Course Notes

Spacecraft Attitude Dynamics and Control

Politecnico di Milano
Prof. Franco Bernelli Zazzera

Part 2:
Attitude sensors and actuators, attitude determination

Note for the reader

These short notes are in support of the course “Spacecraft attitude dynamics”, they are not intended to replace any textbook. Interested readers are encouraged to consult also printed textbooks and archival papers.

Technological evolution provides continuously upgraded and improved sensors and actuators used for attitude determination and control of spacecraft. These notes, therefore, report some basic concept and operating principle of the most common sensors and actuators, and the reader is encouraged to consult the catalogues of the manufacturers to collect information of the actual functional schemes.

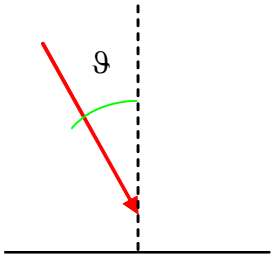
Franco Bernelli Zazzera

Index

Spacecraft attitude sensors

Sun sensors

Sun sensors are normally based on materials that produce an electric signal when illuminated by Sun radiation. Measuring the electric output, in general a current, it is possible to evaluate the Sun incidence angle with respect to the surface illuminated.



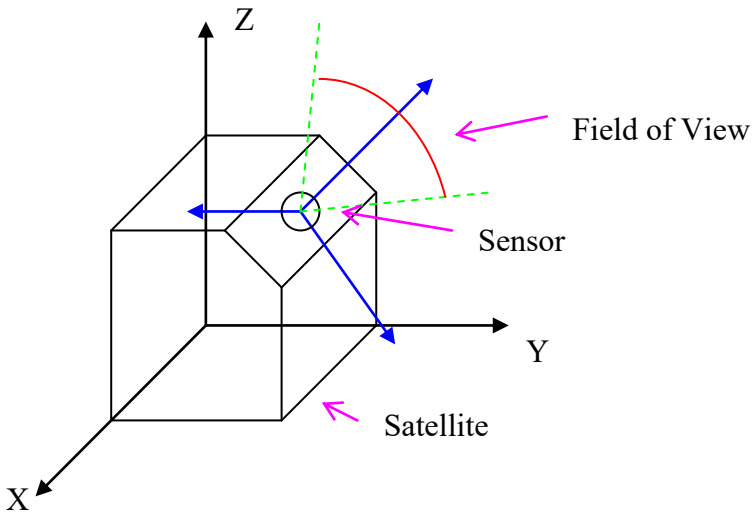
$I = I_0 \cos\vartheta = \alpha SW \cos\vartheta$

where S is the sensor surface, α is a coefficient and W the intensity of the incident radiation.

Sensors are normally built using a series of surfaces with appropriate geometry. Three major categories are identified:

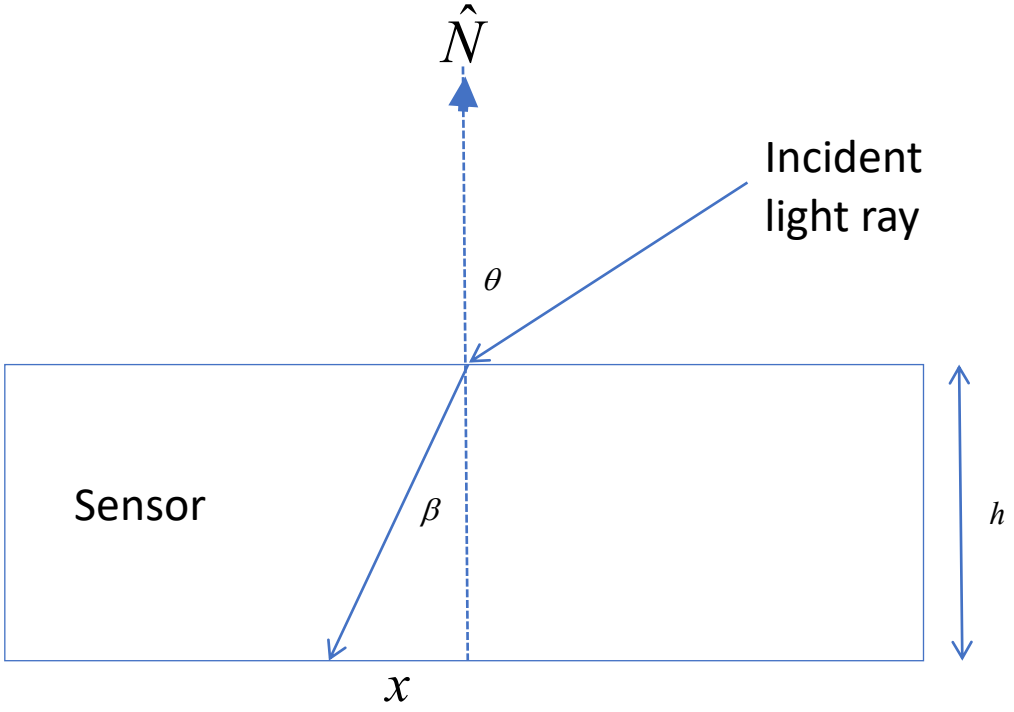
- Sun presence sensors, that essentially detect the presence of the Sun only in a narrow portion of the surrounding space
- Sun sensors, the most used ones, that provide information on the position of the Sun with a wide field of view
- Fine Sun sensors, that provide a precise position of the Sun within a narrower field of view

Every sensor has a well-defined field of view (FOV) and a reference axis that defines the local zero angles. The sensor provides a measure of the relative position of the Sun in the sensor reference. This measurement has to be transformed into information in the satellite reference through a rotation, knowing the orientation of the sensor reference with respect to the principal inertia reference (typically).



The Sun sensor can be calibrated according to the mission of the satellite. For geocentric orbits the sensor will always see the Sun under an apparent angle of 0,53 degrees, that is the apparent Sun diameter at 1 AU, while for interplanetary orbits the angle depends on the position. Normally, Sun sensors are calibrated for use in geocentric orbits.
Sun sensors are widely adopted. They are used not just for attitude determination purposes, but also for instrument pointing, solar panel pointing and/or thermal requirements verification.

One-axis Sun sensor optics can be described using the Figure below:



Snell's law gives $n\ sin\,\beta = sin\,\theta$ where n is the refractive index and, from simple geometry, we have

$$sin\,\beta = \frac{x}{\sqrt{x^2 + h^2}}$$

It follows that

$$\sin\theta = \frac{nx}{\sqrt{x^2 + h^2}}$$

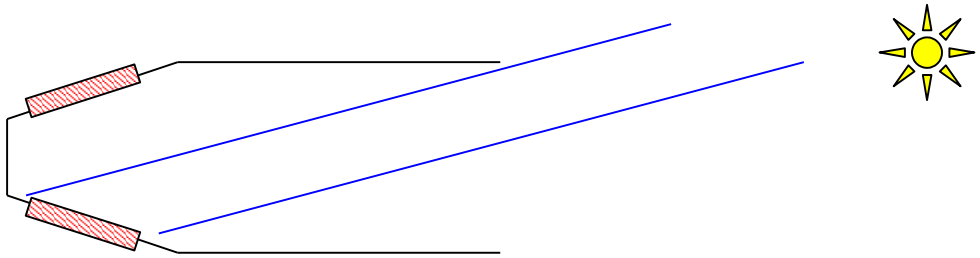
Given that the thickness h is known only the distance x is required to compute the angle between the Sun and the normal to the sensor which can be measured with a digital detector. Three axis Sun sensors are used which yield three independent measurements $\theta_1, \theta_2, \theta_3$

$$\begin{bmatrix} cos\,\theta_1 \\ cos\,\theta_2 \\ cos\,\theta_3 \end{bmatrix} = \begin{bmatrix} \hat{N}_1 \cdot \underline{s} \\ \hat{N}_2 \cdot \underline{s} \\ \hat{N}_3 \cdot \underline{s} \end{bmatrix} = \begin{bmatrix} \hat{N}_1^T \\ \hat{N}_2^T \\ \hat{N}_3^T \end{bmatrix} \underline{s}$$

from which the Sun direction vector $\underline{s}$ can be computed.

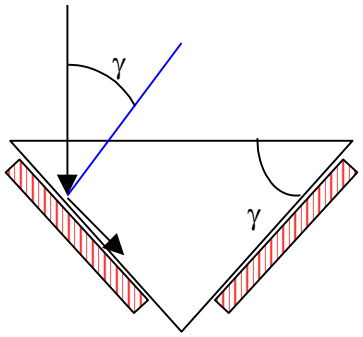
Sun presence sensors

Sun presence sensors provide a sort of binary information: the Sun is detected or it is not detected in the sensor field of view, which is usually restricted. The cosine law is not applicable in this case, due to the presence of shields or support structures. It is common practice to adopt configurations with more than one photocell, as in the following picture.



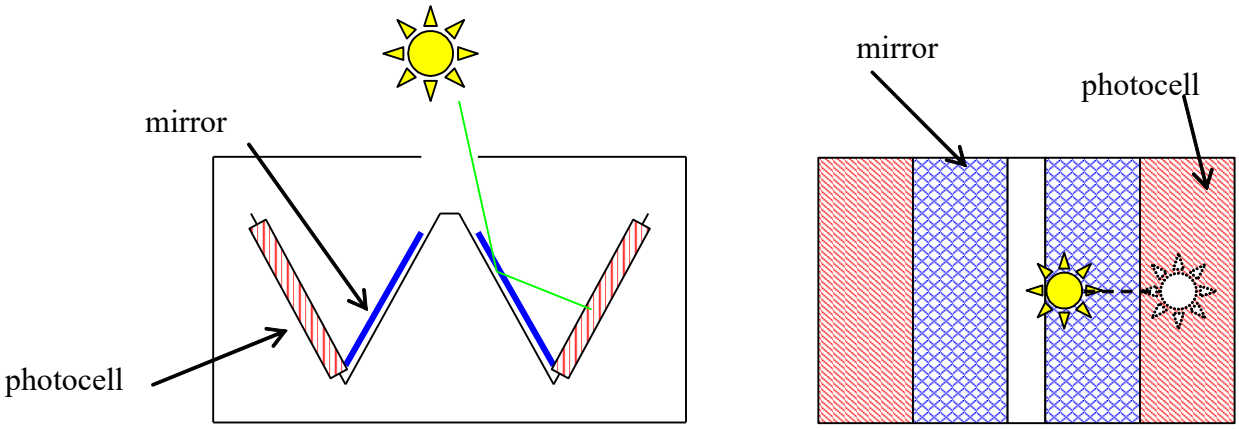
When the Sun moves, either one, both or no photocell is illuminated. Comparing the output signal of the two photocells we can then have the information of the Sun presence and a rough indication of the position.

One alternative geometry uses the principle of refraction in a prismatic configuration.



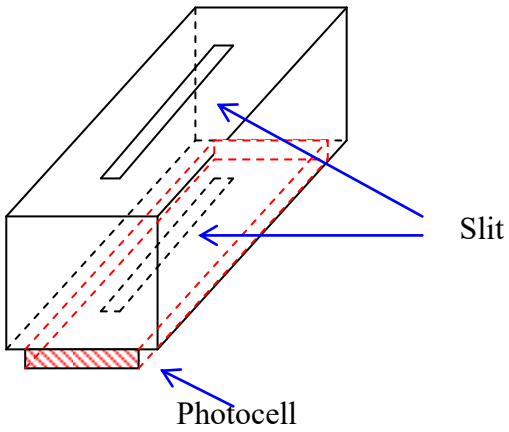
If n•sin γ =1, with n indicating the refraction index of the prism, in the condition of the above figure radiation does not hit the photocell since it is refracted. As the angle γ varies, on one photocell the incidence will be greater than γ and on the other smaller, therefore on one photocell radiation will be totally reflected and on the other partially absorbed, providing an electric output.

A further geometry is the following:

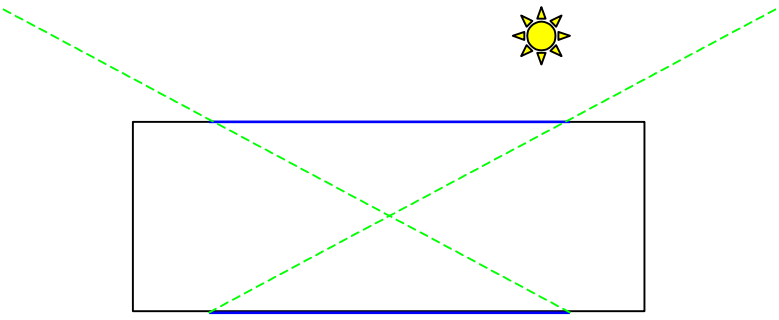


As the Sun moves from the sensor axis, the reflecting surface is more and more illuminated up to the point when the Sun image is completely on the reflecting surface. From this condition the output would be constant even for further Sun inclination, the sensor is saturated.

We can now look at a binary sensor:



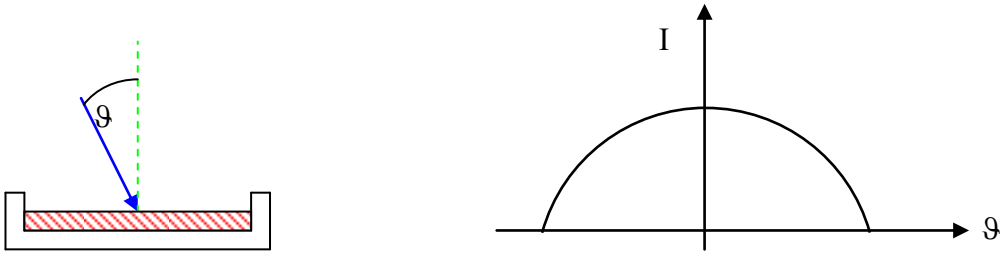
When the Sun ray passes through both slits, it means the Sun is in the plane of the slits and in a definite region of that plane, as seen in the following figure:



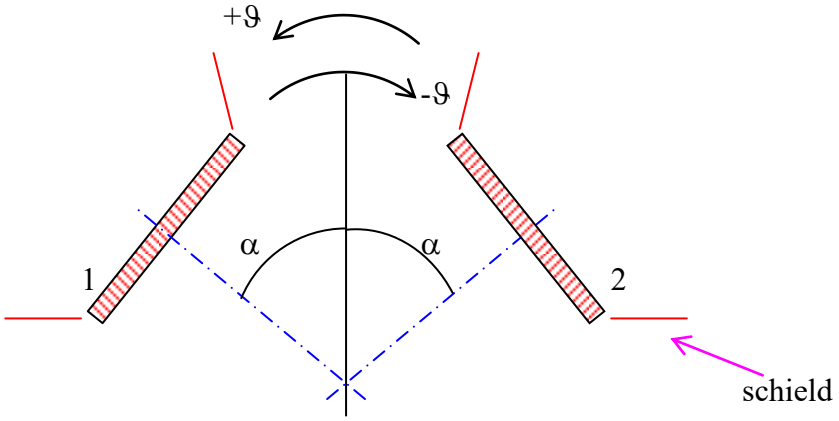
Analog and digital Sun sensors

Sensors providing the position of the Sun can be divided into two main categories: analog sensors and digital sensors. We start by considering the analog sensors.

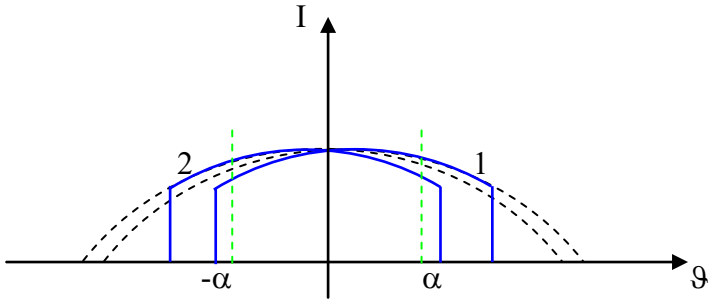
Using the cosine law it is not possible to distinguish positive angles from negative ones. Furthermore, small errors in the measurement close to the sensor zero (vertical) will deliver large errors in the evaluation of the Sun angle, since the variation of the cosine is small close to zero angles.



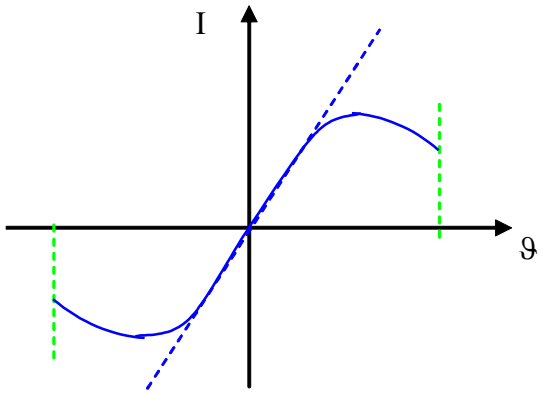
The problem can be limited by adopting two photocells inclined with respect to the zero-axis of the sensor:



Looking now at the intensity of the output current of the two photocells, within the limits guaranteed by the presence of the shields, we have:



Taking now the difference of the two:

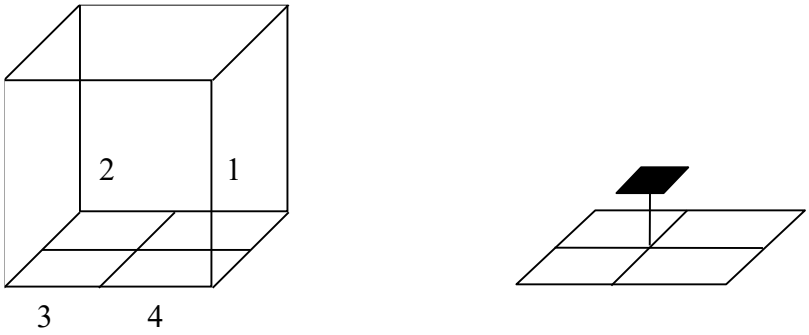


The combined output curve is expressed as:

$$cos(\vartheta - \alpha) - cos(\vartheta + \alpha) = cos\,\vartheta\,cos\,\alpha + sin\,\vartheta\,sin\,\alpha - cos\,\vartheta\,cos\,\alpha + sin\,\vartheta\,sin\,\alpha$$
$$= 2\,sin\,\vartheta\,sin\,\alpha = k \cdot sin\,\vartheta$$

where the constant 2sinα depends on the sensor geometry. The output signal is then sinusoidal, and for small angles it can be approximated by a linear function. If the angle α becomes large, then the useful operating field of view becomes smaller, since the output can be considered reliable only when both photocells are hit by Sun radiation.

With a two-cell sensor we can retrieve the information on the Sun position in one plane, while to obtain the position in space a four-cell configuration has to be adopted, as depicted in the following figure.



