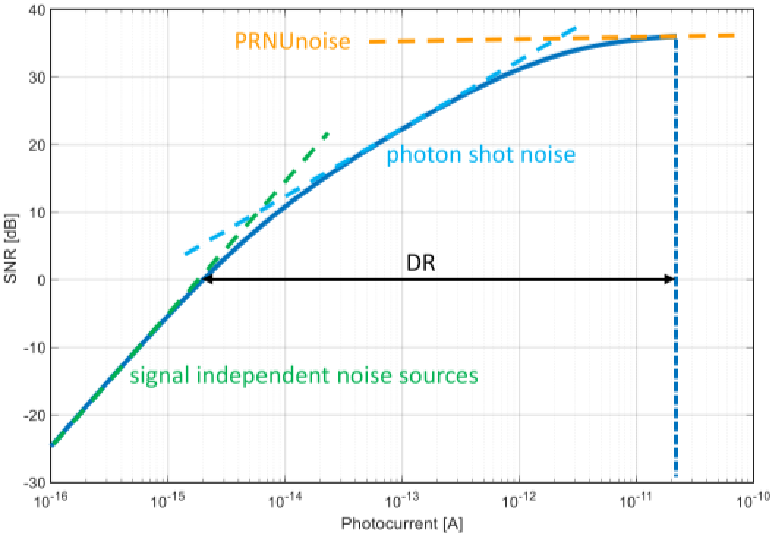
To compensate PRNU one can take the mean curve from the previous graphs and use it as a 1st order correction. Instead, for DSNU correction things are easier: just subtract (pixel by pixel) the measured DSNU, accounting for its dependency on the integration time.

9.9.3 Overall noise

Including PRNU, DSNU and quantization in the SNR formula we get:

$$SNR = 20 \cdot \log_{10}\left[\frac{i_{ph} \cdot t_{int}}{\sqrt{\underset{\text{Shot noise}}{q\left(i_{ph}+i_d\right) \cdot t_{int}} + \underset{\text{kTC}}{k_B T C} + \underset{\text{Quantization}}{\left[\left(C_g + C_{fd}\right)\frac{V_{DD}}{2^{N_{bit}}\sqrt{12}}\right]^2} + \underset{\text{DSNU}}{\left(i_d\, t_{int}\frac{\sigma_{DSNU,\%}}{100}\right)^2} + \underset{\text{PRNU}}{\left(i_{ph}\, t_{int}\frac{\sigma_{PRNU,\%}}{100}\right)^2}}}\right]$$

A sample corresponding SNR vs photocurrent complete graph is shown below.
Note the flattening at large photocharges due to the PRNU and the graphical calculation of the Dynamic Range.



9.10 Limitations of the 3T topology

9.10.1 Limitation 1: gain vs FF (active area)

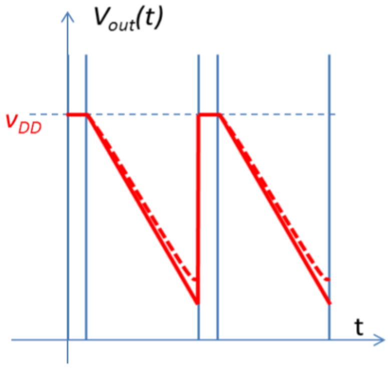
In a 3T APS photodiode the photosensing node (i.e. the photodiode junction, where the charge is collected) is also the conversion node (direct integration). This dual role of the photodiode makes difficult to improve the conversion gain G (from photo-generated electrons to output voltage) while keeping a good fill factor (or a large photodiode area). Indeed, the pixel voltage is proportional to the inverse of C_{dep}, and thus decreases for large junction areas, whereas the fill factor increases, thus generating a trade-off.

$$G = \frac{q}{C_{\text{dep}} + C_g} \underset{\substack{\text{for large}\\ \text{pixel area}}}{\longrightarrow} \propto \frac{1}{A_{\text{diode}}} \qquad\qquad FF = \frac{A_{\text{diode}}}{A_{\text{overall-pixel}}} \propto A_{\text{diode}}$$

9.10.2 Limitation 2: 3T pixel non-linearity vs FF (active area)

Consider the expression of the depletion width and of the depletion capacitance of a reversely biased diode:

$$x_{dep} = \sqrt{\frac{2\varepsilon_0\varepsilon_{Si}\left(v_{bi} + V_{pix}\right)}{qN_A}} \qquad\qquad C_{dep} = \frac{\varepsilon_0\varepsilon_{Si}A_{diode}}{x_{dep}}$$



As the capacitance depends on the reverse voltage, which varies during integration, the pixel response is not linear. In particular, the pixel output during the integration time becomes more nonlinear when the integration capacitance is dominated by the diode's capacitance. Large FF (large photodiode area) means large depletion capacitance with respect to the gate capacitance, so a high nonlinearity (reduced linear Q_{max}). We thus conclude that it is challenging to increase the DR by just increasing the area.

