

4 Sensors and actuators for space applications

4.1 Objectives

TBC

4.2 Sensors

There are several criteria for classifying sensors for attitude (and orbit) determination. Here we shall distinguish between

- 1) inertial sensors like accelerometers and gyroscopes, capable of sensing movement variations without any communication link or field-of-view,
- 2) attitude sensors, like the classical Sun and Earth sensors (requiring a field-of-view) and magnetometers (no field-of-view), and the modern sensors like star trackers (field-of-view) and differential GPS (Global Positioning System) or equivalent (radio link),
- 3) position sensors like GPS and radio links.

Table 1 provides performance and constraints of attitude and angular rate sensors.

Table 1. Attitude and angular rate sensors					
No	Sensor	Accuracy [mrad]	Weight [kg]	Power [W]	Limitations and typical FOV [rad×rad]
0	Sun sensor	0.17 ÷ 175	0.2÷2	0÷3	Sensible to planet albedo. FOV 1×1
1	Earth horizon sensor	0.88 ÷ 17.5	2÷5	0.3÷10	Accuracy limit due to horizon uncertainty. Applicable to Low-Earth-Orbit satellites.
2	Star tracker	0.005 ÷ 0.29	3÷7	5÷20	(i) Angular rotation, (ii) Sun, Earth and Moon stray light. Line of-sight separated from stray-light sources of about 0.5 to 0.7 rad. FOV 0.25×0.25
3	Magnetometer	8.8 ÷ 52	0.6÷1.2	<1	Low-Earth-Orbit satellites; accuracy limit due to Earth

Table 1. Attitude and angular rate sensors					
No	Sensor	Accuracy [mrad]	Weight [kg]	Power [W]	Limitations and typical FOV [rad×rad]
					magnetic field uncertainty
4	Differential GPS	8.8 ÷ 175	3	3 ÷ 10	Accuracy limit by multipath beam and carrier tracking errors
5	Gyroscope	0.005 μ rad/s			Subject to drift

4.3 Attitude sensors

4.3.1 Sun sensors

4.3.1.1 Introduction

Sun sensors are mounted on every spacecraft, because of its luminosity and its small apparent angular radius ($\rho_s = 4.67$ mrad) as seen by spacecrafts near the Earth (at 1 Astronomic Unit, AU) compared to the Earth globe. At AU distance sun can be kept as a point-wise source at an infinite distance, whose size can be neglected for the most part of sun sensors. Sun sensors are usually employed during the Sun acquisition phase, which is essential for battery recharging and power supply, and for the initial attitude determination when a bright and unambiguous light source in the celestial sphere is essential. They are light weight, inexpensive and low power as Table 1 points out.

Solar sensors are rather simple in construction and conception. they can be classified in three types:

- 1) analogue sensors, providing as output an analogue signal as a function of the sun aspect angle α between the sensor line-of-sight \vec{k}_s and the sensor-to-sun direction \vec{s} ,
- 2) detectors, providing as output a binary signal of the presence/absence of the sun beam in the sensor Field-of-view (FOV),
- 3) digital sensors, providing as output a numerical signal as a function of the sun aspect angle α .

The sensor output is a unit vector \vec{s} pointing in the Sun direction, with coordinates \mathbf{s}_s in the instrument frame $\mathcal{R}_s = \{C_s, \vec{s}_1, \vec{s}_2, \vec{s}_3\}$ where \vec{s}_3 is the optical axis (or line-of-sight, LOS). The inertial coordinates \mathbf{s}_i of \vec{s} are known from the satellite position \vec{r} and the Earth position in

its orbit. Denoting the known (calibrated) instrument orientation matrix in the body frame with R_s^b , the following relations are established with the Sun direction \mathbf{s}_b in body coordinates

$$\mathbf{s}_b = R_s^b \mathbf{s}_s = R_i^b(\mathbf{q}) \mathbf{s}_i, \quad (4.1)$$

where \mathbf{q} is the body quaternion.

4.3.1.2 Analogue sensors

The basic principle comes from a single photocell ('cosine sensor') whose output current is proportional to the cosine of the Sun aspect (or incidence) angle α between the Sun direction \vec{s} and the photocell normal $\vec{n}_s = \vec{s}_3$, and to the solar energy flux (or irradiance) Φ_s [W/m²]. The solar power delivered to a photocell area A holds

$$W_s = A \vec{P}_s \cdot \vec{n}_s = A \Phi_s \vec{s} \cdot \vec{n}_s = A \Phi_s \cos \alpha, \quad |\alpha| < \pi/2, \quad (4.2)$$

where \vec{P}_s is the Poynting vector of the solar electromagnetic radiation. Supplying the cell with constant voltage and assuming constant solar flux, the cell current I becomes proportional to $\cos \alpha$ as follows

$$I(\alpha) = I(0) \cos \alpha, \quad (4.3)$$

where $I(0) \approx 0.1$ mA is the peak current depending on A , Φ_s and other sensor parameters. The sensor field-of-view¹ (FOV) $2\alpha_{\max}$ is lower than π due to a dead zone $\delta \leq 0.2$ rad of the photocell

$$2\alpha_{\max} = \pi - 2\delta. \quad (4.4)$$

When the incident radiation is normal to the cell surface, the circulating current reaches the maximum value, but the current variation about this condition, $\alpha \cong 0$, results to be very low due to cosine law in (4.3) thus reducing sensor gain (sensibility) and relevant accuracy as it holds

$$\Delta I(\alpha) = -I(0) \alpha \sin \alpha \cong -I(0) \alpha^2. \quad (4.5)$$

Such a configuration can only be used to detect the cone (axis and aperture) where the Sun is visible from the sensor. In practice the sensor can be employed as a Sun detector (they are usually referred to as Coarse Sun Sensors, CSS) and not as an accurate direction sensor. In addition the signal (4.3) is affected by sign ambiguity.