The sensor box or shield produces a shadow on each photocell, and the sensor will provide reliable information only when all 4 cells are illuminated. For each cell we have:

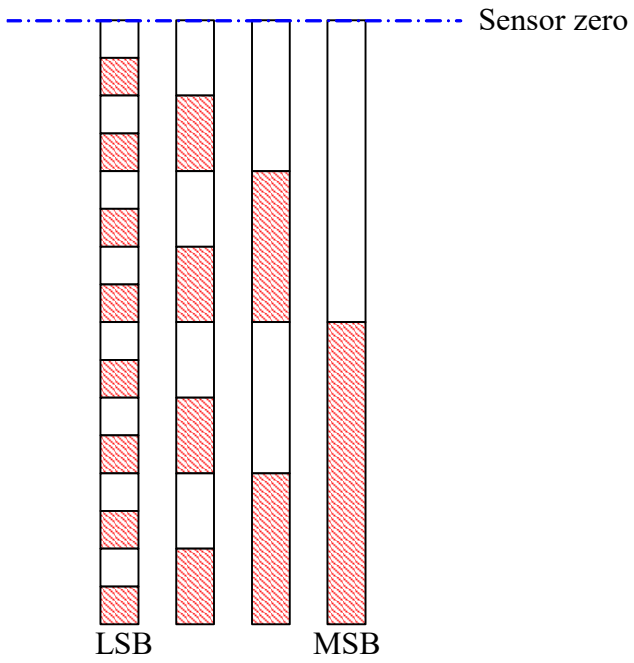
$$I = \mathrm{cA}_0 \cdot \underline{i} \cdot \underline{n}$$
$$A_0 = f(\vartheta, \phi)$$

A_0 is the illuminated area and it is a function of the Sun direction and the sensor geometry. In fact, the product $\underline{i} \cdot \underline{n}$ is known from the Sun distance, so that each output will be proportional to A_0 which in turn can be evaluated as a function of the Sun position.

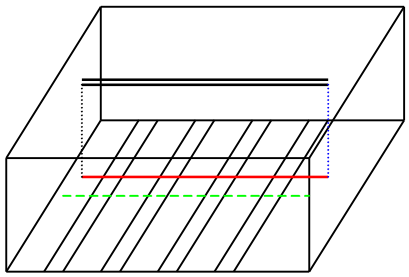
There exist also digital Sun sensors, which provide directly binary information dependent on the Sun position. To understand the sensor principle of operation, start looking at a classical binary table:

Analog	Binary	Analog	Binary
0	0000	8	1000
1	0001	9	1001
2	0010	10	1010
3	0011	11	1011
4	0100	12	1100
5	0101	13	1101
6	0110	14	1110
7	0111	15	1111

The corresponding sensor can be built by using four sensor strips with appropriate mask. In the following figure the red zones correspond to bit set to 1.

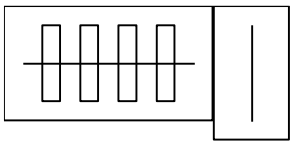


The photocells have all the same properties but their masks are different. When the cells are placed at the bottom of a box with one slit, the Sun produces a line of light that hits each photocell in the same position, generating a signal according to the mask of the series of photocells. This signal can be interpreted as a binary signal. The cell with the finest grid is the least significant bit (LSB), while the cell with the coarse grid is the most significant bit (MSB).



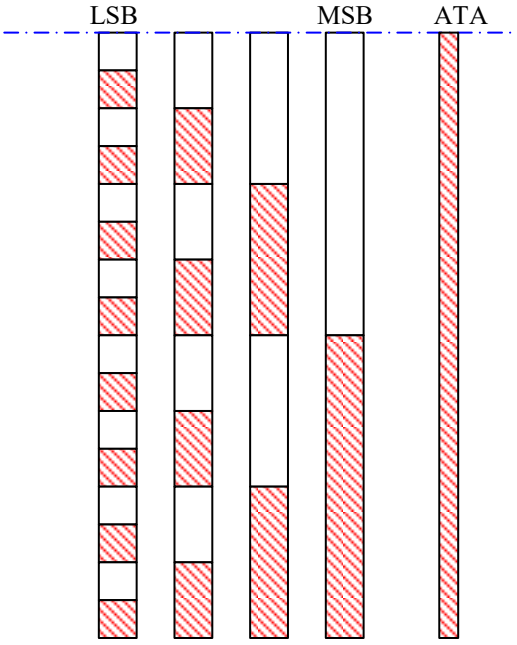
This sensor provides a meaningful signal only when the Sun is in a plane orthogonal to the top slit, otherwise some photocells might not be hit by radiation and their output would become not dependent on the Sun angle. To avoid this, the sensor includes always a Sun presence sensor with disposition

orthogonal to the top slit, as in the figure below, and the output of the Sun presence sensor triggers the decoding of the sensor signal.



At 1 AU distance, the Sun is seen under an apparent angle of approximately 0.5 degrees, so this is the thickness of the light line hitting the photocells. The finest grid should then have a 0.5 degrees sequence. The LSB is then providing a change in the output at every 0.5 degrees change in the Sun position. The overall sensor field of view (FOV) is then a function of the number of strips (bits) of the sensor, since each additional bit doubles the field of view. With 7 bits we have 64 degrees FOV, with 8 bits 128 degrees FOV, that is the maximum since it is useless to go beyond 180 degrees.

To decode the output signal, a threshold has to be defined above which the signal is interpreted as 1, below which it is interpreted as 0. The threshold can be made automatic and position-independent by including a further photocell strip, of width half the width of the other strips and with no cover grid, called Automatic Threshold Adjust (ATA).



Comparing the output signal of the ATA with the other photocells, each bit is set to 1 if the ATA has a lower output; it is set to 0 if the ATA has a higher output. This because the output of the ATA is always half the maximum output of any other photocell, due to the half width of the ATA, irrespective of the position of the Sun. In fact, calling j the grid spacing of the LSB, k the width of the LSB and k/2 the width of the ATA, we have:

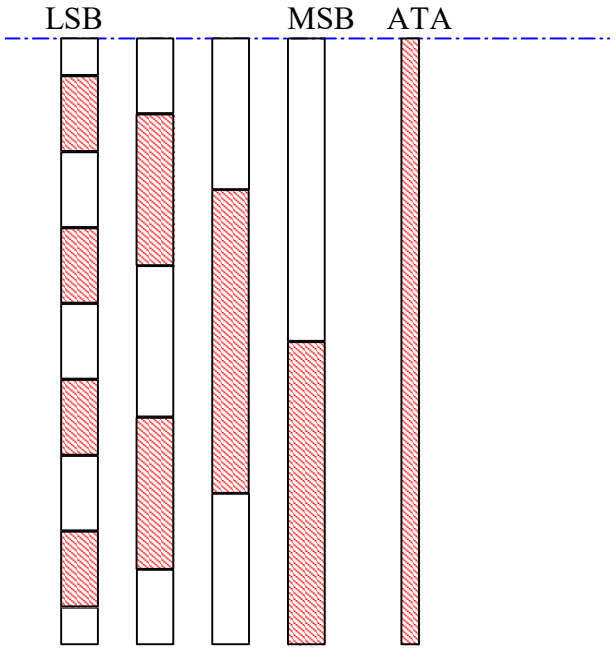
$$I_{ata} = \alpha \frac{jk}{2} \cos\theta$$
$$I_{lsb} = \alpha x k \cos\theta$$

where x is the portion of the LSB illuminated by the radiation. If $I_{ata}>I_{bit}$, then k/2>x and the bit is set to 0 because the LSB is illuminated by less than half. On the contrary, if $I_{ata}<I_{bit}$, then k/2<x and the bit is set to 1 because the LSB is illuminated by more than half. This procedure does not suffer from the changes of the output due to the cosine law and any variation in the intensity of the incident radiation. In a similar way, all bits are compared to the ATA to decode them and set them to 0 or to 1.

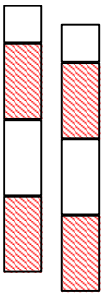
When any bit strip is illuminated exactly for half the maximum, then any measurement error can trigger the detected output to 1 or to 0, in a rather random way. This is a great problem when a series of bits are in this condition, such as between the decimal value 7 and 8. In this case, any measurement noise can trigger any bit to 1 or to 0, so the actual decoded signal would not look like 0111 (decimal 7) or 1111 (decimal 8), but it could be any combination of 0 and 1, including 0000 (decimal 0) and 1111 (decimal 15). This is obviously a huge error, and to avoid this situation the binary code has to be modified in order to avoid having to consecutive integer numbers differing by more than 1 bit. This appropriate code is called Gray code.

Analog	Binary	Gray
0	0000	0000
1	0001	0001
2	0010	0011
3	0011	0010
4	0100	0110
5	0101	0111
6	0110	0101
7	0111	0100
8	1000	1100

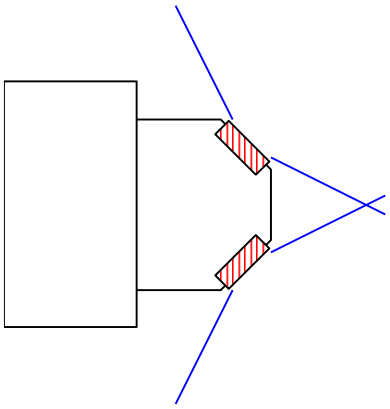
The grid spacing of the Gray code is twice as large as the classical binary code, but is has a symmetric configuration so that the resolution is not changed. With this code, random errors due to misinterpretation of bits are limited to 1 sensor unit, in this case 0,5 degrees.



By placing one additional grid besides the LSB, but shifted by half unit, the sensor resolution can be improved. Once the basic bits are interpreted, a further analysis of the LSB and the additional bits provides a further half unit of resolution (one quarter of degree) if only one further bit is added, and a quarter unit resolution in the case of two added bits. This has the effect of providing resolutions up to approximately 1/8 of degree. For higher resolution, specific and mission dependent solutions must be implemented.

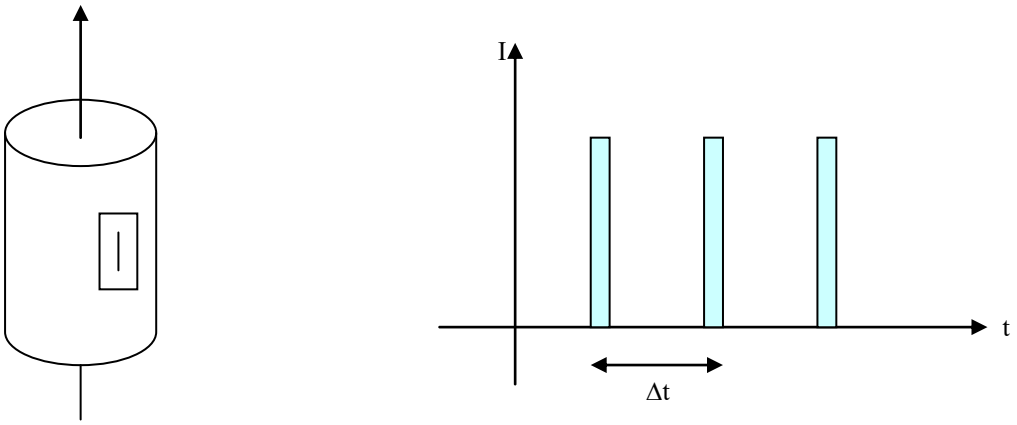


In order to have a full 180 degrees FOV, at least two sensors must be used with overlap of their individual FOVs.

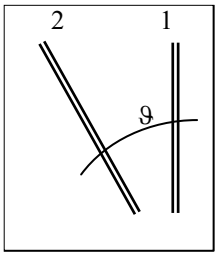


Use of Sun presence sensors on board simple spin satellites

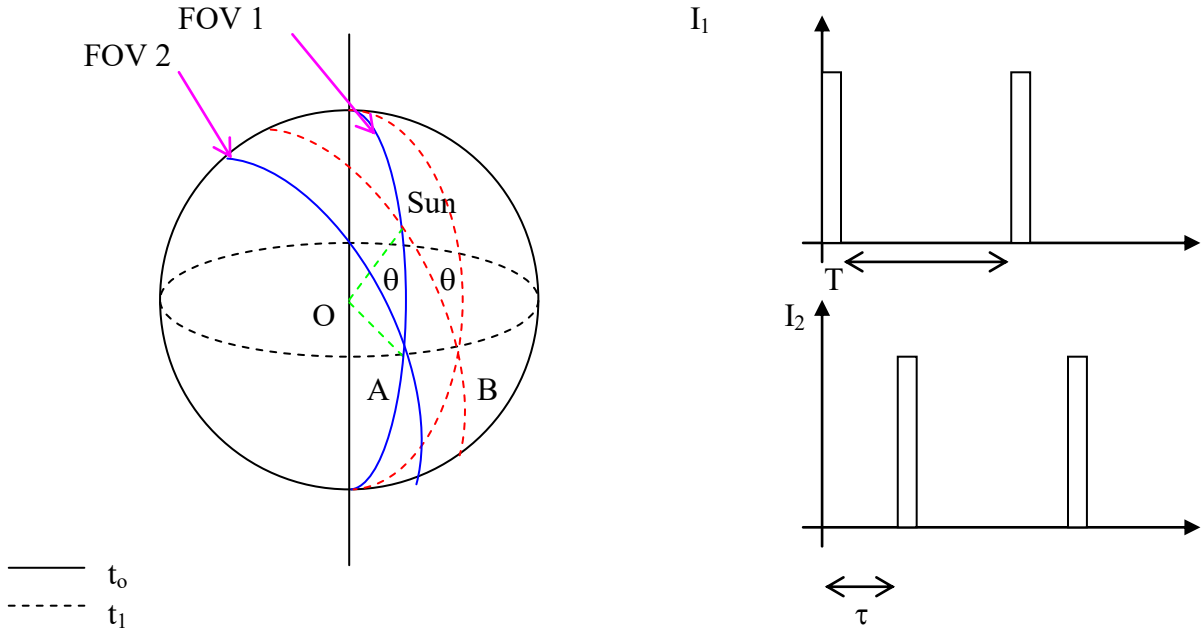
The following example shows the possibility to use a Sun presence sensor to infer the Sun angle when mounted on a simple spin satellite. Consider the simple spin satellite with a two-slit Sun presence sensor mounted on the lateral surface. At every rotation, a signal is produced by the sensor if the Sun is in the sensor FOV.



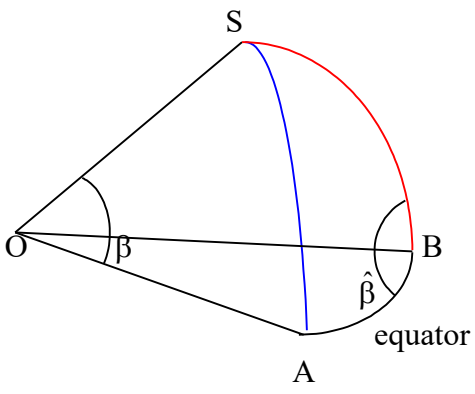
The Sun angle can be calculated by the analysis of the time interval between two consecutive electric pulses produced by a pair of two-slit sensors, with relative inclination ϑ.



Looking at the celestial sphere centered on the satellite, we can draw the following geometry of the problem:



The reference system is centered on the sensor and the equator is orthogonal to the sensor itself. Notice that the line of sight of sensor 1 is orthogonal to the equator. The interval T is connected to the spin angular velocity of the satellite, while the interval τ depends on the spin angular velocity and on the slit relative inclination ϑ. Consider the spherical triangle OABS:



We have:

$$\omega = 2\pi/T$$
$$\mathrm{AB} = \omega\tau$$
$$\hat{B} = \frac{\pi}{2} - \vartheta$$
$$\hat{A} = \frac{\pi}{2}$$

Using the cosine theorem for spherical triangles we evaluate β. Hence, measuring τ and ω it is possible to evaluate the Sun angle β, considering that some limitations to the general FOV of the sensor exist due to the relative inclination between the two slits (angle ϑ).

Starting from the measurement τ, the equations to solve the problem are:

$$\mathrm{tg}\left(\frac{\pi}{2} - \beta\right) = \frac{\mathrm{tg}\vartheta}{\sin(\omega\tau)}$$

$$\mathrm{cotg}\ \beta = \frac{\mathrm{tg}\ \vartheta}{\sin(\omega\tau)}$$

In a simulation environment, it is also possible to analyze and predict the sensor output, so evaluating the time interval τ as a function of the problem geometry, with the following equation

$$\tau = \frac{1}{\omega}\sin^{-1}(\mathrm{tg}\beta\ \mathrm{tg}\ \vartheta)$$

This is useful to define the minimum sampling time of the signal, sufficiently lower than the interval τ. Considering also some possible construction and mounting errors for the sensor, the above equation should be modified as:

$$\mathrm{cotg}^2\beta = \left[\frac{\mathrm{tg}(\vartheta + \Delta\vartheta + \varepsilon) - \mathrm{tg}\varepsilon\ \cos(\omega\tau - \delta)}{\sin(\omega\tau - \delta)}\right]^2 + \mathrm{tg}^2\varepsilon$$

where:

$\Delta\vartheta$	error on the relative angle between slits (not known)
δ	alignment error of slit 1 and spin axis
ε	in-plane rotation angle error of slit 1

In general, the simplest equation is used as a solution method, while the more complete model is used only to model the sensor and infer the possible operating errors, in connection also with the sampling frequency of the signals.

Horizon sensors (Earth sensors)

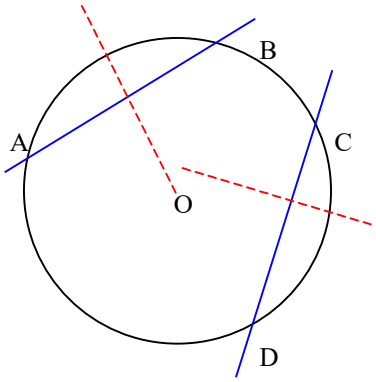
If the spacecraft is orbiting close to an extended body, such as a planet, the information on the presence of the planet within the FOV of a sensor is not sufficient to have information on the satellite attitude. To do so, the precise position of the center of the planet has to be evaluated. Horizon sensors work by analyzing the areas of the FOV that are not illuminated (deep space) compared to the illuminated areas (planet), and in general do not work in the visible spectrum to reduce the effects of interferences. It is common practice to design infrared sensors, for which the ratio of the Earth radiation compared to Sun radiation is "only" 1 to 400 (it is 1 to 30.000 in the visible spectrum). For Low Earth Orbits, the Earth can occupy 40% of the celestial sphere.

Earth horizon sensors use thermopile detectors to compute the direction vector of the Earth with respect to the spacecraft. A thermopile is a device for measuring the power of incident electromagnetic radiation via the heating of a material with a temperature-dependent electrical resistance. This sensor operates in one of two principal modes. The first is based on scanning the Earth's horizon to locate the centre of the Earth using a narrow field of view sensor (2 – 7 degrees); the second mode is based on static determination of the location of the Earth's contour inside the instrument's field of view (FOV is slightly larger than the Earth). Both kinds of sensors will be briefly explained.