9.10.3 Limitation 3: thermally generated dark current i_d

At typical operating temperatures, dark carriers are generated by thermal excitation. This thermal generation is more likely to happen in those regions where the Si crystalline structure has defects or "traps" (i.e. low time τ_n).
The generation rate $[\frac{C}{s \cdot m^3}]$ and dark current density are:

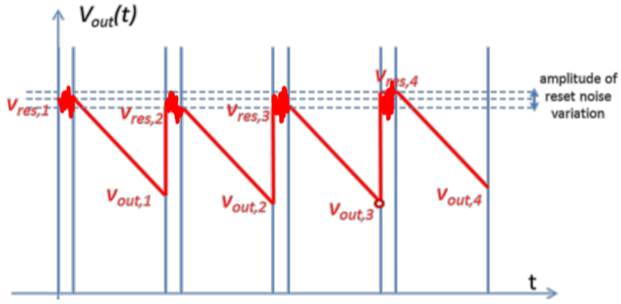
$$G_d = \frac{qn_i}{2\tau_n} \qquad\qquad J_d = \frac{qn_i}{2\tau_n}x_{dep}$$

Dark current contributions come from two regions : bulk dark current, generated in the Si volume and surface dark current, generated at the Si-SiO$_2$ interface. In the 3T topology the collecting anode is directly placed at the Si surface so that it collects all dark current contributions, i.e. both the bulk- and the surface-generated carriers along the junction surface perimeter.

9.10.4 Limitation 4: 3T pixel Reset Noise

The on-resistance of the M_{RST} transistor is characterized by its own Johnson noise.

$$S_{V,R_{on}} = 4\ k_B\ T\ R_{on}$$



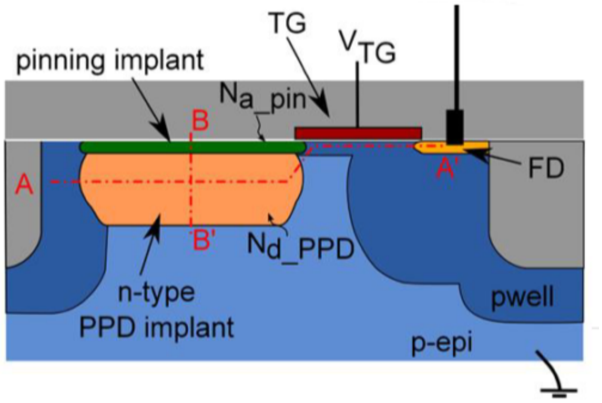
When the reset switch opens, one random voltage value within the noise power spectral density is "sampled" and frozen as starting voltage for the integration.

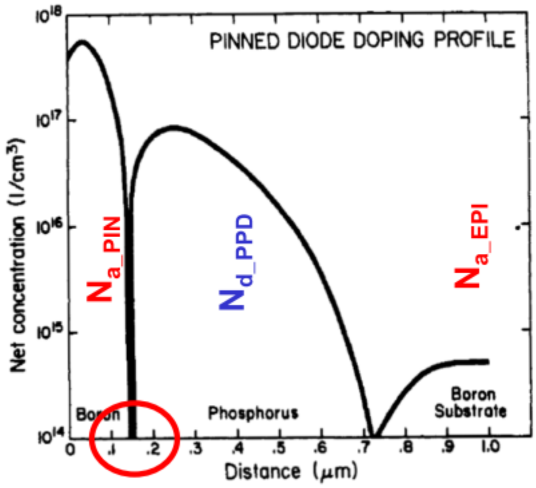
$$\sigma^2_{q,kTC} = k_BT\left(C_{dep} + C_g\right)$$

Repeating the integration several times for an identical signal (or once for identical pixels), the output will vary due to this noise contribution.

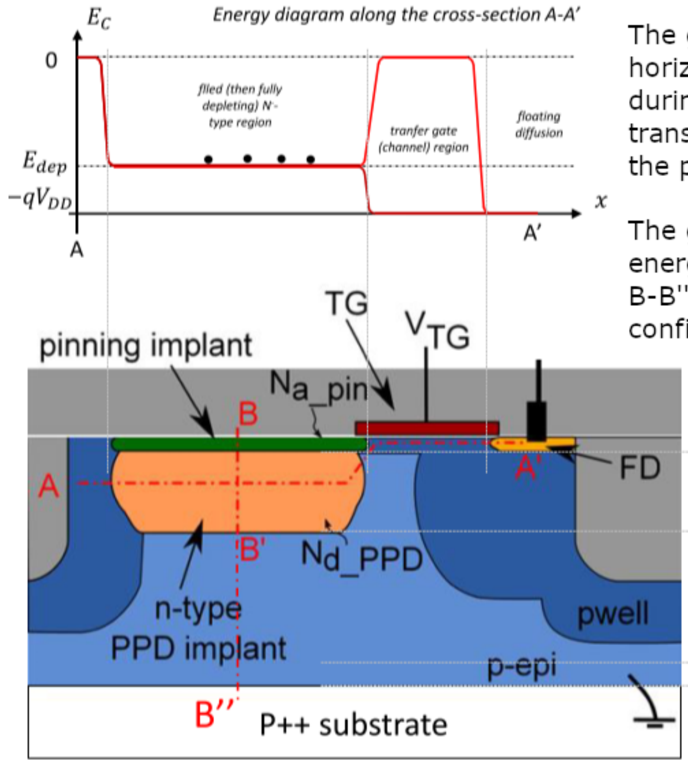
9.11 Pinned photodiode and 4T topology

This kind of photodiode is formed by an N^- implant (N_{d_PPD}) "pinned" (blocked) at the surface by a very shallow (e.g. around 50-nm deep) heavily-doped and annealed P-type implant (N_{a_PIN}). A sample doping profile across the BB' cross section is shown below.