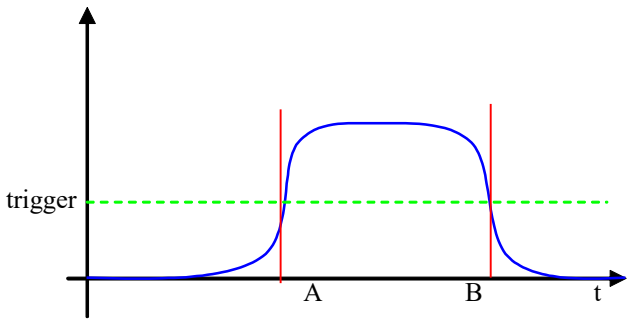
Scanning sensor

Scanning Earth horizon sensors use a spinning mirror or prism and focus a narrow beam of light onto a sensing element usually called a thermopile. The spinning causes the device to sweep out the area of a cone, and electronics inside the sensor detect when the infrared signal from Earth is first received and then lost. A factor that plays into the accuracy of such sensors is the fact the Earth is not perfectly circular.

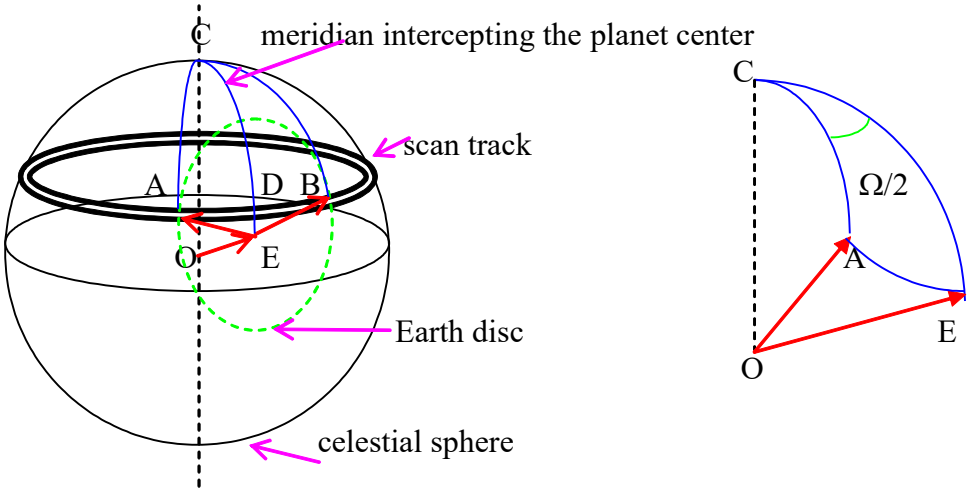
One option to calculate the planet center position is to detect two planet chords, assuming the planet is a sphere.



The sensor must be capable of detecting 4 points on the horizon, and associate the points to build the two chords. In general the sensor has a narrow field of view (2 degrees), but it can scan the horizon to produce the required output. When the sensor FOV is on the planet, the output is non-zero, when the sensor FOV is in deep space the output is zero.



Given a reference frame on the satellite, position of points A and B is known, and in similar way positions of points C and D and then the planet center. There are obviously some critical issues, since the boundary of the planet (points A and B) might not be so neat due to the transient nature of the sensor output induced by planet atmosphere. This can be solved by choosing an appropriate trigger level of the signal, above which we consider the signal as correct and below which we assume there is no signal. Points A and B will then correspond to the adopted trigger value. This value should be not fixed, but allowed to vary as a function of the peak signal or on the average of the signal itself, in order to consider also seasonal changes in the radiation emitted or day/night transitions. It is also to notice that the intensity of the sensor signal is not constant, due to change in the distance from ground as the sensor scans the horizon.
The center of the Earth can be detected using spherical geometry. Assuming the Earth is a perfect sphere, knowledge of the orbit altitude provides the information of the Earth apparent radius, and so that only one chord becomes necessary. In the following picture the case of scanning above the equatorial plane is illustrated.



For the spherical triangle CAEO we know $\overline{CA}$ and angle $\hat{C}$. Arc $\overline{AE}$ is the apparent radius of the Earth, known from the altitude, and calling θ the inclination of the scan direction with respect to the satellite horizon we have:

$$\overline{\mathrm{CA}} = \gamma \quad (\text{latitude}) = \frac{\pi}{2} - \theta$$
$$\hat{C} = \frac{\Omega}{2} = \frac{\omega \Delta t}{2} \text{ detected bysensor} = \text{arcAD}$$
$$\overline{\mathrm{AE}} = \rho$$

From spherical triangle geometry we then have:

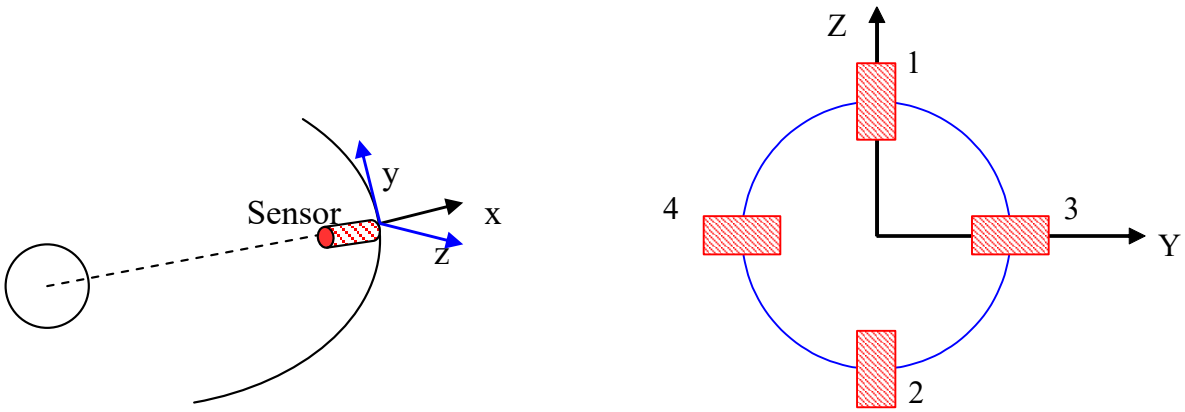
$$\frac{\Omega}{2} = \cos^{-1}\left(\frac{\cos\rho - \cos\gamma\cos\eta}{\sin\gamma\,\sin\eta}\right)$$

from which $\eta = \overline{CE}$

This type of sensor is rather heavy due to the presence of precise scan mechanism, that must guarantee a scan velocity much higher that the spin velocity of satellite.

Static sensor

Static Earth horizon sensors contain a number of sensors and sense infrared radiation from the Earth's surface with a field of view slightly larger than the Earth. The accuracy of determining the geocenter is 0.1 degrees in near-Earth orbit to 0.01 degrees at GEO. Their use is generally restricted to spacecraft with a circular orbit. An example configuration of 4 thermopile sensors is given below.

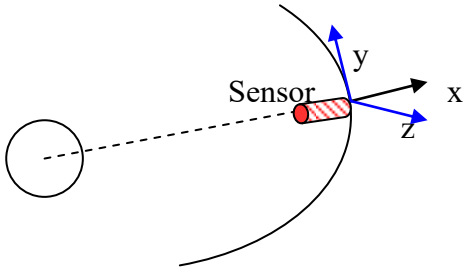


When the Earth is in the center of the FOV, the four detectors have all the same output, so that:

$$(1) - (2) = 0$$
$$(3) - (4) = 0$$

When the image of the Earth is not centered, the four detectors have different outputs and taking appropriate differences we have direct information of the misalignment angles:

$$(1) - (2) = \Delta z \text{ proportional to } \alpha_y$$
$$(3) - (4) = \Delta y \text{ proportional to } \alpha_z$$



Given these two angles we can then compute the Earth horizon vector using spherical coordinates with $\hat{h}_B = [cos\,\alpha_y \quad sin\,\alpha_z\,sin\,\alpha_y \quad -cos\,\alpha_z\,sin\,\alpha_y]^T$

Given that we can also have an orbital model on-board or from our orbit determination we also have the direction vector of the Earth with respect to the spacecraft measured in the inertial frame $\hat{\underline{h}}_N$. Therefore, we have

$$\hat{\underline{h}}_B = A_{B/N}\hat{\underline{h}}_N$$

We could parameterize $A_{B/N}$with Euler angles so that there are only three unknowns and then we would have three nonlinear equations and three unknowns. Since it is nonlinear we can try to solve for the Euler angles using a numerical optimization scheme that minimizes the function:

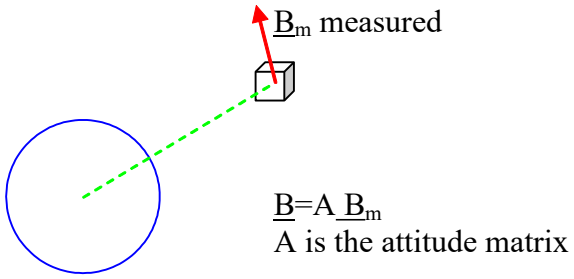
$$J = \left\|\hat{\underline{h}}_B - A_{B/N}\hat{\underline{h}}_N\right\|$$

However, since $\hat{\underline{h}}_B$and $\hat{\underline{h}}_N$ are unit vectors the problem is underdetermined, so it cannot be solved. This is intuitive since with a direction vector only, we cannot know the angle of the roll axis. So we cannot retrieve the attitude matrix just with an Earth horizon sensor.

For simple spin or dual spin satellites, with nominal angular momentum aligned with the z axis, measurement of α_z is always required since the z equation is decoupled from x and y equations. The x and y equations are instead coupled, so that measuring only α_x or α_y we can in any case reconstruct the satellite attitude with a state observer. Therefore this sensor, that measures the two angles α_y and α_z, is suitable for a full state observation in the linear satellite dynamics model.
This sensor is not capable of detecting rotations around the yaw axis (x), and furthermore the optical head must be designed in order to have a FOV that is twice the Earth apparent radius. Therefore this sensor is commonly used only for circular orbits, at constant altitude, and on Earth pointing satellites.

Magnetic field sensor

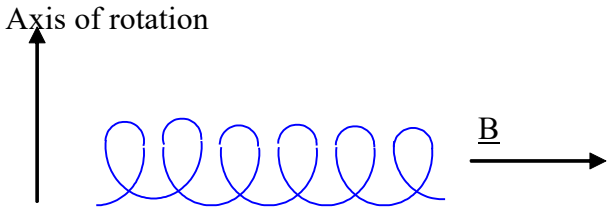
This sensor can provide the measurement of one vector. It has to be coupled with a mathematical model of the magnetic field and to the knowledge of the position of the satellite in order to provide attitude information.



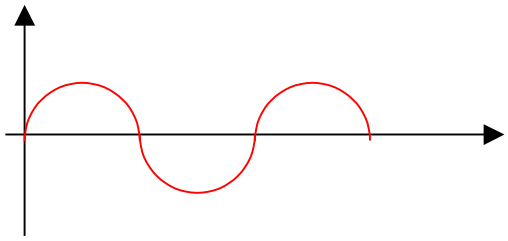
It is not simple to measure a constant magnetic field; in fact typically the rate of change of the flux of the magnetic field vector can be measured by a coil. Indicating with dove ϕ_B the flux of the magnetic field and μ the magnetic permeability, we have

$$I = \frac{d\varphi_B}{dt}\mu$$

This could be in principle obtained by having one coil rotating around a defined axis

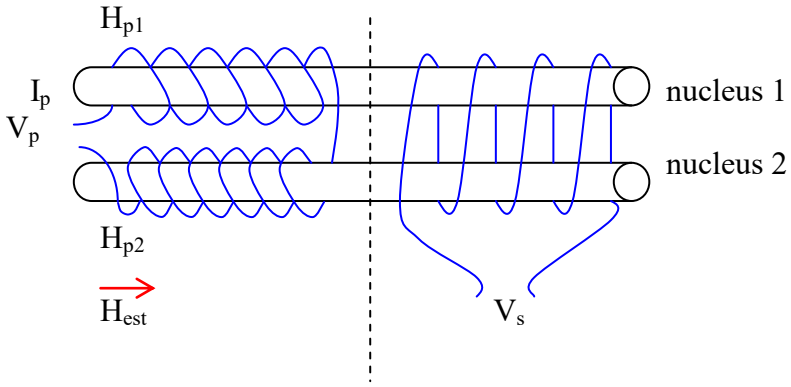


For a constant B vector the measurement would be sinusoidal



and the derivative of the measured signal is the projection of the B vector onto the coil axis, rotating inside the satellite. This procedure is complicated and provides a measurement not directly useful for attitude determination purpose, therefore alternative sensors are built, called fluxgate magnetometers.

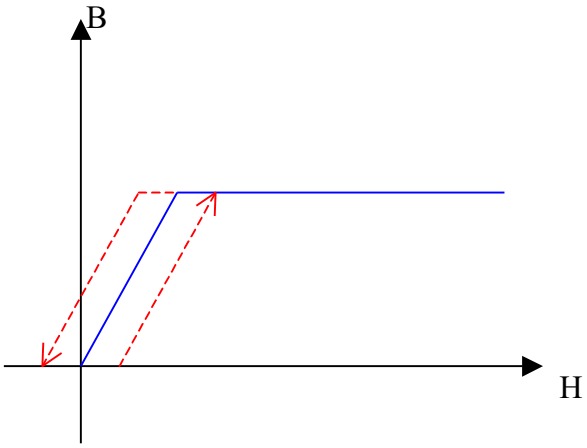
Consider two nuclei of ferromagnetic material, with two coils winded as shown in the figure below



On the left side the coil, called primary coil, is winded in such a way to produce on the two nuclei exactly the same magnetic induction but with opposite sign. Since the voltage at the secondary coil (on the right) is

$$V_s = -\frac{d\varphi(B_1 + B_2)}{dt}$$

In the case of zero external magnetic field the voltage would be zero. If the external magnetic field is non-zero, it is added to the induction due to the primary coil and the voltage at the secondary coil would be non-zero.
The coils create a magnetic induction field that generates the magnetic field B. This is normally a nonlinear mapping, which for sake of clarity will be modeled as linear with saturation, neglecting nonlinearity and hysteresis.

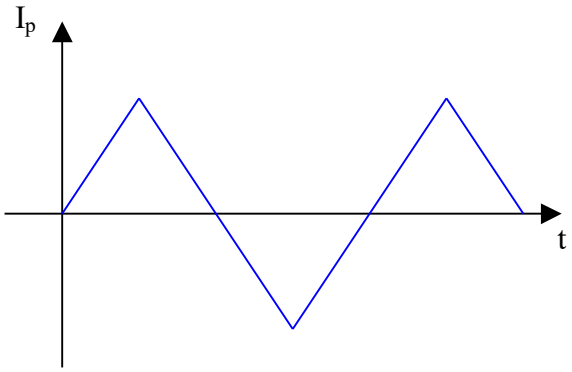


The current flowing in the primary coil is usually generated with a triangular wave form, so that the generated magnetic induction will also have a triangular waveform.

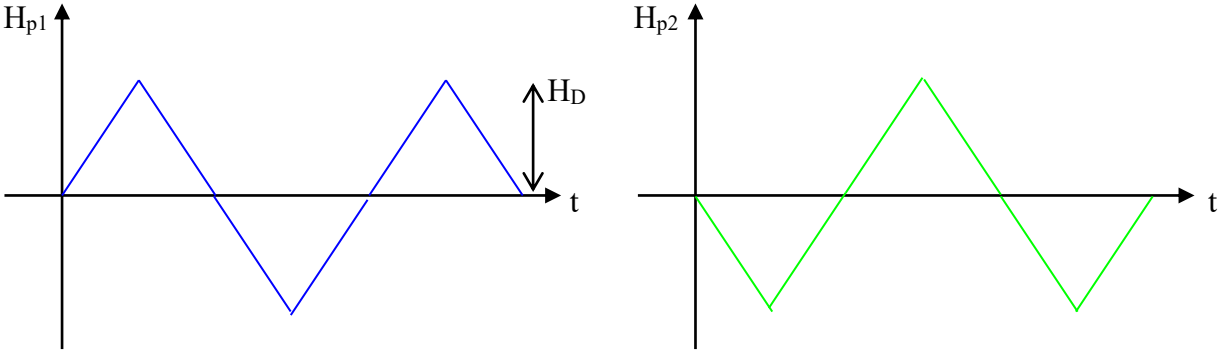
$$V = L\frac{di}{dt} = -\frac{dH}{dt}$$

It will be clear later on that the triangular waveform as input generates a pulsed output, easy to convert into a magnetic field measurement.

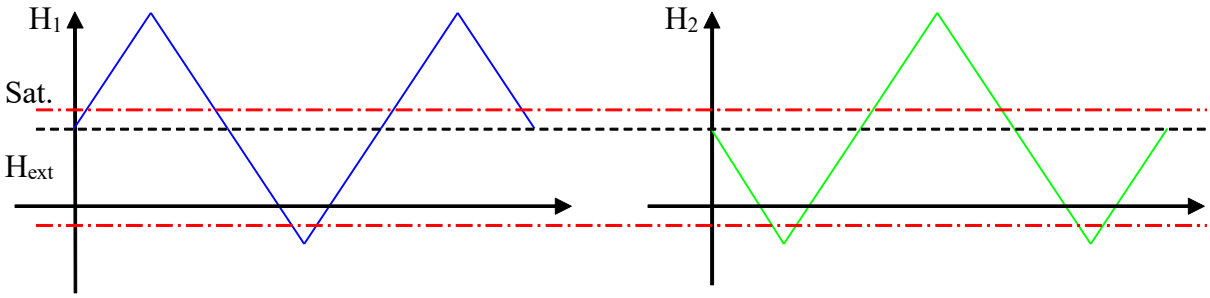
Define the following input to the primary coil



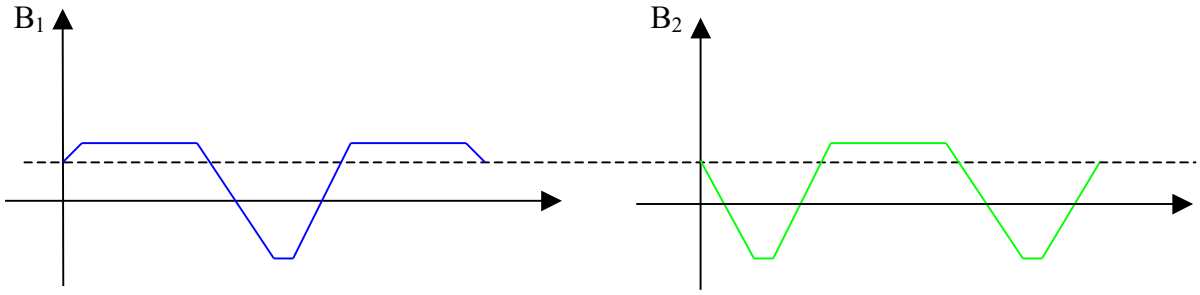
The two magnetic induction in the two nuclei are



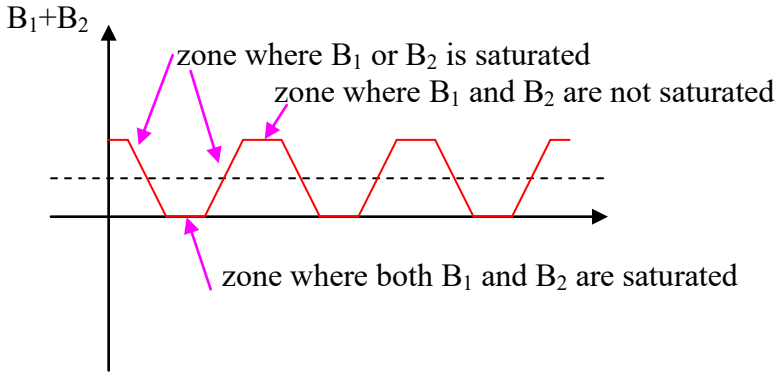
Adding the effect of the external magnetic field we get to



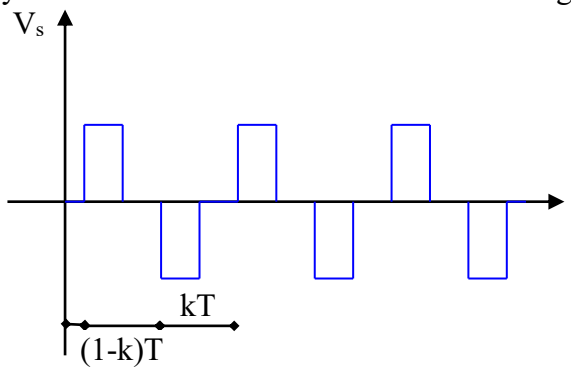
The magnetic field in the two nuclei becomes, considering nucleus saturation



We can now add the two magnetic fields generated, to analyze what is included in the secondary coil on the right part of the sensor.



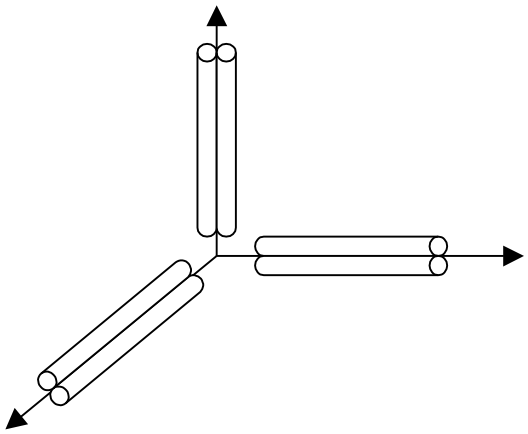
The voltage at the secondary coil is the derivative of the flux of the magnetic field vector induced



The spacing between the impulses is related to the component of the external magnetic field in the direction of the coil axis. The amplitude of the impulse is instead depending on the peak current in the primary coil.

$$H_{\text{ext}} = (1 - 2k)H_D$$

Since the principle of operation of this sensor relies on the saturation of the two nuclei, the value of the current in the primary coil must consider the expected value of the external magnetic field. In case no saturation of the two nuclei occurs, or if one nucleus is permanently saturated and the other never saturated, then the sensor would produce no meaningful output.
Using three pairs of nuclei, as shown below, we can measure the components of the magnetic field along three orthogonal directions, so the B vector.



They have an accuracy of about 300 arc minutes (5 degree) at 200km altitude to about 1 degree at 5000km. They are the least accurate attitude reference sensors as there is a lot of uncertainty in the modelling and measurement of the magnetic field. The magnetometers are also subject to noisy measurements and interference from on-board electronics. Variations in the solar wind, solar storms can lead to large inaccuracies in the on-board model.
Note that this measurement is taken in a body axes frame ($\underline{b}_B$), while the mathematical model of the magnetic field produces a vector expressed in a local horizontal frame ($\underline{b}_H$ or $\underline{b}_{r,\varphi,\vartheta}$). Detecting the rotation that maps one vector onto the other leads to the attitude determination in the local horizontal frame. In principle, we can think about looking for the solution to this problem

$$\left\| \underline{b}_B - A_{B/H}\underline{b}_H \right\| = 0$$

This can be solved using a numerical optimization procedure that seeks to find A that minimises the function

$$J = \left\| \underline{b}_B - A_{B/H}\underline{b}_H \right\|^2$$

If we are using DCM then to guarantee that the solution is orthonormal, a constrained optimal control problem must be solved, such that *J* is minimized subject to the constraints:

$$A_{B/H}^T A_{B/H} = I, det(A_{B/H}) = 1$$

There is however always uncertainty in the rotation around the direction of the measured $\underline{b}_B$. This is the reason why the sensor is always coupled to another measurement in order to provide reliable attitude determination.

Star sensors

The best precision attainable with Sun sensors is approximately 1/8 of degree, and in addition the Sun might not be always visible if the orbit has a phase in eclipse. Star sensors allow improving the accuracy of attitude determination and, with appropriate choice of the stars observed, no eclipse problem arises. On the other hand, star sensors are extremely expensive, in terms of constructions and in terms of operations to be performed on board for data processing.

Star sensors have to perform the following operations:

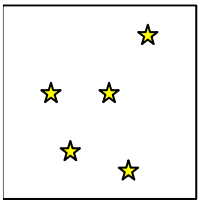
Phase 1: data acquisition
Phase 2: correct positioning in space of the acquired data
Phase 3: interpretation of the data acquired (star identification) that requires analysis of a star catalogue
Phase 4: attitude determination

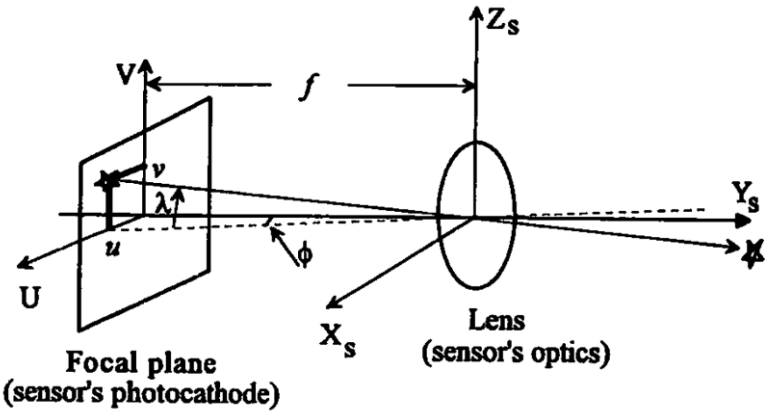
The interpretation of the data is an extremely consuming activity in terms of computation resources. It requires the comparison of the sensor data with a map of the stars, available on board the sensor, in terms of light intensity (magnitude of the star) and radiation spectrum.

Star sensors are intrinsically precise sensors, but to achieve maximum accuracy the sensor has to be also extremely sensible, so that spurious radiation can deteriorate the performances and the operating

lifetime of the sensor. A Sun presence sensor can be used to switch off the star sensor when its FOV is too close to the Sun, and light shields can also be used to limit interference with other sources of light (even Earth radiation can be source of disturbances).

In the case of star-sensors the number of measurements available are often much greater than 3 (typically 20 spacecraft-star measurements are given). Each star direction vector can be computed by detecting the point at which it intersects the focal plane. Modern sensors are built with a matrix sensor (CCD), on which the star image is projected.





The direction of the star is computed from the projection of the image on the focal plane

$$\tan\phi = \frac{u}{f} \quad ; \quad \tan\lambda = \frac{v}{f}\cos\phi \quad ; \quad \hat{O} = [-sin\phi\, cos\lambda \quad cos\phi\, cos\lambda \quad -sin\lambda]^T$$

It is pointed out that, even with narrow FOV, the stars captured by the sensor might be numerous. If the sensor has tilting capacity it can detect one star and tack it so that it is always in the center of the FOV. This sensor is called star tracker.

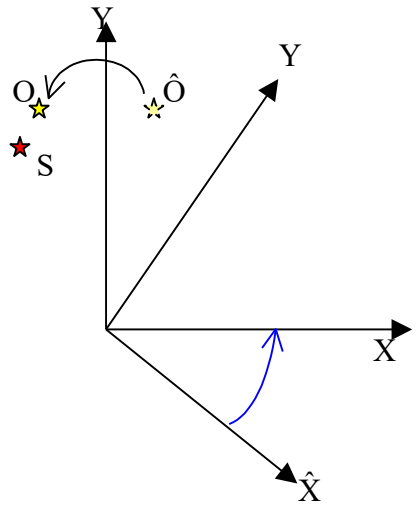
A different possibility is to keep the sensor fixed on the satellite and to reconstruct the map of the observed sky, in what is called a star mapper sensor. In this case the problem is to have a sufficiently fine CCD sensor.

Attitude determination requires first the identification of the stars observed; now we look at how this process can be completed.

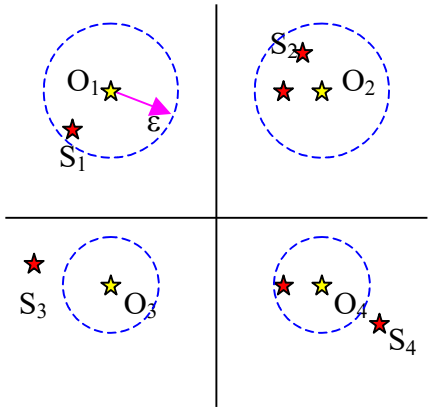
Assume to have S stars in the catalogue and Ô observations available, the goal is to associate each observation (in the sensor reference frame) with one star of the catalogue (in the inertial frame).

$$\begin{array}{ccc} \hat{O}_1 \overset{A}{\rightarrow} O_1 \rightarrow S_1 & & \\ \hat{O}_2 \overset{A}{\rightarrow} O_2 \rightarrow S_2 & & \\ \hat{O}_3 \overset{A}{\rightarrow} O_3 \rightarrow S_3 & & \\ \vdots \quad \vdots \quad \vdots & & \\ \hat{O}_n \overset{A}{\rightarrow} O_n \rightarrow S_n & & \hat{O}_i \quad \text{local reference} \\ & & O_i \quad \text{inertial reference} \end{array}$$

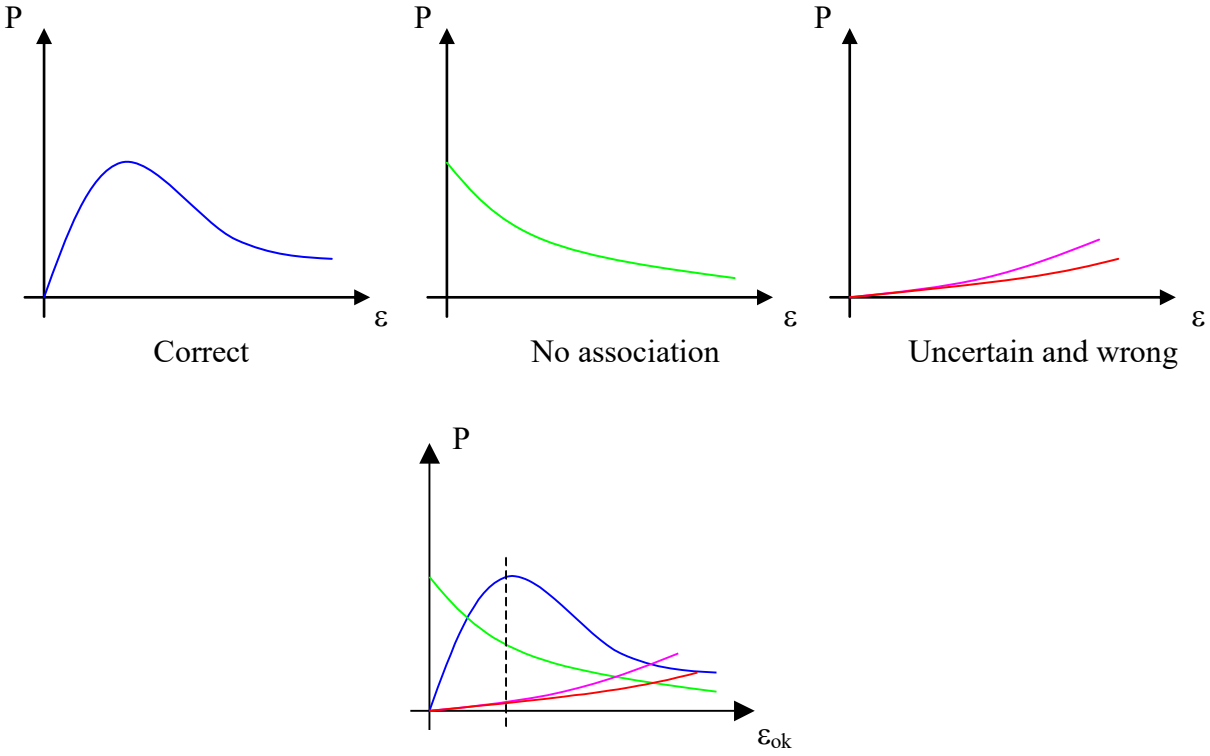
The process requires a first guess estimation of the attitude matrix A, to refine after the correct association of the observed stars with the catalogue.



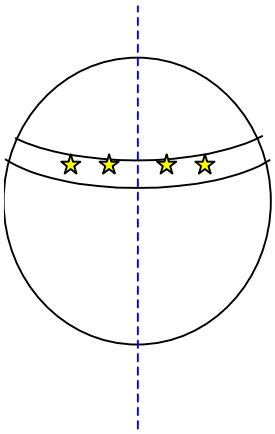
Once the observed star Ô is rotated in the catalogue system (O), the distance between O and S is used to correct the estimation of the attitude A, but this operation is correct only if we are sure that O is the correct image of S and not of a different star. To have a higher probability of success, the identification can be based also on relative geometry of a set of stars.



First, the star closer to O_1 is searched until one is found. If the association is correct, the remaining image stars should be immediately associated to the respective catalogue stars. This is called direct matching of stars. Some problem might arise in the case represented in the case O_2 in the above figure, where a multiplicity of stars can be associated to the observation, or the case O_3 in the above figure, where no stars in the catalogue are sufficiently close to the observation (no identification). The worst case is however the case 4 above, where the wrong star in the catalogue is closer to the observation, leading to an erroneous association. If the search area is enlarged (increase of neighborhood ε), the situation of O_3 is improved, leading to possible correct association, but as a general rule the number of uncertain associations or wrong associations might increase as ε increases. The optimal search radius ε maximizes the probability to have, for each observed star, a unique and correct association with the catalogue.

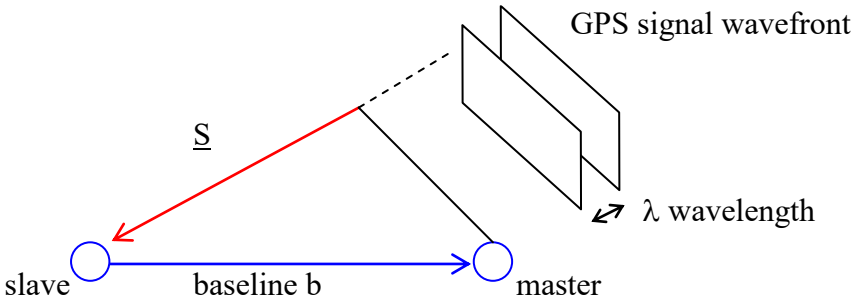


The optimal value of ε depends on the completeness of the catalogue and on the density of stars of the same magnitude in the observed area. Sky areas with a diffuse radiation can create problems since the sensor can erroneously associate one star to a portion of image, even if this is only diffuse background radiation.
A different method can be used if a circular scan of a portion of sky is performed, and if this image strip is compared to the catalogue. The correct attitude would match the highest number of image stars to the stars in the catalogue. This method is useful only if stars on the strip are not uniformly distributed, and if at least the altitude of the scan is known to reduce the search space within the catalogue.



Use of GPS sensors for attitude determination

GPS sensors can be used to determine attitude, provided at least 3 antennas are available.



Slave and master are two antennas mounted on the satellite. If their distance b is known we can detect its orientation in space. Taking the projection of b on the direction S of the incoming GPS signal, we have the path difference of the signal received by the two antennas.

$$\underline{S}^T b$$

The position of the GPS satellite is known, so that direction S is known in geocentric coordinates if we know the position of the master antenna. Measuring the path difference

$$\Delta r = \underline{S}^T b$$

and transforming the vector S from geocentric to body frame:

$$\underline{S} = AS$$
$$\Delta r = S^T A^T b$$

the unknown is the rotation matrix A. Notice that most of the times instead of measuring Δr we measure a phase difference:

$$\Delta\varphi = \Delta r \, 2\pi/\lambda$$

The problem can be solved if we have at least three independent measurements, that requires the use of at least two different baselines (3 antennas) receiving signals from at least two GPS satellites. In this case the available measurements are:

$$\Delta r_{11} = S_1^T A^T b_1 \quad \text{baseline 1 GPS satellite 1}$$
$$\Delta r_{21} = S_2^T A^T b_1 \quad \text{baseline 1 GPS satellite2}$$
$$\Delta r_{12} = S_1^T A^T b_2 \quad \text{baseline 2 GPS satellite 1}$$
$$\Delta r_{22} = S_2^T A^T b_2 \quad \text{baseline 2 GPS satellite2}$$

Note that two baselines are required; having three GPS satellites in view of one single baseline would not allow to determine the rotation around the baseline. To determine the attitude, the following function can be minimized:

$$J = \sum_{i=1}^{N.S} \sum_{j=1}^{N.b} \left(\Delta r_{ij} - S_i^T A^T b_J \right)$$

Note that when considering phase differences, the integer number of wavelengths K has to be correctly determined:

$$\Delta\phi = (\Delta r + K\lambda)\frac{2\pi}{\lambda} + \nu$$

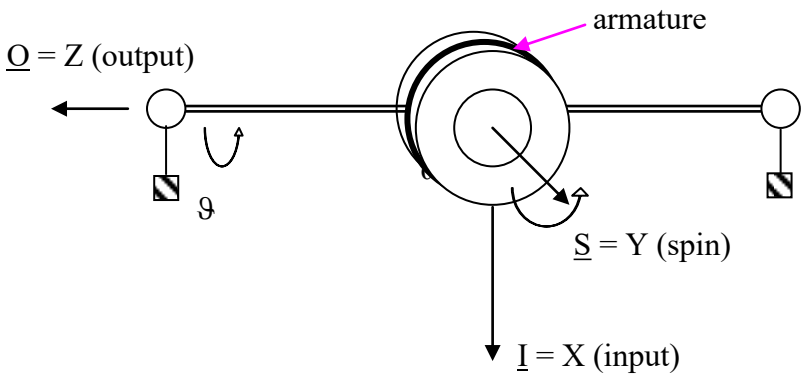
In addition, proper account of the measurement noise ν must be taken. This noise can have various sources, among which reflection of incoming signal is predominant.