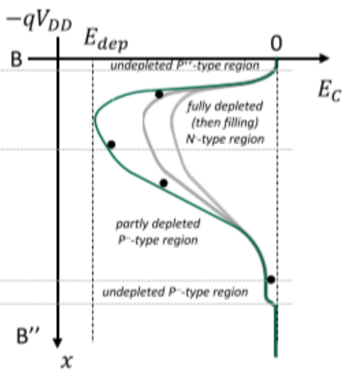
Note that the photodiode anode N_{d_PPD} has no direct ohmic electrical connection: for this reason, the N_{d_PPD} implant can be connected to an N-type floating diffusion (FD), through the channel of a MOS, if the transfer gate (TG) is activated.



lateral P wells act as reflecting walls for electrons, being $N_{A,p\text{-}well} > N_{A,p\text{-}epi}$

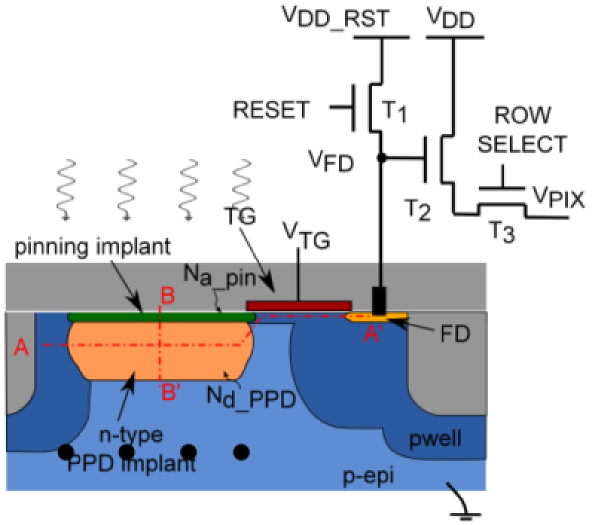
The graph on the left shows the horizontal (A-A' segment) confinement during integration, along with the transfer operation which empties again the pinned photodiode region (in red).

The graph below shows the potential energy of through the vertical segment B-B'', along with its rise (in gray) when confined photoelectrons fill this region.



Energy diagram along the cross-section B-B''

9.11.1 Working principle of the 4T topology



The pixel architecture is formed by:

- **3 n-MOS transistors** (reset, source follower and select), like in the 3T APS;
- **1 further transfer gate** (the fourth T), which has one implant used as the photons collecting junction (the N_{d_PPD});
- a very **shallow pinning P-type implant** (N_{a_PIN}) that separates the N-type anode from the surface.

When light is impinging, signal charge is generated and then collected in the fully depleted N-type pinned diode anode.

Assuming that the pinned photodiode is fully depleted by the former readout operation, we can distinguish 4 operations:

1. **RESET:** with $V_{TG} = 0\,V$, the sense (floating diffusion) node is reset to V_{DD} by closing the reset ($T1 = M_{RES}$). The floating diffusion is thus charged at $v_{FD} \approx V_{DD}$.
2. **INTEGRATION:** the reset transistor is then opened, but charge is accumulated only in the pinned N_{d_PPD} region. The sense node voltage does not change: $v_{FD} = V_{DD}$
3. **TRANSFER:** the TG is closed: (V_{TG} risen): charge is wholly transferred to the floating diffusion (which empties the N_{d_PPD} region), causing a sudden change in the sense node:

$$v_{FD} = V_{DD} - \frac{(i_{ph} + i_d)\, t_{int}}{C_S}$$

4. **READOUT:** the selection transistor closes ($T3 = M_{SEL}$), activating the source follower ($T2 = M_{RD}$) for the readout.

Let's see the advantages we have when we use the 4T topology compared to the 3T topology:

1. **Linearity**
   The capacitance of the sense node (the floating diffusion) can be made low as this small N^+ implant is decoupled from the wide buried N^- implant used for photons collection. As a consequence the sense node capacitance – which is the parallel of the FD junction capacitance and the follower capacitance – is now (small and) dominated by the follower gate, even at large collecting areas (and FF). Therefore, the overall capacitance at the sense node depends very poorly on the accumulated charge. Such a pixel thus avoids typical nonlinearity of 3T topologies even at high FF, becoming very linear.

2. **FF and conversion gain**
   In the 3T topology there is a marked trade-off between the active area and the conversion gain. In the 4T configuration the sensing node has a very small active area, so that the capacitance is dominated by the source follower. Nevertheless, the FF is almost unchanged.

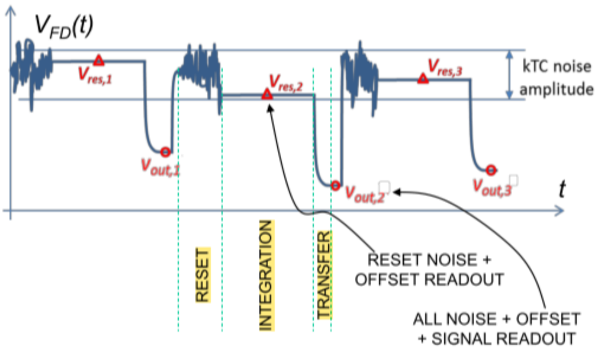
   $$G = \frac{q}{C_{dep,FD} + C_g} \approx \frac{q}{C_g}$$

   A further consequence is that the conversion gain (output voltage variation with respect to the input signal charge) can be larger than in the 3T topology, even at larger collection areas. As a consequence, it becomes also advantageous, for high-end applications, to increase the area, increasing the DR. For instance, a decrease by a factor 10 in the capacitance of the sensing node makes its capacitance negligible over the source follower capacitance, partially increasing gain regardless of the FF.

3. **Dark current**
   The presence of the pinning shallow P-type implant has also a positive effect on the dark current. Indeed the surface generated charges are no more within the N-type region and typically recombine before reaching it (high P doping → short τ_n). Therefore the surface generated current can be almost neglected and the overall dark current density decreases. At the low end of the DR, shot noise associated to the dark current is thus decreased
   $\Longrightarrow$ DR improves for long t_{int}.

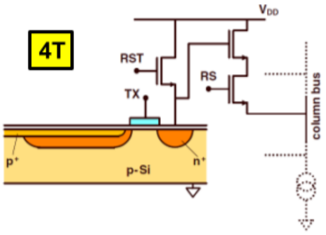
4. **Correlated Double Sampling** (**CDS**)
   We have seen that kTC noise is a sort of freezing of the thermal noise, that occurs when a capacitance is disconnected from a resistive noise source.
   The Correlated Double Sampling (CDS) working principle in a 4T topology is based on the following 4-phase operation:

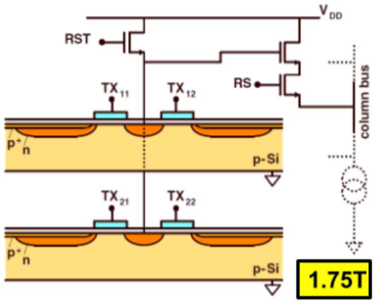
   (a) reset (ends when T1 opens);
   (b) reset noise + eln offset sampling (in the meanwhile, integration occurs in the PPD) (we can read because V_{TG} doesn't change after reset!);
   (c) charge transfer (activation of TG);
   (d) signal + all offset + all noise sampling



When we do a subtraction O-Δ (i.e. $V_{out,i} - V_{res,i}$), we cancel both electronic offset and reset noise (differential readout: step d - step b). Unfortunately, shot noise and dark current are not subtracted, indeed they appear only during integration.

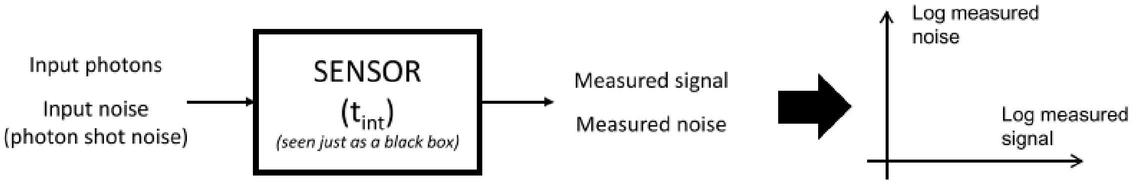
5. **Pixel Sharing the same electronics**
   The presence of a transfer gate enables sharing the floating diffusion, the reset transistor, the source follower and the selection transistor among different pixels. This enables equivalent topologies like the **1.75T** shown below. 4T is suitable for both high performance, large area pixels, but also (still) for small area pixels!



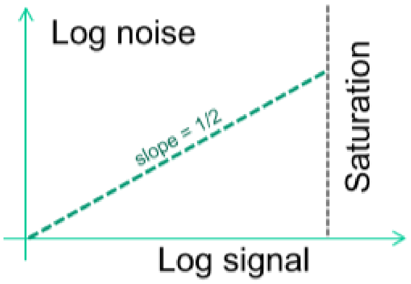


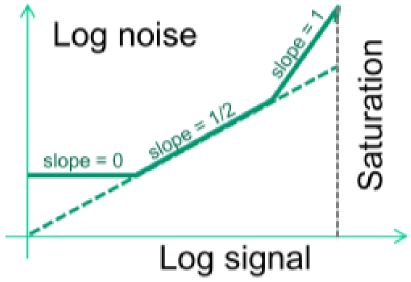
9.12 Photon transfer curve

One standardized test procedure to characterize digital sensors noise (including FPN) is the so-called **Photon Transfer Curve** (**PTC**) method, which is used by several manufacturers as well as by important scientific laboratories. From this single curve one can gather several information on noise, dark current, full-well capacity, sensitivity, dynamic range and linearity. The idea is to provide a graph of the output noise of the camera versus the measured output signal (which in practice is linear with the input signal). Signal and noise are given in the same unit.