Experiments performed to date have guaranteed a precision in the order of some fraction of degree.

Gyroscopes

Mechanical gyroscopes

Gyroscopes, or gyros, are used to measure angular velocities. The classical gyro is a mechanical sensor, with a spinning rotor and one degree of freedom in the rotor support mechanism.



The rotor spins around the S axis, while the support mechanism and the armature are free to rotate around axis O. The angular momentum in the XYZ body axes is then:

$$h_T = I_R \omega_R \underline{y} + J_z \dot{\vartheta} \underline{z}$$

I_R rotor

J_Z rotor + armature + support

We can write the Euler equation of the system, as:

$$\underline{\dot{h}} + \underline{\omega} \wedge \underline{h} = \underline{M}$$

where $\underline{\omega}$ is the composition of the angular velocity of the body on which the sensor is mounted, typically the satellite, and the relative angular velocity between gyro and satellite:

$$\underline{\omega} = \begin{Bmatrix} \omega_x \\ \omega_y \\ \omega_z + \dot{\vartheta} \end{Bmatrix}$$

The rotor angular velocity ω_r does not appear since the rotor is totally free inside the armature.

The Euler equations are then written as:

$$\begin{cases} \omega_y J_z \dot{\vartheta} - I_R \omega_R \left(\omega_z + \dot{\vartheta}\right) = M_x \\ I_R \dot{\omega}_R - \omega_x J_z \dot{\vartheta} = M_y \\ J_z \ddot{\vartheta} + I_R \omega_R \omega_x = M_z \end{cases}$$

The third equation is used to evaluate the angular velocity ω_x while the first two equations can be used to determine the support reaction forces. If M_z is zero, from the third equation we evaluate ω_x by measuring the second derivative of the output angle ϑ. In general, to avoid having a constant acceleration at the output axis, a restrain torque is present. Assuming this has the general viscoelastic form, we have:

$$M_z = -k\vartheta - c\dot{\vartheta}$$

The third equation then becomes:

$$J_z\ddot{\vartheta} + c\dot{\vartheta} + k\vartheta = -I_R\omega_R\omega_x$$

The steady state solution is then:

$$\overline{\vartheta} = -\frac{I_R\omega_R\omega_x}{k} \qquad\qquad \omega_x = -\frac{k\overline{\vartheta}}{I_r\omega_r}$$

Measuring the steady state rotation of the output axis we measure the angular velocity component along the input axis. This instrument is known as RATE GYRO (RG). In a similar way, if we avoid using an elastic support and have only a viscous torque at the output axis, the steady state output will be a constant angular rate, which can still be used to measure the angular velocity along the input axis:

$$\overline{\dot{\vartheta}} = -\frac{I_R\omega_R\omega_x}{c} \qquad\qquad \omega_x = -\frac{c\overline{\dot{\vartheta}}}{I_r\omega_r}$$

This sensor is known as RATE INTEGRATING GYRO (RIG).
The transient terms in the output equation have an influence on the response time of the sensor, since the assumption is made that measures are taken at steady state.

Three gyros can be used to measure the three components of the angular velocity. It is pointed out that, since the measure implies a rotation of the output axis and therefore also of the input axis, it is almost impossible to measure directly the angular velocity in principal axes. To avoid this, and measure the angular velocity along a fixed reference frame, a control torque M_z can be applied at the sensor output axis such that the output angle ϑ is zero. In this condition, the angular velocity is measured by measuring the control torque M_z, and the measure is always corresponding to the same input axis.

$$M_z(\vartheta) = I_R\omega_R\omega_x$$
$$\omega_x = \frac{M_z(\vartheta)}{I_R\omega_R}$$

An alternative solution to keep the input axis I fixed is to rotate the whole support platform, which in this case becomes an inertial platform. Measuring the rotation angles of the platform with respect to the satellite we would directly retrieve the information of the inertial attitude of the satellite.

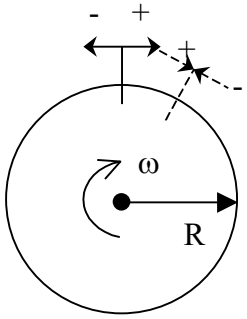
The precision of the gyro depends on the term $I_R\omega_R$; the larger it is, the smaller is the full scale of the sensor but the sensor would be more precise. Furthermore, the bandwidth of the sensor depends on J_z, since

$$\lambda = \sqrt{\frac{k}{J_z}}$$

To have a fast sensor we need a high bandwidth and so to reduce J_z and as a consequence also I_R. Also, ω_R can be increased, but this can be a problem above a certain limit due to friction and dissipation.
The mechanical gyro is suitable to measure relatively high angular velocities, for low velocities the nonlinearities and the internal friction make the measurement highly uncertain.

Laser gyroscopes

A conceptual scheme of a laser gyroscope can be reduced to an optical path in which a signal is injected by a transmitter. The same signal is then retrieved by a receiver at the end of the optical path. For simplicity, the description will assume a circular optical path.



The transmitter produces two signals (+ and -) in two opposite directions of the optical path. Since the gyroscope is rotating, the receiver, collocated with the transmitter, will receive the two signals at different time instants since the two signals will have travelled for different lengths.

$$ct^- = (2\pi - \omega t^-)R$$
$$ct^+ = (2\pi + \omega t^+)R$$

The instants in which the signals are received are then:

$$t^- = \frac{2\pi R}{c + \omega R}$$
$$t^+ = \frac{2\pi R}{c - \omega R}$$

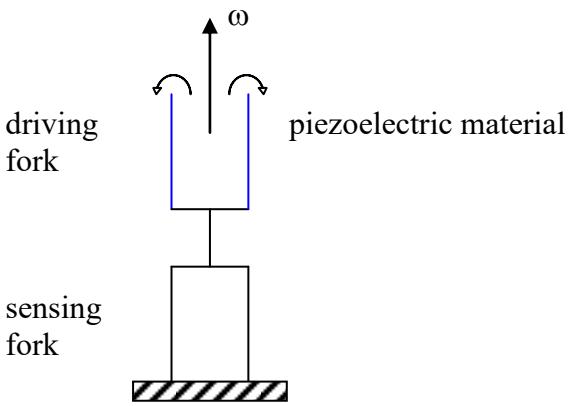
Measuring the time difference we compute the angular velocity:

$$\Delta t = (t^+ - t^-) = \frac{2\pi R(c+\omega R) - 2\pi R(c-\omega R)}{(c^2 - \omega^2 R^2)} = \frac{4\pi R^2 \omega}{(c^2 - \omega^2 R^2)} \cong \frac{4\pi R^2 \omega}{c^2} = \frac{4A\omega}{c^2}$$
$$\omega = \frac{\Delta t c^2}{4A}$$

This instrument is known as RING LASER GYRO (RLG). If the optical path is made of a fiber optic, it is possible to wind the fiber optic for more than one winding, improving the instrument sensitivity. This instrument is known as FIBER OPTIC GYRO (FOG). These gyroscopes measure the angular velocity in a reference frame fixed with the satellite, and are not subject to friction and motion problems.

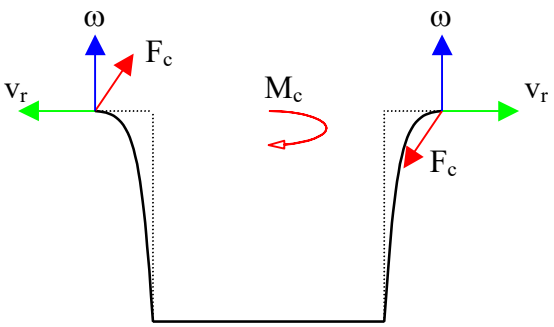
Piezoelectric gyroscopes

One further scheme of angular velocity sensor is based on the use of piezoelectric materials. Two forks made of piezoelectric materials are used. If subject to an electric field, the fork is deformed, while if deformed it produces an electric signal.



If the sensor rotates with velocity ω, and the two prongs are electrically driven in phase opposition, the Coriolis force, perpendicular to ω and the tip velocity, generates a torque. The torque is transmitted to the base prongs that, due to their deformation, generate a second electric signal that can be measured.

$$F_c = -2\underline{\omega} \wedge \underline{v_r}$$



Also this sensor measures the angular velocity along a fixed direction, and has no moving parts. However, as in part also true for laser gyros, it is extremely sensitive to temperature oscillations, which cause a thermal deformation that has an effect on the scale factor of the sensor.

Simple Gyro noise modelling

Solid-state gyroscopes output noisy measurements, that can be modelled as

$$\underline{\omega}_i^M = \underline{\omega}_i + \underline{n} + \underline{b}$$

Where $\underline{b}$ is the gyro bias

$$\underline{n} = \sigma_n \underline{\zeta}_n$$
$$\dot{\underline{b}} = \sigma_b \underline{\zeta}_b$$

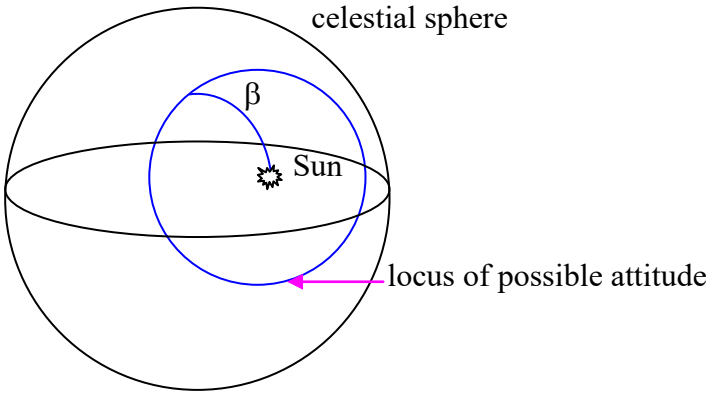
$\sigma_n \underline{\zeta}_n$ is white Gaussian noise with zero mean and standard deviation σ_n . This noise is known as angular random walk (ARW) attributed to thermo mechanical noise of the system and $\sigma_b \underline{\zeta}_b$ is white Gaussian noise with zero mean and standard deviation σ_b known as rate random walk. (RRW) attributed to electronic noise (flicker). These are the main contributions of noise in a gyro and manufacturers will often provide values for ARW and RRW.

Attitude determination

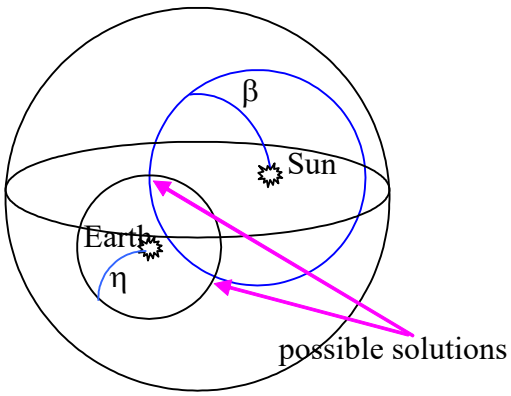
The sensors in general provide either a measure of one angle or of a direction in space, which must be converted into an attitude determination of one axis, for some simple spin satellites, or of three axes. Attitude determination methods are classified as geometric methods, algebraic methods or statistical methods.

Geometrical methods

Let suppose that one angle of the Sun or of Earth is known in the principal inertia reference, then the orientation in space of the sensor axis can be determined as a circle on the celestial sphere. This means the optical axis of the sensor has been placed on a geometrical locus.

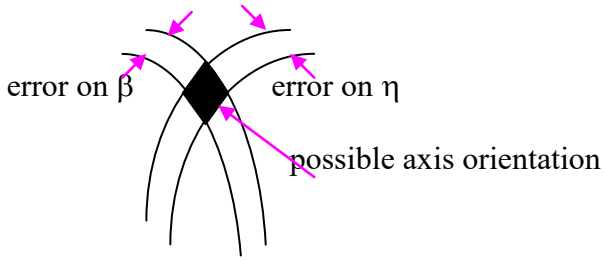


This because any rotation by the angle β starting from the great circle would rotate the optical axis of the sensor to match the direction of the Sun. If a second measurement of the same kind, referring to a different celestial body, is available, the intersection of the two cones provides the attitude of one axis of the satellite in inertial reference. Two solutions are obtained.

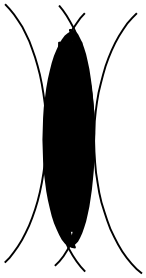


If the satellite is spinning, taking two sets of measurements with a given time separation it is expected that the orientation of the spin axis will not change. The real solution will then be the one that moves less in the inertial space, while the two circles will rotate around this point, since the relative geometry of the satellite with respect to Sun and Earth will change. The solution depends in the end on sensor geometry.

In the real situation, the intersection is not between two lines, but between two narrow strips with amplitude given by the measurement errors. The exact orientation of the spin axis will then be determined inside a small square surface.

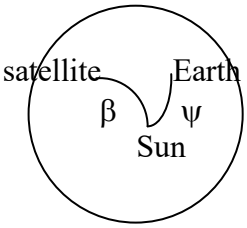


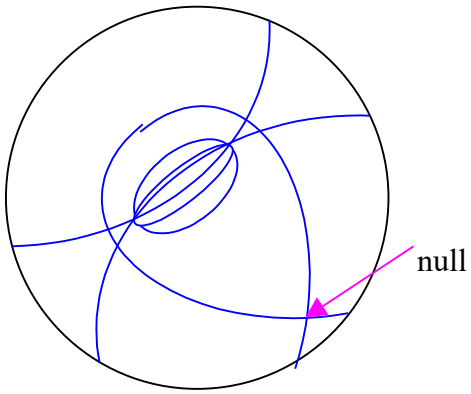
The error on the identified position of the spin axis is then higher than the error on a single measurement. If in addition the two strips do not intersect orthogonally, along one direction the error can be greatly amplified:



This situation depends on the relative position of the two detected celestial bodies, and the worst situation corresponds to the two bodies in the equatorial plane of the satellite. Notice that this event is totally independent from the sensor precision, and to minimize it the only solution is the appropriate selection of the celestial bodies to detect.

A second alternative is based on the measurement of one absolute angle and one relative angle between two celestial bodies. The shape of one locus will be still a circle on the celestial sphere, while the second locus is rather complex.





The actual shape of the second locus depends on ψ. The null point is a singular condition that appears when the angle ψ is equal to the angle β.

Algebraic methods

Knowledge of two angles (two scalar measurements) allows detecting the orientation of one axis in space. The complete knowledge of the attitude is yet undetermined, and requires at least one further scalar measurement to fix the phase of the reference frame around the detected axis.

With the exception of the GPS sensor, that can provide directly the attitude in the GPS reference system, or the fixed Earth sensor that provides the two small angles α_y and α_z, from which a state observer can reconstruct the complete attitude, the general case requires measures of two or more directions in space (celestial bodies) to correctly determine the attitude.

Calling s_i the unit vectors measured by the sensors, in body-fixed frame, v_i the unit vectors of the same celestial bodies but referred to an inertial frame or in any case a given reference frame, the following holds:

$$s_i = Av_i$$

The solution in terms of A requires at least two independent measurements, and the method of solution can differ for the cases with two measurements (s_1 and s_2), three measurements (s_1, s_2 and s_3) or more than three measurements. The simplest case refers to 3 measurements:

$$[s_1 \quad s_2 \quad s_3] = A[v_1 \quad v_2 \quad v_3]$$

If the measurements are independent, V is non-singular so that:

$$S = AV$$
$$A = SV^{-1}$$

V is invertible provided that v_i are independent, which means that the 3 measurements v_i refer to non-coinciding directions in space. This means that the selection of the target celestial bodies must be appropriate, and some degrees of freedom can also be exploited in order to improve the numerical accuracy of the solution. The best condition refers to 3 orthogonal measurements.

In case the number of measurements available is greater than 3:

$$[s_1 \quad s_2 \quad s_3 \quad s_4] = A[v_1 \quad v_2 \quad v_3 \quad v_4]$$

The system is redundant, so that if we want to make use of all available measurements the solution requires the evaluation of the pseudo inverse of matrix V:

$$SV^T = AVV^T$$
$$SV^T(VV^T)^{-1} = A$$
$$V^* = V^T(VV^T)^{-1}$$
$$A = SV^*$$

An alternative method would require the selection of the best three measurements and revert to the previous method; nevertheless using all available measurements can provide a better filtering of random errors.

In case only two measurements are available, the solution can be obtained with the following algorithm. Call the two measurements p and q, and call a and b their corresponding directions in inertial space. Associate the measurement p to unit vector s_1 and the corresponding direction a to unit vector v_1. Two orthogonal frames can then be associated to the measurements and to the reference directions as follows:

$$s_1 = p$$
$$s_2 = \frac{p \wedge q}{|p \wedge q|}$$
$$s_3 = p \wedge s_2$$

$$v_1 = a$$
$$v_2 = \frac{a \wedge b}{|a \wedge b|}$$
$$v_3 = a \wedge v_2$$

Since by construction the 3 unit vectors s_1, s_2, s_3 and v_1,v_2, v_3 are orthogonal, we can write:

$$V^{-1} = V^T$$
$$A = SV^{-1} = SV^T$$

To minimize errors, vector p should be measured with the maximum possible precision and q should be as orthogonal to p as possible. If p and q are almost aligned, as could be the case with one Earth sensor and one magnetometer in proximity of the poles, the attitude could not be determined.

Some filtering of the errors can be provided if we take:

$$p^* = \frac{p + q}{2}$$
$$q^* = \frac{p - q}{2}$$

And then start the attitude determination algorithm. In practice, this method detects first of all the plane containing the two measurements and then the perpendicular direction to the plane.