The only unavoidable input related noise is the photon shot noise (we know how to predict it). All other output noise components are due to the sensor. The figures below show the ideal behavior of the photon noise and a typical PTC from a sensor on a log-log graph for a given t_{int}.





In the photon transfer curve graph can be found three distinct regions:

1. an **initial flat region** where the signal is low and only read noise (kTC and quantization noise), dark current shot noise and dark signal non-uniformities (DSNU) are visible;
2. a region where photon shot noise becomes dominant over other noise contributions (**slope 1/2** in the log-log-plot);
3. a final region where noise is linear with the signal (**slope 1** in the log-log plot): this region is due to photo response non-uniformities (PRNU).

$$\sigma_{\text{shot },d} = \sqrt{q i_d t_{int}} \qquad \sigma_{kTC} = \frac{\sqrt{k_B T\,(C_g + C_{fd})}}{R_{kTC}} \qquad \sigma_{DSNU} = i_d t_{int} \frac{\sigma_{DSNU,\%}}{100}$$

$$\sigma_{ADC} = (C_g + C_{fd})\,\frac{V_{DD}}{2^{N_{bit}}\sqrt{12}} \qquad \sigma_{\text{shot,ph}} = \sqrt{q i_{ph} t_{int}} \qquad \sigma_{PRNU} = i_{ph} t_{int} \frac{\sigma_{PRNU,\%}}{100}$$

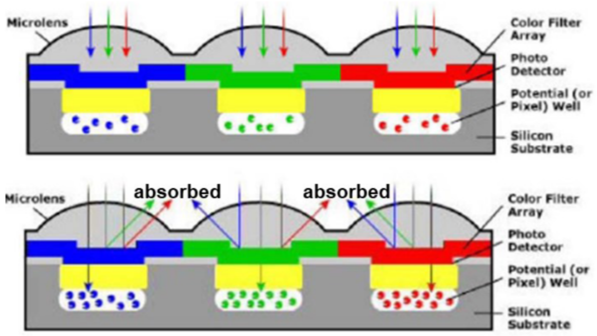
In conclusion, the 3T topology, though advantageous because of its simplicity, carries intrinsic drawbacks that limit its performance. Adding a thin implant and one transistor, a lot of issues are solved:

- dark current (and noise): elimination of the contribution from the junction surface;
- kTC noise, through CDS: Correlated Double Sampling is made possible only by the 4T operation;
- linearity: decoupling of the gain element from collection area;
- transistors sharing among different pixels: only the pinned photodiode and the TG are individually designed for each pixel, all other elements can be shared!

4T topology represents the state of the art of CMOS image sensors.

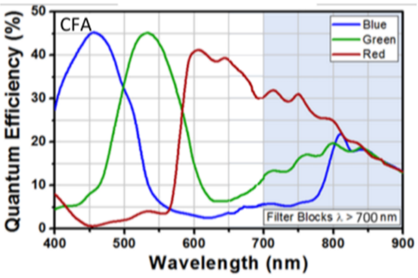
9.13 Color acquisition

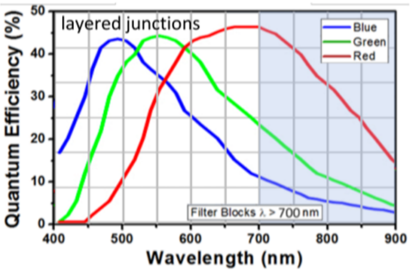
The human visual system is based on a spatial distribution of photoreceptors (cones) that allow **structured** vision. Those photoreceptors can have three different spectral sensitivity functions (L, M or S cones), that allow **color** vision. The result is our capability to perceive **structured color images**.



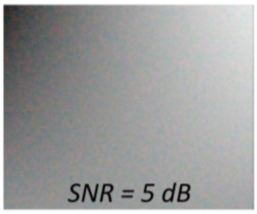
In order to obtain the same results with a digital camera, we need a **matrix of pixels**, to obtain structured (2D) images, and **three** (or more) different **sensitivity functions** of the incoming wavelength. The easiest approach to obtain a color sensor is the **deposition of pixel-level CFA**(Color Filter Array): when the light focused by µ-lenses reaches the CFA surface, only light transmitted by the corresponding filter passes through, generating electrons in the active layer of the pixel.

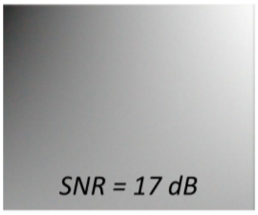
CFA transmittances can have a FWHM (full width at half maximum) around 80-150 nm. In general, the narrower filters have a higher overall light absorption (and loss in signal), whereas the larger filters have a larger color crosstalk (typical of non-CFA pixels).
A sizeable portion of the incoming radiation in the visible range (roughly the 70%) is practically absorbed, and thus wasted, if the CFA approach is used.