Statistical methods

Alternative methods to determine the attitude need at least two vector measurements and knowledge of the relative precision of the sensors, in order to build and minimize a suitable weighted error function:

$$J = \frac{1}{2}\sum_{i=1}^{N} \alpha_i |s_i - \mathrm{Av}_i|^2$$

where N is arbitrary and α_i is a constant identifying the sensor precision, more precise sensors are associated to the higher α_i parameters. In the following, without loss of generality, it will be assumed that the vector of weights α is normalized to 1

$$\sum_{i=1}^{N} \alpha_i = 1$$

The real solution is the one minimizing J. In this approach, we take for granted the impossibility to solve exactly the problem when measurement errors are considered, so we try to find the optimal solution that minimizes the cumulative sum of errors. A minimum of two independent measurements are needed.

$$J = \frac{1}{2}\sum_{i=1}^{N} \alpha_i (s_i^T s_i + v_i^T A^T \mathrm{Av}_i - 2\mathrm{s}_i^T \mathrm{Av}_i)$$

Since A is a rotation matrix, A^TA is identity matrix, and since s_i and v_i are unit vectors the first two terms are equal to 1, so that:

$$J = 1 - \sum_{i=1}^{N} \alpha_i (s_i^T \mathrm{Av}_i)$$

The optimal solution minimizes J(A), and therefore maximizes

$$\tilde{J} = \sum_{i=1}^{N} \alpha_i (s_i^T \mathrm{Av}_i) = \mathrm{tr}(\mathrm{AB}^T)$$

where tr indicates the trace operator and B matrix is

$$B = \sum_{i=1}^{N} \alpha_i s_i v_i^T$$

A direct solution in terms of direction cosine matrix is computationally expensive, and is most often searched adopting singular value decomposition techniques. If the B matrix is non-singular, which means having at least 3 observed vectors, then

$$A = (B^T)^{-1}(B^T B)^{1/2} = B(B^T B)^{-1/2}$$

However, after the construction of B, this requires the tricky computation of $(B^TB)^{-1/2}$. In fact, the square root of a matrix can have multiple solutions. In a general case, matrix B can be decomposed using singular value decomposition as

$$B = U\Sigma^T V^T = \mathrm{U}\ \mathrm{diag}[\Sigma_{11} \quad \Sigma_{22} \quad \Sigma_{33}]^T V^T$$

where U and V are unitary orthogonal matrices, representing the eigenvectors of matrices BB^T and B^TB respectively, and UU^T=I and VV^T=I. The singular values Σ are the square roots of the eigenvalues of matrix B^TB or of matrix BB^T. Then

$$\Sigma = diag\,[1 \quad 1 \quad (det\ U)(det\ V)]$$
$$A = \mathrm{U}\ \mathrm{diag}\,[1 \quad 1 \quad (\det \mathrm{U})(\det \mathrm{V})]V^T$$

A simpler solution can be sought if we replace A by its representation in terms of quaternions

$$A = \left(q_4^2 - \underline{q}^T\underline{q}\right) I + 2\underline{q}\underline{q}^T - 2\mathrm{q}_4[q\ \wedge]$$

In this case the cost function to maximize becomes

$$\tilde{J} = q^T \mathrm{Kq}$$

where the matrix K, square of order 4, is assembled as

$B = \sum_{i=1}^{N} \alpha_i s_i v_i^T = [b_{\mathrm{ij}}]$	square matrix of order 3
$S = B^T + B$	square matrix of order 3
$z = [b_{23} - b_{32} \ , \ \mathrm{b}_{31} - b_{13} \ , \ \mathrm{b}_{12} - b_{21}]^T$	vector of order 3
$\sigma = \mathrm{tr}(B)$	scalar
$K = \begin{bmatrix} S - \sigma\mathrm{I} & z \\ z^T & \sigma \end{bmatrix}$	square matrix of order 4

To maximize function $\tilde{J}$, the constraint $q^Tq = 1$ must be added with the associated Lagrange multiplier λ, and then the function $\tilde{G} = q^T\mathrm{Kq}\text{-}\lambda(q^T\mathrm{q}\text{-}1)$ is maximized with no constraints. Evaluating the gradient of $\tilde{G}$ and equating the gradient to zero we have

$$\mathrm{Kq} = \lambda\mathrm{q}$$

so that, back substituting into the expression of $\tilde{J}$ we have

$$\tilde{J}(q) = q^T\mathrm{Kq} = q^T\lambda\mathrm{q} = \lambda$$

The maximum value of the cost function is then associated with the maximum eigenvalue of the K matrix. As a consequence the optimal attitude, from a statistical point of view, is the eigenvector associated to the maximum eigenvalue of matrix K.
In a normal operating condition, with all sensors operational, it should be expected that the errors in attitude determination are rather small, which means that we expect the cost function J to be close to zero. In this condition, function $\tilde{J}$ will be close to one, so that we expect that the maximum eigenvalue of K is close to 1. With this assumption, we can avoid the evaluation of eigenvalues and eigenvectors of matrix K, and try to solve an approximate problem that leads to the evaluation of the Gibbs vector

giving the satellite attitude. To do so, rewrite the equation $Kq=\lambda q$ in partitioned form, considering the form of matrix K. We have the following two equations

$$(S-\sigma \mathrm{I})\underline{q}+\mathrm{z}q_4=\lambda_{\mathrm{MAX}}\underline{q}$$
$$z^T\underline{q}+\sigma q_4=\lambda_{\mathrm{MAX}}q_4$$

Dividing the first equation by q_4 and taking the approximation $\lambda_{MAX}=1$, we have

$$(S-\sigma \mathrm{I}-I)\underline{g}+z=0$$

that represents a linear system with unknown vector $\underline{g}$ and coefficient matrix $(S-\sigma I-I)$. Solving this linear system we evaluate the approximate Gibbs vector that provides the satellite attitude.

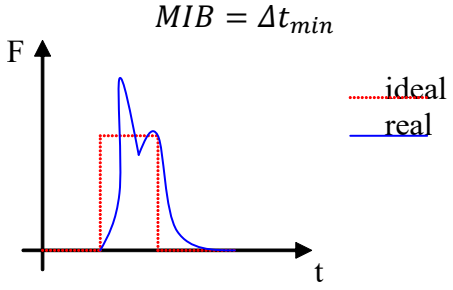
Actuators for spacecraft attitude control

Thrusters for attitude control

The simplest way to create torques is to create a set of forces with direction not aligned with the center of mass, and this can be obtained by mass expulsion techniques. Jet thrusters pose some operational problems due to the ignition transient, and besides it is not simple to finely control the magnitude of the force, so these devices are not used for fine attitude control. In addition, for control quite often the force needed is rather small (milli Newton-meters), while chemical thrusters produce forces in the order of at least some Newton. To make them compatible with attitude control, they are switched on and off with a given modulation, but this enhances the problems due to ignition transients and can cause mechanical ware of the thruster.

These problems can be solved by adopting electric propulsion thrusters, based on electrodynamic acceleration of a suitable ionized propellant, that need de-ionization immediately after expulsion to avoid charging electrically the spacecraft. These thrusters can be easily modulated in amplitude, have a high specific impulse (over 3000) that allows a reduced propellant consumption. The thrust produced can be in the order of a few Newton down to 10^{-6} Newton, so they are well suited to fine control actions. Unfortunately, electric thrusters are extremely power consuming, more or less 90% is devoted simply to keep it ready to use and only 10% is due to the thrust produced, therefore electric propulsion units are often coupled to extremely large solar panels.

With conventional (chemical) thrusters it is not possible to control the amplitude of the thrust; they are either switched on or off. The transient delay and the presence of hydraulic circuits make the actual thrust profile quite different from the ideal one, requiring a careful calibration for proper command selection.
An important characteristic of the thruster for attitude control is the minimum impulse bit (MIB) which is a measure of the accuracy and related to the precision it can point.



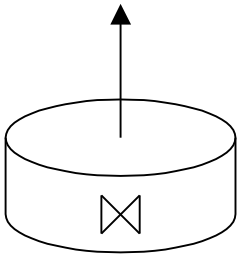
Cold-gas thrusters use a non-reactive gas, stored at high pressure (around 30 MPa), commonly use Nitrogen~SP 70s or Helium~SP 175s. Helium saves mass but is more prone to leakage and more expensive.

Monopropellant uses a propellant that is decomposed catalytically. Hydrazine~SP 242s is by far the most common. They do not require high pressure but are highly toxic.

Electric Thrusters use electric fields to accelerate the expelled ionized mass at high velocity. High SP 500-3000s.

Use of thruster on spinning satellites

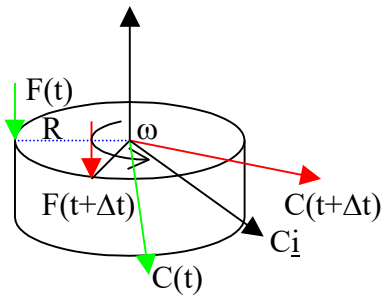
In case of spinning satellites, to control the spin velocity the thrusters must be located on the side of the satellite and the thrust direction must be orthogonal to the angular velocity:



$I\dot{\omega} = M$

Control of the direction of the spin is more critical. It requires thrust in the same direction of the angular velocity, but shifted with respect to the satellite centerline. In this case the ignition delay does not affect the magnitude of the angular velocity (or, the angular momentum), but affects the direction of the velocity. In fact, due to satellite spin, a delay in the application of the control force implies that in inertial space the force is applied with a different line of action, so the torque applied has a different direction in inertial space and hence the different final direction of $\underline{\omega}$ and $\underline{h}$.

$$\Delta\phi = \omega\Delta t$$
$$dM = FR\cos\phi\, d\phi \underline{i}$$



If the force is constant, and applied for a rotation angle equal to 2φ, also the torque will be constant, orthogonal to the average force direction and directed along the centerline of the rotation, characterized by the angleφ:

$$M_{TOT} = 2\int_0^{\phi} FR\cos\phi\, d\phi \underline{i} = 2FR\sin\phi\, \underline{i}$$

In general, to know the direction of the resulting torque the exact thrust profile must be known, in order to evaluate the average direction of the thrust.
If the thruster is switched on for one full spin rotation of the satellite, thus for φ equal to π, M_{TOT} will be zero. The net effect is null even if some propellant has been expelled. The maximum effect is obtained for φ equal to π/2, so that the thruster should be switched on and off every half spin rotation. This is however not efficient, since at the beginning and end of the thrust the effects would tend to be opposite. The efficiency of the process can be evaluated by taking the ratio of the impulse of the torque applied divided by the impulse of the force.
Since:

$$dt = \frac{d\phi}{\omega}$$

and since the maximum torque is equal to FR, the torque impulse is:

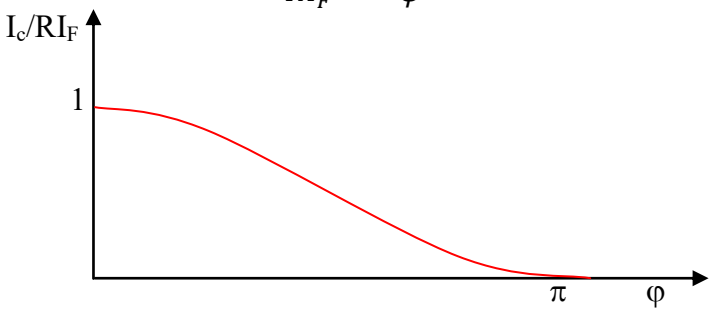
$$I_c = 2\int_0^t FR\cos\phi\, dt = 2\int_0^\phi FR\cos\phi\frac{d\phi}{\omega} = \frac{2FR}{\omega}\sin\phi$$

while the force impulse is:

$$I_F = 2\int_0^t F dt \Rightarrow$$
$$RI_F = 2R\int_0^t F dt = 2R\int_0^\phi F\frac{d\phi}{\omega} = \frac{2FR}{\omega}\phi$$

The ratio of the two gives an indication of the efficiency of the process, as:

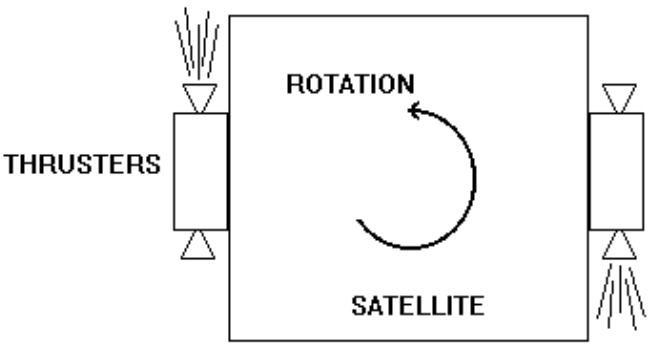
$$\frac{I_c}{RI_F} = \frac{\sin\phi}{\phi}$$



This tells us how much torque we are obtaining along the average direction compared to the total torque applied. If φ tends to zero the efficiency tends to 1. It is obvious that the torque obtained is low, but with a repeating sequence of short pulses, one per rotation, the total torque obtained can be high.

Attitude Thruster configurations

In the simplest configuration, thrusters operate in couples as displayed below:



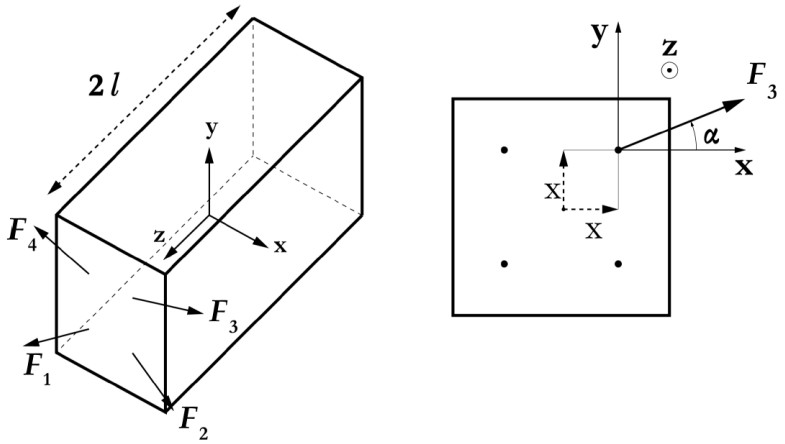
Then the torque and force are then

$$\underline{T} = \underline{r}_1 \wedge \underline{F}_m + (\underline{r}_2 \wedge -\underline{F}_m) = (\underline{r}_1 - \underline{r}_2) \wedge \underline{F}_m$$
$$\underline{F} = \underline{0}$$

In a more general configuration, with an arbitrary number of thrusters, the total torque generated becomes

$$\underline{T}_{tot} = [\hat{R}]\underline{F}$$

where matrix $\hat{R}$ maps the force generated by the set of thrusters into the torque with respect to the body axes. The example below shows the map of the forces generated by 4 thrusters.



$$[\hat{R}] = \begin{bmatrix} l\sin\alpha & 0 & 0 \\ 0 & l\cos\alpha & 0 \\ 0 & 0 & x\sin\alpha - x\cos\alpha \end{bmatrix} \begin{bmatrix} 1 & 1 & -1 & -1 \\ -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 \end{bmatrix}$$

At first glance the problem of solving for $\underline{F}$ seems straightforward since we have seemingly to invert the equation $\underline{u}_{ideal} = [\hat{R}]\underline{F}$. In general, $[\hat{R}]$ is not a square matrix so we can think to apply the Moore-Penrose pseudo inverse. Then we can write $\underline{F} = [\hat{R}]^* \underline{u}_{ideal}$. A general pseudo inverse can be defined that also satisfies this constraint:

$$\underline{F} = [\hat{R}]^* \underline{u}_{ideal} + \gamma w$$

where γ is a scalar defined by

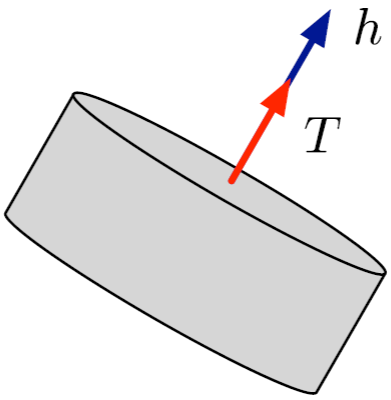
$$\gamma = max_{i=1,..,N}([\hat{R}]^* \underline{u}_{ideal})_i / w_i$$

where $w > 0$ is the null space vector of $\hat{R}$. Therefore, we can apply this method if the null space $w > 0$ of $\hat{R}$ exists. To compute the nullspace *w* we need to solve $\hat{R}w=0$. For the example given above, with 4 thrusters, $w = [1 \quad 1 \quad 1 \quad 1]^T$. Thus, the generalized pseudo inverse can be applied to this configuration.

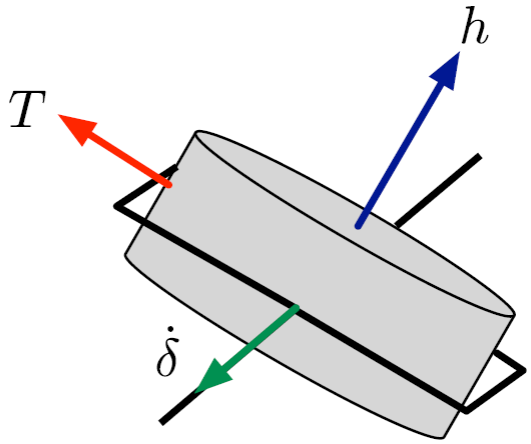
Inertia and reaction wheels

These actuators are based on acceleration and deceleration of spinning rotors. Normally, when the nominal condition is with zero angular velocity the actuator is called Reaction Wheel (RW); those that have a non-zero nominal velocity are called Inertia Wheels (IW). In some cases the distinction between the two categories is not evident, and all actuators can be classified as RW. A further class of actuators is based on wheels that spin at a constant rate but can be tilted around an axis orthogonal to the spin axis, and their principle of operation is that of gyroscopes. These actuators are called

Control Moment Gyroscopes (CMG). All these actuators can be considered as one family of actuators; all can only exchange angular momentum with the satellite but not produce a net torque.



RW and IW torque diagram

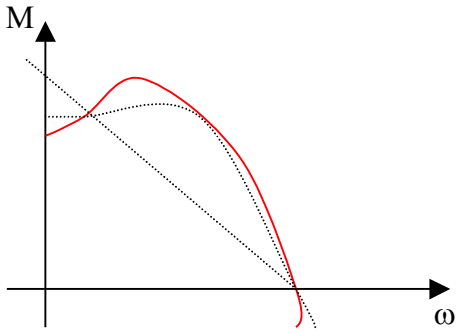


GMG torque diagram

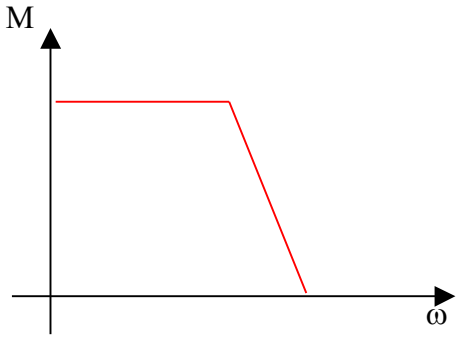
Looking at the equations for a dual spin satellite with no external torques applied, in fact, we can write:

$$I_r\dot{\omega}_r = M_r$$
$$I_z\dot{\omega}_z = -I_r\dot{\omega}_r$$

The torque M_r is in general provided by an electric motor that has its typical operational curve. One possible curve is the following:



Despite the variability of electric motor characteristic, it is commonly accepted that there is a velocity for which the torque falls to zero, this is called saturation speed. When the wheel reaches this speed it can provide no additional torque, either positive or negative, and requires a special desaturation mechanism to make it again operational. An approximation of the characteristic curve can be the following:



In general, actuators are operated in order to keep them within the speed range for which the torque is constant, so that the actuator can be somehow modeled as a linear device if commanded by a current in the armature.
If the torque required by the control is more or less periodic, then appropriate sizing of the actuator can be made in order to have it always within the linear functioning regime. If the torque required is the combination of a periodic component and a secular component, it will be inevitable to reach the saturation speed. Desaturation mechanisms must apply a net torque to the satellite; therefore they can be based on thrusters or on magnetic torquers. In case the secular component of disturbances is persistent, several desaturation maneuvers will be required.
Considering that desaturation maneuvers are studied apart with specific models, the general problem of control with RW, IW or CMG can be modeled by writing the appropriate set of Euler equations, including the effects of the n actuators in the expression of the angular momentum

$$\underline{h} = I\underline{\omega} + A\underline{h}_r$$

Matrix A is composed by 3 rows and as many columns as number of actuators. Each column represents the direction of the axis of rotation of the wheel. Vector $\underline{h}_r$ has as many elements as actuators, and represents the angular momentum of each actuator around its spin axis.
Assuming as usual that the term $I\underline{\omega}$ considers the presence of the rotors with no relative velocity, and that $\underline{T}$ represents the disturbance torque, the dynamics becomes:

$$I\dot{\underline{\omega}} + \underline{\omega} \wedge I\underline{\omega} + \dot{I}\underline{\omega} + \dot{A}\underline{h}_r + A\dot{\underline{h}}_r + \underline{\omega} \wedge A\underline{h}_r = \underline{T}$$

Additional equations for the relative dynamics of the rotors must be added to the system:

$$\dot{A}\underline{h}_r + A\dot{\underline{h}}_r = \underline{M}_r$$

We can now analyze the individual terms of the equation:

$\dot{I}\underline{\omega}$	exists only for CMG, and can be neglected if these are small compared to the satellite
$\dot{A}\underline{h}_r$	exists only for CMG, and indicates the relative attitude rotation in the spacecraft body frame
$A\dot{\underline{h}}_r$	in general it exists for RW and IW, not for CMG since these have often constant spin velocity
$\underline{\omega} \wedge A\underline{h}_r$	exists for any actuator, could be neglected for RW since they are characterized by nominal $\underline{h}_r = 0$

To simplify the solution of the control problem, the four terms can be grouped together in one term. Upon the determination of the control torque required, a separate equation can be solved including the actuator terms to infer the effective control command for the actuator. We can consider one example in which we have only RW.

$$\begin{cases} I\underline{\dot{\omega}} + \underline{\omega} \wedge I\underline{\omega} + \underline{\omega} \wedge A\underline{h}_r + A\underline{\dot{h}}_r = \underline{T} \\ M_c = -\underline{\omega} \wedge A\underline{h}_r - A\underline{\dot{h}}_r \\ I_r\underline{\dot{\omega}}_r = \underline{M}_r \end{cases}$$

We can write the equation for the evaluation of the control law as:

$$I\underline{\dot{\omega}} + \underline{\omega} \wedge I\underline{\omega} = \underline{T} + \underline{M}_c$$

Any linear control design technique can be used to calculate $\underline{M}_c$ after which $\underline{M}_r$ can be evaluated, i.e., $\underline{\dot{h}}_r = I_r\underline{\dot{\omega}}_r$ with the following procedure:

$$A\underline{\dot{h}}_r = -\underline{M}_c - \underline{\omega} \wedge A\underline{h}_r$$
$$\underline{\dot{h}}_r = -A^* \left(\underline{M}_c + \underline{\omega} \wedge A\underline{h}_r\right)$$

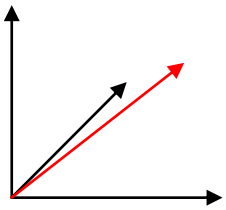
where A^* is the pseudo inverse matrix of A, in the general case, and becomes the inverse only if 3 actuators are used. Notice that the pseudo inverse can be calculated as

$$A_{3,n}x_{n,1} = y_{3,1}$$
$$A_{n,3}^T A_{3,n}x_{n,1} = A_{n,3}^T y_{3,1}$$
$$\left(A_{n,3}^T A_{3,n}\right)_{n,n} x_{n,1} = A_{n,3}^T y_{3,1}$$

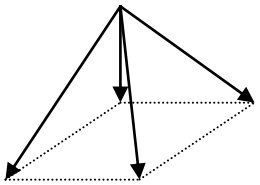
Assuming $\left(A_{n,3}^T A_{3,n}\right)_{n,n}$ is invertible:

$$\left(A_{n,3}^T A_{3,n}\right)_{n,n}^{-1} \left(A_{n,3}^T A_{3,n}\right)_{n,n} x_{n,1} = \left(A_{n,3}^T A_{3,n}\right)_{n,n}^{-1} A_{n,3}^T y_{3,1}$$
$$x_{n,1} = \left(A_{n,3}^T A_{3,n}\right)_{n,n}^{-1} A_{n,3}^T y_{3,1}$$
$$x_{n,1} = A_{n,3}^* y_{3,1}$$

In general 3 rotors are enough to solve exactly the control problem, in practice a fourth rotor is added for redundancy with a geometry that allows a full control capacity even with one rotor failed. This means that the 4 rotors have distinct rotation axes, and any combination of 3 rotors would allow the invertibility of the rotor A matrix.
The two most popular configurations are with 3 rotors aligned with the principal axes and the fourth with equal components along the three axes, or a full pyramid configuration with no actuator aligned with any of the principal axes. In this second configuration no direction is privileged, and the pyramid parameters can be tuned to have different control capacity along the three axes, this can be useful if the disturbances are not uniform along the principal axes.



Three axes and diagonal



pyramid

Analyze now the configuration with 3 rotors aligned with the principal axes and the fourth with equal components along the three axes:

$$A = \begin{bmatrix} 1 & 0 & 0 & 1/\sqrt{3} \\ 0 & 1 & 0 & 1/\sqrt{3} \\ 0 & 0 & 1 & 1/\sqrt{3} \end{bmatrix}$$

We can extract any of the actuators, and the remaining three will guarantee the invertibility of A. The pseudo-inverse is:

$$A^* = \begin{bmatrix} 5/6 & -1/6 & -1/6 \\ -1/6 & 5/6 & -1/6 \\ -1/6 & -1/6 & 5/6 \\ 1/2\sqrt{3} & 1/2\sqrt{3} & 1/2\sqrt{3} \end{bmatrix}$$

And the system is solved as:

$$\dot{\underline{h}}_r = \underline{M}_r = -A^* \left(\underline{M}_c + \underline{\omega} \wedge A\underline{h}_r \right) = A^* \underline{M}_e$$

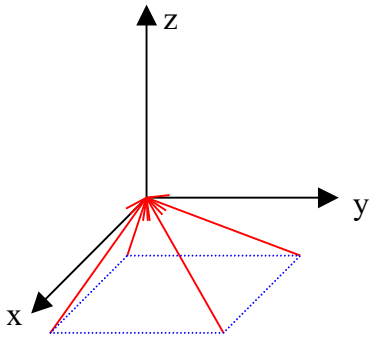
Assume now that the control torque must be aligned with the x-axis, we see that multiplying A^* with the required torque $\{1,0,0\}^T$ we have a control command for each actuator in the system, not simply the actuator aligned with x. In this case, the norm of the wheel control vector $\underline{M}_r$ is

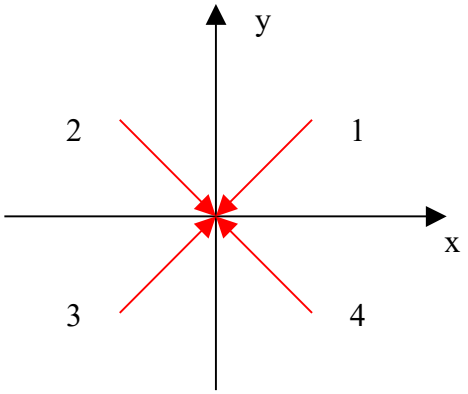
$$\underline{M}_c = \begin{Bmatrix} 1 \\ 0 \\ 0 \end{Bmatrix}$$

$$\underline{M}_r = \begin{Bmatrix} 5/6 \\ -1/6 \\ -1/6 \\ 1/2\sqrt{3} \end{Bmatrix}$$

$$\left\| \underline{M}_r \right\| = \sqrt{\frac{27}{36} + \frac{1}{12}} = \sqrt{\frac{30}{36}}$$

Using the pyramid configuration, with all actuators with equal components along the principal axes we have instead:





$$A = \begin{bmatrix} -a & a & a & -a \\ -a & -a & a & a \\ a & a & a & a \end{bmatrix}$$

$$a = \frac{1}{\sqrt{3}}$$

The pseudo-inverse is:

$$A^* = \begin{bmatrix} -b & -b & b \\ b & -b & b \\ b & b & b \\ -b & b & b \end{bmatrix}$$

$$b = \frac{\sqrt{3}}{4}$$

Assume now that the control torque must be aligned with the x-axis, we see that multiplying A^* with the required torque $\{1,0,0\}^T$ we have the same control command on all actuators, with different signs. The norm of the wheel control vector $\underline{M}_r$ is:

$$\|\underline{M}_r\| = \sqrt{\frac{3}{4}}$$

The norm is lower than the previous case, so this configuration is preferred. In case one rotor fails, as example rotor number 4, we still have control capability since:

$$A_{1,2,3}^{-1} = \begin{bmatrix} -\frac{\sqrt{3}}{2} & 0 & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} & 0 \\ 0 & \frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}$$

Any torque component is provided by activating two rotors with equal acceleration.
As a general rule, adoption of IW can be preferred to RW since there are minor friction problems. In addition, a suitable configuration of IW can provide a non-zero nominal angular momentum, useful for passive attitude stability. In this case, it must be remembered that the overall rotor angular momentum must be aligned with the satellite angular momentum, so for some geometries of the actuator set the nominal angular velocities of the actuators might not be the same.

Considering the pyramid configuration seen above, if all rotors have the same velocity then the only nonzero component of $\underline{h}$ is along axis z, and we can control the satellite exactly as if we had one IW along axis z and 2 RW along axes x and y. Accelerating rotors 3 and 4 and slowing rotors 1 and 2 the net effect is equivalent to a RW along axis y, slowing rotors 2 and 3 and accelerating rotors 1 and 4 the net effect is equivalent to a RW along axis x. Accelerating all rotors together the effect is equivalent to IW along axis z.
Notice that, due to dynamic coupling, it is not always required to have one actuator for each axis, since 2 actuators can guarantee controllability of the system and a third actuator can be used for redundancy.

Note that RW will have a maximum momentum storage or in other words a saturation limit on their angular velocity. When the RW reaches the saturation limit no additional exchange of momentum may occur in the same direction and therefore the RW must be de-saturated. For example, for a RW with dynamics $\dot{\underline{h}}_{rw} = \underline{T}_{rw}$ where $\underline{T}_{rw}$ is the torque applied to the RW we can use a simple feedback control to de-celerate the wheel $\underline{T}_{rw} = -\underline{T}sign\left(\underline{h}_r\right)$. However, de-saturating the wheel will result in a generally undesired momentum exchange with the spacecraft e.g a de-tumbled spacecraft will begin to move when the RW is de-saturated. Thus, during the de-saturation of the RW a torque must be applied to the spacecraft using another actuator such as a magnetic torquer or thruster to maintain the spacecraft orientation, for example $\underline{T}_{sc} = -\underline{T}sign\left(\underline{h}_r\right)$. When the RWs velocities fall within a prescribed tolerance the applied torque to them can be switched off while $\underline{T}_{sc} = -\underline{T}sign\left(\underline{h}_r\right)$is applied to de-tumble the spacecraft due to angular velocities induced by the mismatch of the applied reaction wheel torques and the other actuator used.

Control moment gyroscopes

Control moment gyroscopes (CMG) generate a torque by gyroscopic effect. A rotating disc is tilted around one axis orthogonal to the spin axis of the disc, obtaining a torque long the third axis of an orthogonal frame. The angular momentum is then exchanged with the satellite not by a variation in the spin of the actuator, but by changing its relative orientation within the satellite. This effect allows on one side to obtain torques in general higher than those obtained with RW, but on the other side the evolution of the actuator configuration can create some singularity problem for specific configurations.
Following the same steps as shown for the RW, we can formulate the general control problem with CMG actuators as

$$I\dot{\underline{\omega}} + \underline{\omega} \wedge I\underline{\omega} + \dot{A}\underline{h}_r + \underline{\omega} \wedge A\underline{h}_r = \underline{T}$$

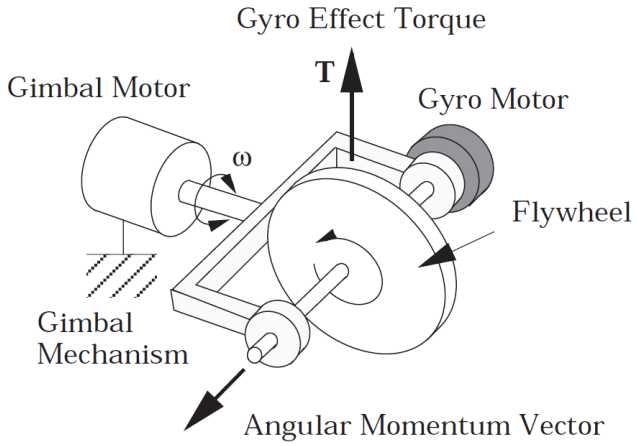
$$\begin{cases} I\dot{\underline{\omega}} + \underline{\omega} \wedge I\underline{\omega} + \dot{A}\underline{h}_r + \underline{\omega} \wedge A\underline{h}_r = \underline{T} \\ M_c = -\underline{\omega} \wedge A\underline{h}_r - \dot{A}\underline{h}_r \end{cases}$$

$$I\dot{\underline{\omega}} + \underline{\omega} \wedge I\underline{\omega} = \underline{T} + \underline{M}_c$$

$$\dot{A}\underline{h}_r = -\underline{M}_c - \underline{\omega} \wedge A\underline{h}_r$$

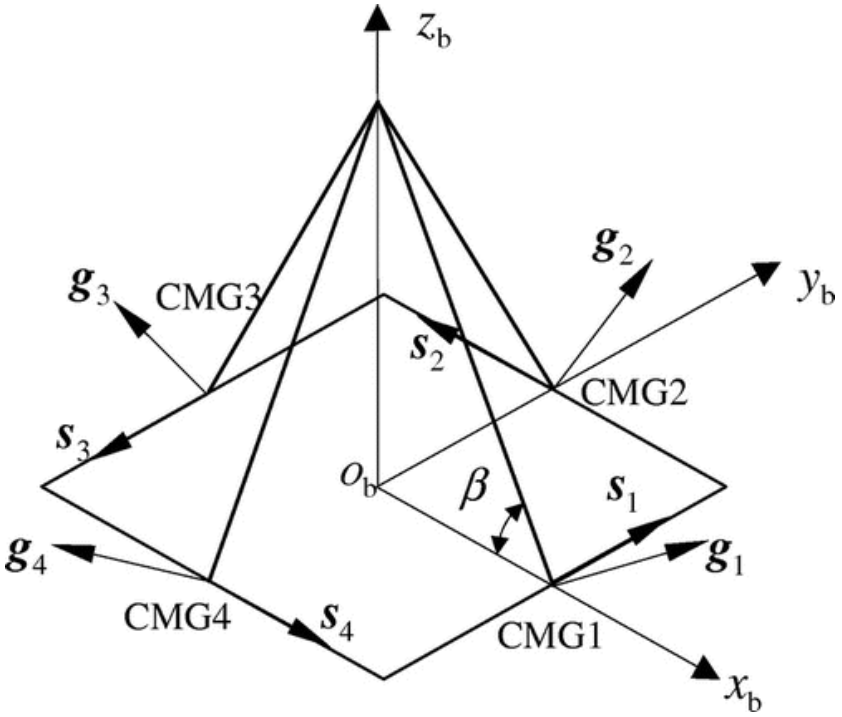
The problem is solved when the matrix $\dot{A}$ is evaluated. With the aid of one example we will show that the evaluation of $\dot{A}$ can be reformulated as the evaluation of a vector of tilt angle rates, which is easier to solve.