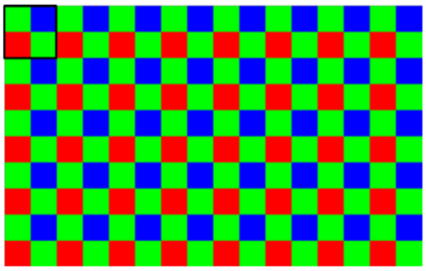
Assume a 75% loss of light (25% average CFA light transmittance), the decrease in SNR with respect to a case where there are no filters is easily quantified:



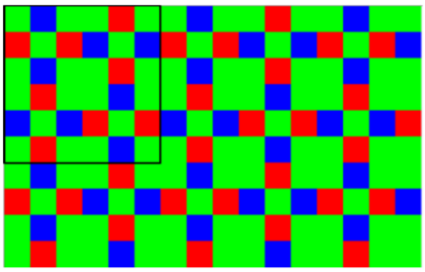


- @ low t_{int}: $20 \cdot log_{10}[0.25] = -12\ dB$
- @ large t_{int}: $20 \cdot log_{10}[\sqrt{0.25}] = -6\ dB$

Color filter arrays can have different spectral transmittance and can be arranged with different patterns: the color information at each pixel position needs to be reconstructed through interpolation algorithms. The GRGB (originally YC_1YC_2) pattern, proposed by Bryce Bayer, a scientist from Eastman Kodak in 1975, quickly became the mostly adopted and ubiquitous pattern. The GRGB pattern, or variations thereof, have more G channel sampling than for R and B channels. This choice is made because the green channel, peaking close to 555 nm, has usually the most similar sensitivity to the photopic curve (the overall perceived intensity by human eyes). The G channel of the CFA is equivalent to "*luminance*" in the opponent colors theory, while B and R channels are considered as blue-yellow and red-green channels.

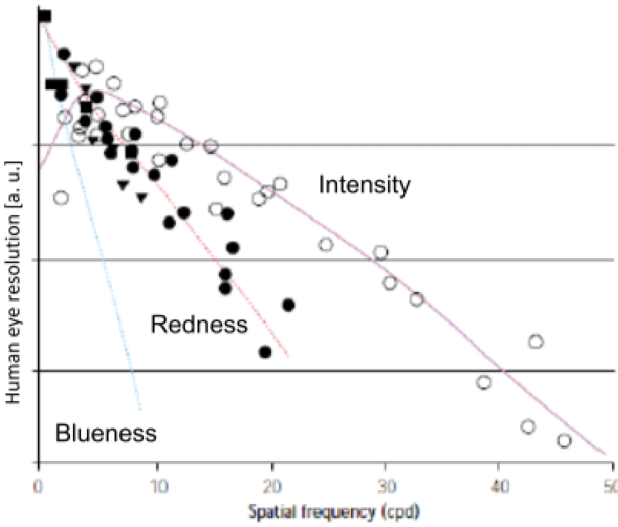


2x2 Bayer pattern



alternative 6x6 GRGB pattern
less prone to aliasing (Fuji)

The figure represents the human eye response vs spatial frequency (i.e. the resolution) obtained from various experiments for the three channels of the opponent colors theory.

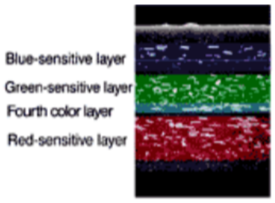


It shows how resolution is better for the intensity information than for the color information. This suggests to operate the capture of color images with higher resolution for the intensity channel (G channel, parent of the photopic curve) than for color channels. Resolution perceived by the eye is most determined by luminance: having twice G-type filter helps enhancing the perceived resolution.

Despite its capillary diffusion, the CFA approach has its own **drawbacks**:

- dedicated **process steps** for filter deposition;
- need for **interpolation** algorithms because only a single spectral function is implemented for each pixel site. The so gathered information is highly incomplete: only 1/2 of G coordinates is measured , while 1/2 is a guess and only 1/4 of RB coordinates is measured, while 3/4 is a a guess. Moreover, interpolation (bilinear, bicubic. . .) requires a large processing computational cost (lowering battery time) and brings artifacts in color reconstruction;
- consistent **loss of SNR** (more than 70% of the incoming signal is wasted). Noise spreads across adjacent pixels through interpolation.

To see how we could find a solution to these problems, lets go back in time to the operation of silver halide film, which is based on absorption properties of randomly distributed and overlapping silver grains and RGB dyes.

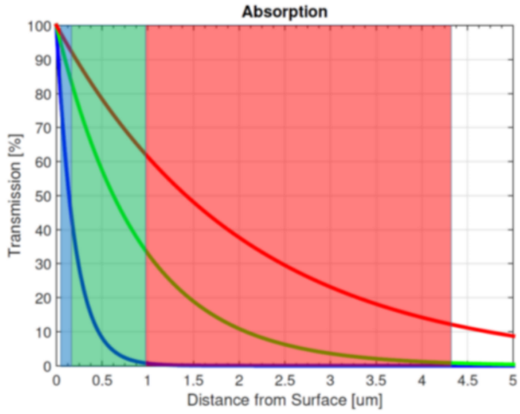


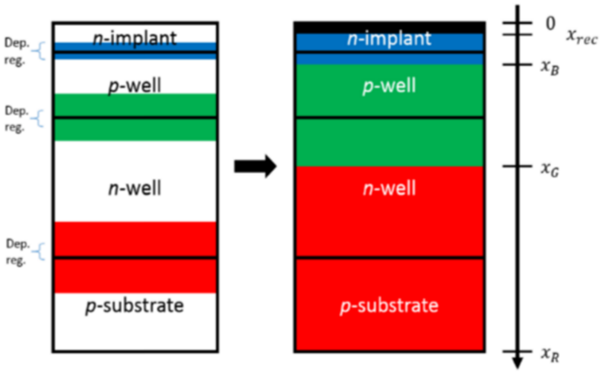
Negative film structure

Color films adopted at least three different alloys in layers: typically the blue-sensitive layer was on top, followed by green and red ones. As a consequence color was sampled almost at each “pixel” (0.2-2 µm) position on all the RGB bands, and with random spatial distribution.

We have seen that also in Silicon visible light is absorbed with different penetration lengths for different λ: we can exploit this property!
This is the basic idea of **filter-less color sensors**, which has the following advantages:

- **No waste of light**, each photon is absorbed in the Silicon active layer without loosing light signal in the CFA.
- **No** need for demosaicking (reduced computational cost): three color coordinates are sampled at each pixel position, so there is no need for interpolation and artifacts arising from this operation are inherently avoided;
- **No CFA deposition** (cheaper process): less process steps in the sensor production;





9.14 Color conversion matrix

Given a digital color camera with **sensitivity functions** $r(\lambda)$, $g(\lambda)$, $b(\lambda)$, the R, G and B values at the output of a pixel illuminated by thespectral distribution $f(\lambda)$ are given by:

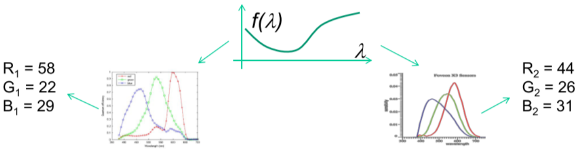
$$R=\int_{\lambda_{\min}}^{\lambda_{\max}} f(\lambda) r(\lambda) d\lambda \qquad G=\int_{\lambda_{\min}}^{\lambda_{\max}} f(\lambda) g(\lambda) d\lambda \qquad B=\int_{\lambda_{\min}}^{\lambda_{\max}} f(\lambda) b(\lambda) d\lambda$$

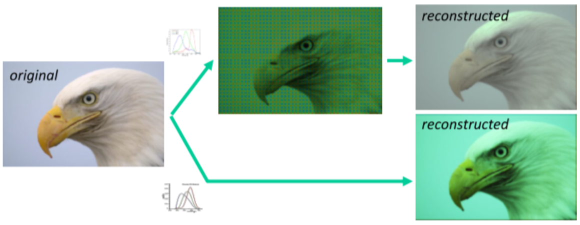
Note: R, G, B can be volts, coulombs, n. of bit, or any normalization thereof.

For the sake of simplicity, the sensitivity functions $r,g,b(\lambda)$ shown below include all effects from optical path $T_{OPT}(\lambda)$, Silicon responsivity $\eta(\lambda)$, CFA transmittance $T_{CFA}(\lambda)$ and electronic gains, and also offset is assumed negligible.

$$\Delta V_{\text{pixout },B}=\phi(\lambda)\frac{\eta(\lambda) q A_{\text{pixel}}\, FF \cdot T_{CFA}(\lambda)\cdot T_{OPT}(\lambda) t_{\text{int}}}{C_{\text{dep }}+C_g} G_{SF}$$

So, $r(\lambda)$, $g(\lambda)$, $b(\lambda)$ represent the 3D camera color space. Each color, defined by R, G, B values, is a 3D vector in this specific color space. Can such RGB values be used by a display?
If we consider cameras with different spectral responses, the output values for the same impinging spectrum are, in general, different! Here, there is an example of the same original image (620 x 380 pixels) captured with different sets of sensitivities, uncorrected (demosaicking if needed, but no color transformation) and rendered by the same display.





From a colorimetric point of view, each camera speaks a different mother language, which is also different from standard (LMS or XYZ) color spaces. On the other side of the imaging chain, displays and printers also speak their own language.
To solve the problem, we need to "teach", to every camera, a common language, through which it can correctly communicate the captured information to other devices.
The teaching operation is named **color conversion** or **camera color calibration** and the common language is known as a **device-independent** (or standard) **color space**. This is defined by analytical sensitivities and usually is a space linearly derived from cones sensitivities space.