Assuming to have only 1 single gimbal GMG, the command becomes $\dot{\delta} = \omega$



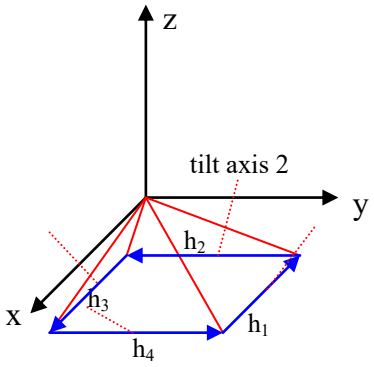
$$h_{\text{CMG}} = A\underline{h}_r = h_r \begin{bmatrix} -\sin\delta \\ \cos\delta \\ 0 \end{bmatrix} \quad \dot{h}_{\text{CMG}} = h_r \begin{bmatrix} -\cos\delta \\ -\sin\delta \\ 0 \end{bmatrix} \dot{\delta}$$

If the gimbal motor is also inclined by a constant angle as in the figure below



$$h_{\text{CMG}} = A\underline{h}_r = h_r \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix} \begin{bmatrix} -\sin\delta \\ \cos\delta \\ 0 \end{bmatrix} = h_r \begin{bmatrix} -\cos\beta\sin\delta \\ \cos\delta \\ \sin\beta\sin\delta \end{bmatrix}$$

Assume a configuration with 4 CMG, whose nominal orientation makes the nominal angular momentum zero and whose tilt directions are along the 4 sides of a pyramid, as in figure



Calling δ the tilt angle of the 4 CMG and β the inclination of the pyramid, assuming all CMG have the same spin speed and therefore the same angular momentum h_r we can write

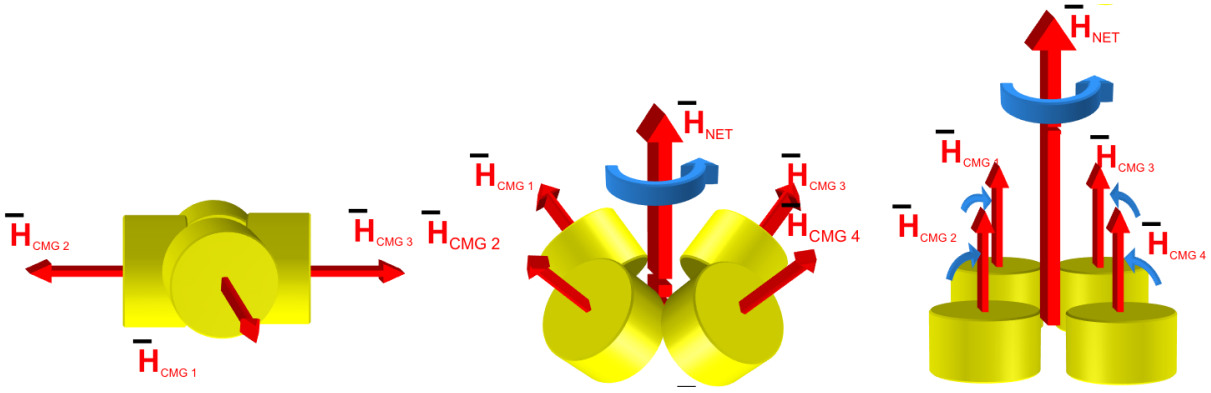
$$h_{\text{CMG}} = A\underline{h}_r = \begin{bmatrix} sin\,\beta\,sin\,\delta_1 & sin\,\beta\,sin\,\delta_2 & sin\,\beta\,sin\,\delta_3 & sin\,\beta\,sin\,\delta_4 \\ -cos\,\delta_1 & cos\,\beta\,sin\,\delta_2 & cos\,\delta_3 & -cos\,\beta\,sin\,\delta_4 \\ -cos\,\beta\,sin\,\delta_1 & -cos\,\delta_2 & cos\,\beta\,sin\,\delta_3 & cos\,\delta_4 \end{bmatrix} \begin{Bmatrix} h_r \\ h_r \\ h_r \\ h_r \end{Bmatrix}$$

$$\dot{h}_{\text{CMG}} = \dot{A}\underline{h}_r = \text{B}\dot{\delta} = h_r \begin{bmatrix} sin\,\beta\,cos\,\delta_1 & sin\,\beta\,cos\,\delta_2 & sin\,\beta\,cos\,\delta_3 & sin\,\beta\,cos\,\delta_4 \\ sin\,\delta_1 & cos\,\beta\,cos\,\delta_2 & -sin\,\delta_3 & -cos\,\beta\,cos\,\delta_4 \\ -cos\,\beta\,cos\,\delta_1 & sin\,\delta_2 & cos\,\beta\,cos\,\delta_3 & sin\,\delta_4 \end{bmatrix} \begin{Bmatrix} \dot{\delta}_1 \\ \dot{\delta}_2 \\ \dot{\delta}_3 \\ \dot{\delta}_4 \end{Bmatrix}$$

The control equation can then be reformulated as

$$\dot{A}\underline{h}_r = -\underline{M}_c - \underline{\omega} \wedge A\underline{h}_r = B\underline{\dot{\delta}} \quad \rightarrow \quad \underline{\dot{\delta}} = -B^* \left[\underline{M}_c + \underline{\omega} \wedge A\underline{h}_r\right]$$

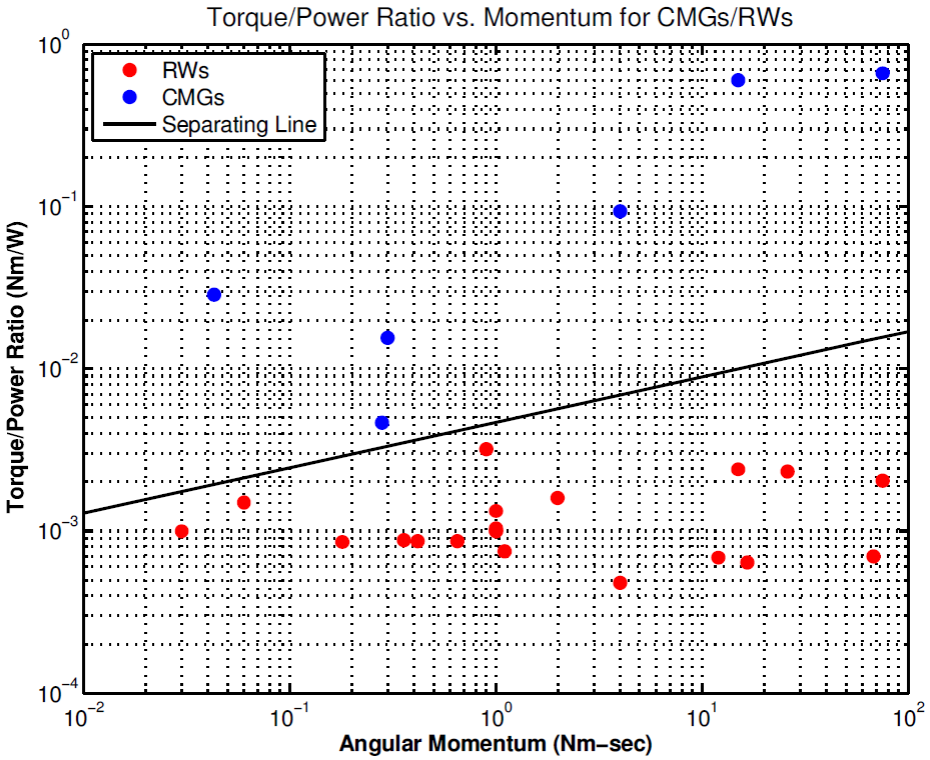
Note that CMGs are subject to saturation, as RWs, if $\underline{\dot{\delta}}$ exceeds the maximum capacity of the drive motor. In addition, CMGs can assume configurations that become singular, which means that the overall control torque can not be arbitrary. This singularity arises from tha fact that the configuration of the set of CMGs changes with time and could bring some of the CMGs to have the same direction of the applicable torque, as shown in the following set of figures.

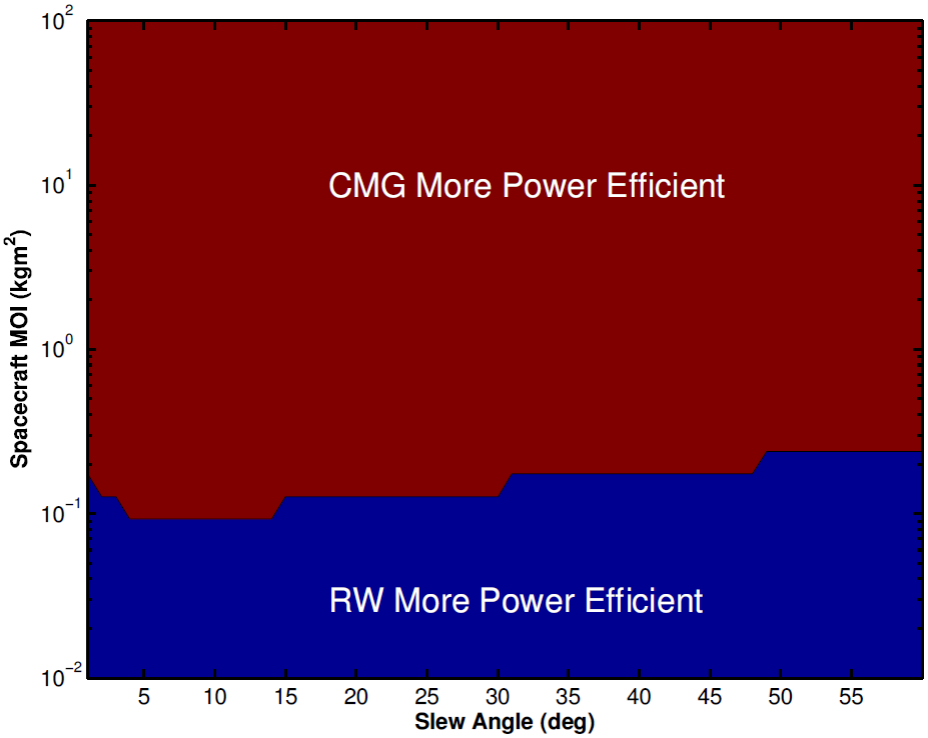


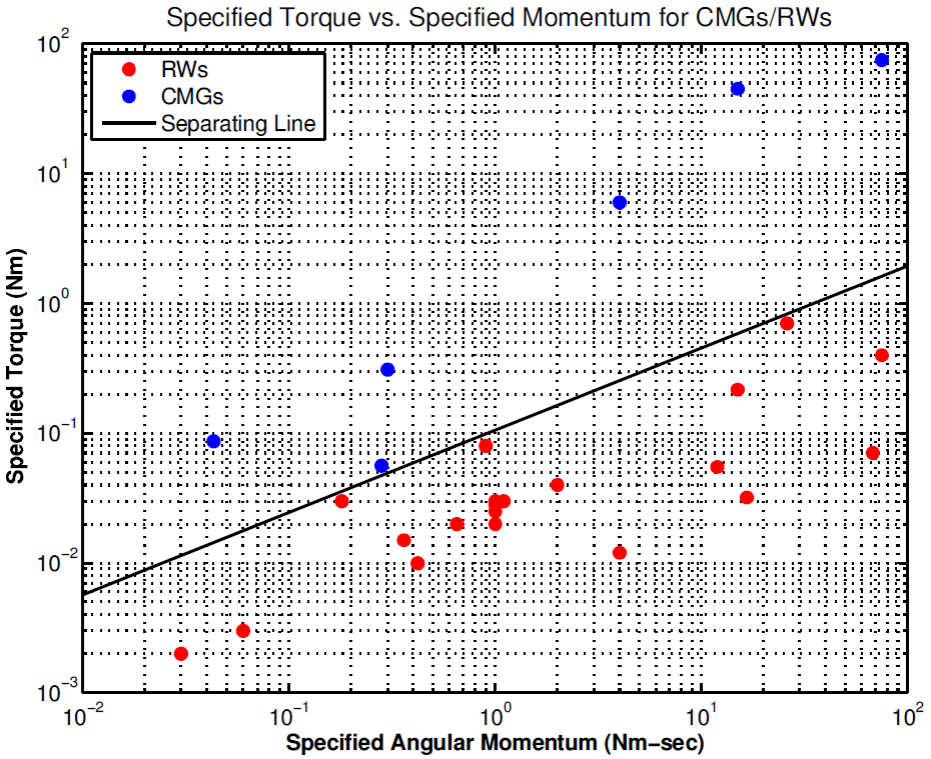
Singular configurations can be avoided by adding a command that generates rotation of the gimbal axes but overall a null torque. This is called null motion command $\underline{\dot{\delta}}_n$, it is the null space of matrix B and can be computed as

$$B\dot{\underline{\delta}}_n = 0 \quad ; \quad \dot{\underline{\delta}} = -B^*\left[\underline{M}_c + \underline{\omega} \wedge A\underline{h}_r\right] + k\dot{\underline{\delta}}_n$$

As a general rule, CMGs have a higher torque capacity compared to RWs and are more efficient for larger spacecraft.






Specified Torque vs. Specified Momentum for CMGs/RWs
RWs
CMGs
Separating Line
Specified Torque (Nm)
Specified Angular Momentum (Nm–sec)
10^{2}
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-2}
10^{-1}
10^{0}
10^{1}
10^{2}

Magnetic actuators

Magnetic actuators generate a torque by inducing a magnetic dipole in a coil that is surrounded by the Earth's magnetic field. The magnetic dipole generated can be modeled as:

$$\underline{D} = \mu n S \underline{I}$$

$\mu =$ magnetic permeability
$n =$ number of windings
$S =$ area of coil
$\underline{I} =$ current intensity

Torque is then given by the vector product of this dipole moment and the external magnetic field:

$$\underline{M} = \underline{D} \wedge \underline{B}$$

The effectiveness of these actuators varies with the orbit height, due to the change in the external magnetic field, and can typically provide torques in the range $10^{-3} \div 10^{-6}$ Nm. They can be customized by changing the number of windings and/or the area of the coils.

The torque generated is perpendicular to the magnetic dipole generated D and to the external magnetic field B, so that it is never possible to generate three independent components of the control torque. However, if the satellite dynamics is coupled on two axes, typically x and y, the availability of control along z and along a second axis guarantees complete controllability.

$$\underline{M} = \underline{D} \wedge \underline{B} = -\underline{B} \wedge \underline{D}$$

Expanding matrix notation we have:

$$\begin{Bmatrix} M_x \\ M_y \\ M_z \end{Bmatrix} = \begin{bmatrix} 0 & B_z & -B_y \\ -B_z & 0 & B_x \\ B_y & -B_x & 0 \end{bmatrix} \begin{Bmatrix} D_x \\ D_y \\ D_z \end{Bmatrix}$$

The control command is D so that we should write:

$$\underline{D} = [-B \wedge]^{-1} \underline{M}$$

Unfortunately $[-B \wedge]$ is a singular matrix, so D cannot be evaluated through this procedure. This reflects the fact that it is not possible to provide three independent components of the control torque. A solution can be found if the dynamics is written in a reference system in which B is aligned with axis z, so that it can be made explicit that the control torque is in the x-y plane of this new reference. This method is quite tedious since it requires a continuous change of reference, and a further transformation of the commands from the rotated reference to the principal axes, in which the actuators are physically located. We can instead reformulate the problem in order to require only 2 components of control, along z and y for example, and write:

$$\begin{Bmatrix} 0 \\ M_y \\ M_z \end{Bmatrix} = \begin{bmatrix} 0 & B_z & -B_y \\ -B_z & 0 & B_x \\ B_y & -B_x & 0 \end{bmatrix} \begin{Bmatrix} D_x \\ D_y \\ D_z \end{Bmatrix}$$

Assuming furthermore that $D_x = 0$ we get to:

$$M_y = B_x D_z \quad \Rightarrow \quad D_z = \frac{M_y}{B_x}$$
$$M_z = -B_x D_y \quad \Rightarrow \quad D_y = -\frac{M_z}{B_x}$$

Now, we notice that the first equation produces still a residual torque M_x:

$$M_{xr} = -M_z \frac{B_z}{B_x} - M_y \frac{B_y}{B_x}$$

This residual is not required by the control, it is due to the interaction between $\underline{B}$ and $\underline{D}$ necessary to produce M_y and M_z. In general, this control structure is used driving to zero the component M_z and using the pitch component of the magnetic field. It is then necessary to add one additional actuator to provide control along the pitch axis z, decoupled from the x and y:

$$D_y = -\frac{M_x}{B_z}$$
$$D_x = -\frac{M_y}{B_z}$$

Passive stability along z can be provided by gravity gradient or by a RW or IW, in such a case the control can be modeled as:

$$\begin{Bmatrix} M_x \\ M_y \\ M_z \end{Bmatrix} = \begin{bmatrix} 0 & B_z & -B_y \\ -B_z & 0 & B_x \\ B_y & -B_x & 0 \end{bmatrix} \begin{Bmatrix} D_x \\ D_y \\ 0 \end{Bmatrix} + \begin{Bmatrix} 0 \\ 0 \\ -\dot{h}_z \end{Bmatrix}$$

or:

$$\begin{Bmatrix} M_x \\ M_y \\ M_z \end{Bmatrix} = \begin{bmatrix} 0 & B_z & 0 \\ -B_z & 0 & 0 \\ B_y & -B_x & 1 \end{bmatrix} \begin{Bmatrix} D_x \\ D_y \\ -\dot{h}_z \end{Bmatrix}$$

The control allocation matrix is now nonsingular, so that the actuator commands can be calculated in order to have three independent components of control torque:

$$\begin{Bmatrix} D_x \\ D_y \\ -\dot{h}_z \end{Bmatrix} = \frac{1}{B_z} \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ B_x & B_y & B_z \end{bmatrix} \begin{Bmatrix} M_x \\ M_y \\ M_z \end{Bmatrix}$$

If the magnetic field component B_z is nonzero the problem can be solved.

A different alternative to evaluate $\underline{D}$ leads to an approximation of the control problem. We can write:

$$\underline{M} = \underline{D} \wedge \underline{B}$$

$$\underline{B} \wedge \underline{M} = \underline{B} \wedge (\underline{D} \wedge \underline{B})$$
$$\underline{B} \wedge \underline{M} = (\underline{B} \cdot \underline{B})\underline{D} - (\underline{D} \cdot \underline{B})\underline{B}$$

Now, we can assume that D is generated in such a way to be always perpendicular to B so that:

$$\underline{D} \cdot \underline{B} = 0$$

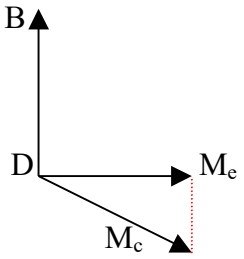
Therefore we can evaluate D:

$$\underline{B} \wedge \underline{M} = B^2 \underline{D} \quad \rightarrow \quad \underline{D} = \frac{\underline{B} \wedge \underline{M_c}}{B^2}$$

However we also have:

$$\underline{M_{\text{eff}}} = \frac{1}{B^2}\left(\underline{B} \wedge \underline{M_c}\right) \wedge \underline{B}$$

and this effective control is equal to the desired control only if the desired control $\underline{M_c}$ is really orthogonal to B, which also means that if D is orthogonal to B we can only calculate the component of control torque M that is orthogonal to B and to D:



This because, having used the vector product to find the solution, this is limited to the plane perpendicular to B, so the method is effective only when the control torques required are by their nature almost perpendicular to the magnetic field.

The above control law can be used to relate an ideal control law to the magnetic dipole. In practice, due to the approximate nature of the control law and the fact that magnetic torquers are an under-actuated controller, the control has only been used for de-saturation and coarse Sun-pointing. For de-saturation of momentum exchange devices, for example, when the RW reaches its saturation limit no additional exchange of momentum may occur in the same direction and therefore the RW must be de-saturated. For example, for a RW with dynamics $\dot{\underline{h}}_{rw} = \underline{T}_{rw}$ where $\underline{T}_{rw}$ is the torque applied to the RW can use a simple feedback control to decelerate the wheel $\underline{T}_{rw} = -k\, sign(\underline{h}_{rw})$. However, de-saturating the wheel will result in a generally undesired momentum exchange with the spacecraft e.g a de-tumbled spacecraft will begin to move when the RW is de-saturated. Thus, during the de-saturation of the RW a magnetic torque can be applied with the control $\underline{M_c} = -M_{max}\, sign(\underline{h}_{rw})$ where $M_{max}$is the maximum possible torque that can be produced by a RW. The torque induced on the spacecraft by decelerating the reaction wheel should be as equal in magnitude as possible. The tuning parameter $k$should be tuned to approximatley $k \approx M_{max}$. This type of control law will only be useful when the angular velocities of the spacecraft are small, since otherwise there is a gyroscopic couling term between the angular velocities of the spacecraft and the wheels that could be non-negligible. In summary we have the control law for a de-saturation maneuver:

Torque applied to the reaction wheel:

$$\underline{T}_{rw} = -M_{max}\, sign(\underline{h}_{rw})$$

And the magnetic dipole feedback:

$$\underline{D} = \frac{M_{max}\, sign(\underline{h}_{rw})}{\underline{b}^2}$$