The definition of the color conversion (or calibration) procedure, i.e. the identification of the transformation matrix, is an operation which is done once, in a camera characterization/calibration phase. The found optimum **color conversion matrix** (**CCM**) is then applied to all the captured images during the entire camera lifetime. The procedure is the following:

1. use a calibrated source (known spectrum) and a color chart (known reflectance spectra), so a set of known target [XYZ] coordinates;
2. measure the [RGB] values obtained through the camera, for every color patch;
3. find the optimum matrix A that turns the [RGB] triplets into approximated [XYZ] triplets (e.g. with the lowest mean square error, using the pseudo-inverse approach).

$$\begin{bmatrix} \hat{X} \\ \hat{Y} \\ \hat{Z} \end{bmatrix} = \boldsymbol{A} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

9.15 White balance operation

The spectrum of the radiation diffused by an object is a function of the object spectral reflectance and of the illuminant spectrum. Therefore, the color (RGB) of an object changes under different sources: e.g. a "blue" object can stimulate more red and green cones than blue ones under specific illuminants. However we have seen experimentally that perceived colors are quite independent of the viewing conditions because the human visual system somewhat adapts to changes in the illuminant. We therefore need to correct the color images captured by a digital still camera (DSC) to account for this phenomenon.

A digital imaging system, as seen so far, simply captures the radiation reflected (or diffused) by an object through three sensitivity functions and converts the information into the XYZ space. To capture images as perceived by human eyes, we thus need to perform a suitable correction.

The usual compensation procedure consists in a correction of the acquired colors so that the coordinates of a white object become those expected under a reference illuminant.

This procedure is thus indicated as white balance.

White balance algorithms operate through two steps:

1. **source identification**: manual or automatic algorithms identify the source (with an intrinsic assumption that we always have an achromatic object in our scene);
2. **White balance correction**: color coordinates are adjusted to match the coordinates of the target white.

There are three type of source identification:

- **MANUAL**: for professional or semi-pro photographer, there exist suitable objects called **perfect reflectors**. Their spectral reflectance is almost even across the visible range. White color coordinates can be manually measured by pre-capturing an image of the perfect reflector in the actual illuminant conditions and telling the camera that this is the measured white:
$$W_{meas} = \left\{\hat{X}_W, \hat{Y}_W, \hat{Z}_W\right\}$$
- **GRAY WORLD**: this algorithm assumes that, given an image with sufficient color variations, the average of reflectance spectra of a scene should be achromatic:
$$X_{average} = Y_{average} = Z_{average}$$
If this is not true, the shift from gray of measured averages on the three channels is due to the illuminant. The XYZ coordinate of the image white (gray) are identified from the averages.
$$\begin{cases} \hat{X}_W = X_{average} \\ \hat{Y}_W = Y_{average} \\ \hat{Z}_W = Z_{average} \end{cases} \implies \quad W_{meas} = \left\{\hat{X}_W, \hat{Y}_W, \hat{Z}_W\right\} = \{X_{average}, Y_{average}, Z_{average}\}$$

- **WHITE POINT**: assuming that there is always some white in the scene, this algorithm looks for it in the image and considers it coordinates to be those of the source.
  The algorithm assumes that the brightest non saturated pixel corresponds to an object point on a glossy or specular surface, if not from pure white color.
  The algorithm identifies this pixel as the max X, max Y and max Z pixels neighborhood found in the image. The measured white thus is found as:

$$W_{meas} = \left\{ \hat{X}_W, \hat{Y}_W, \hat{Z}_W \right\}$$

WB correction through Von Kries approach. The correction is then operated on all the sensor pixels by using the following correction coefficients, based on a reference white:

$$W_{ref} = \{X_{W,ref}, Y_{W,ref}, Z_{W,ref}\}$$

$$\alpha = \frac{X_{w,\text{ ref}}}{\hat{X}_W} \qquad \beta = \frac{Y_{w,\text{ ref}}}{\hat{Y}_W} \qquad \gamma = \frac{Z_{w,\text{ ref}}}{\hat{Z}_W}$$

To write the WB correction using the same matrix formalism seen for color conversion, it is useful to write the coefficients as the following diagonal matrix:

$$W = \begin{bmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{bmatrix}$$

So that the two operations of color conversion and white balance can be written as in the following expression:

$$\begin{bmatrix} \hat{X}_{WB} \\ \hat{Y}_{WB} \\ \hat{Z}_{WB} \end{bmatrix} = \begin{bmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{bmatrix} \cdot \begin{bmatrix} \hat{X} \\ \hat{Y} \\ \hat{Z} \end{bmatrix} = \begin{bmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{bmatrix} \cdot A \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix} = W \cdot A \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

To sum up:
A camera is not only a matrix of sensing elements: we discussed the relevance of the optics and we have seen how to obtain full color images, rendered by suitable algorithms, possibly as perceived by the human eye.
This occurs through (at least) these steps:

- **3-dimensional color capture** through three different spectral responses, either CFA plus interpolation or layered-junction;
- **color conversion** from camera RGB space into a standard 3-D space (device independent), usually derived from the human visual system;
- **white balance** through the measurement of the actual illuminant, compared to a reference illuminant.