To dispose of an accurate measurement of Sun direction especially about the sensor normal and to eliminate the sign ambiguity, multiple cells are adopted (Fine Sun Sensors, FSS). Here

¹ The FOV is the defined as twice the maximum Sun incidence angle α that the sensor can measure.

we limit to a pair $k = 0, 1$ of symmetrically inclined photocells ('single axis sensor') with respect to the sensor normal \vec{n}_s as in Figure 1. Denoting tilt of the k photocell normal \vec{n}_{sk} wrt to \vec{n}_s with $s_k \alpha_0$, where $s_0 = 1$, $s_1 = -1$, the photocell currents become

$$I_k(\alpha) = I_k(0) \cos(\alpha_0 + s_k \alpha). \quad (4.6)$$

Taking the difference and assuming equal zero-angle current, the differential current

$$\Delta I = I_1 - I_0 = I(0) (\cos(\alpha_0 - \alpha) - \cos(\alpha_0 + \alpha)) = 2I(0) \sin \alpha_0 \sin \alpha, \quad (4.7)$$

becomes proportional to the signed Sun angle. As a limitation the FOV reduces to

$$2\alpha_{\max} = \pi - 2(\alpha_0 + \delta). \quad (4.8)$$

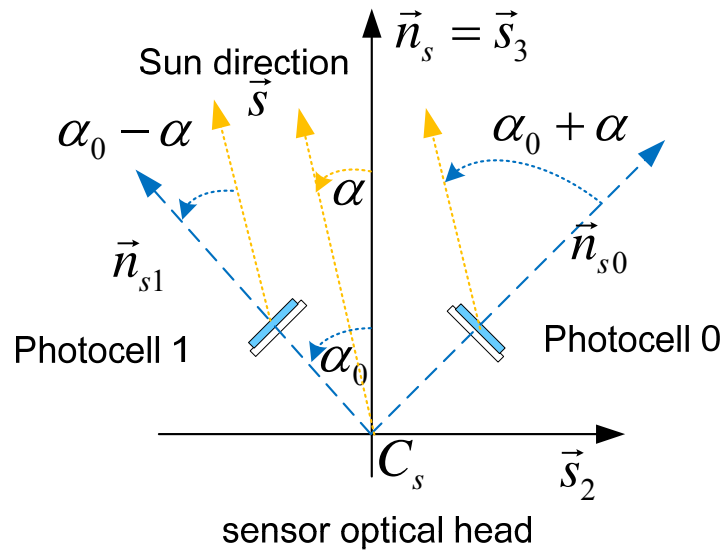


Figure 1. 2D layout of single-axis sun sensor.

Outside the FOV the sensor behaves as a single photocell as in (4.3), thus affected by sign ambiguity. Figure 2 shows the voltage response $V(\alpha) = RI(\alpha)$ versus incidence angle of a cosine sensor as in (4.3) and the single-axis response accompanied by the positive and negative components. Response parameters are

$$\alpha_0 = \pi/4, I(0) = 0.1 \text{ mA}, R = 50 \text{ k}\Omega, \delta \cong 0.1 \text{ rad}. \quad (4.9)$$

Note the single-axis response is rather linear within the FOV $2\alpha_{\max} \cong 1.4 \text{ rad}$. The largest fractional deviation $\partial\alpha$ from linearity (linearity error) occurs at $\alpha = \pm\alpha_{\max}$ and holds

$$\partial\alpha = \frac{\alpha_{\max} - \sin(\alpha_{\max})}{\sin(\alpha_{\max})} \cong 0.087. \quad (4.10)$$

Narrowing the FOV it can be further reduced, but linearity error can be compensated by employing the nonlinear response (4.7).

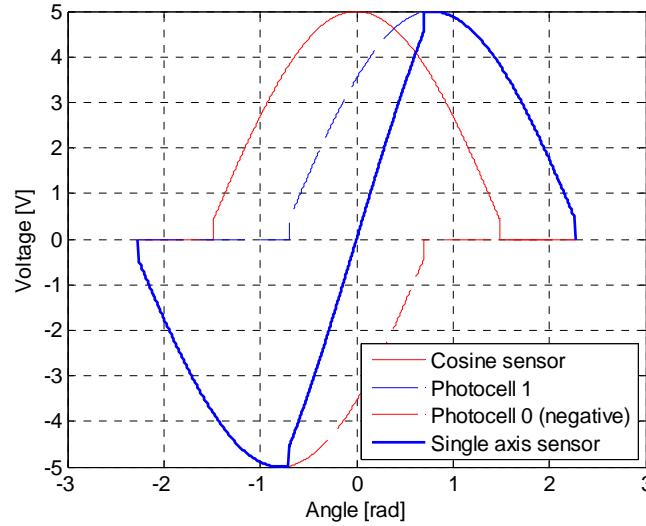


Figure 2. Cosine sensor (centre top) and single-axis sensor (solid line) with photocell signals (dashed lines).

To retrieve Sun direction, two incidence angles α and β must be measured, which is achieved by mounting a pair of single-axis sensors 90° apart (one along \vec{s}_1 and another along \vec{s}_2) with the same optical axis \vec{n}_s . Then assuming a unit component along the optical axis the sun unit vector in the instrument frame holds with the help of

$$\mathbf{s}_i = \frac{1}{\sqrt{1 + \tan^2 \alpha + \tan^2 \beta}} \begin{bmatrix} \tan \alpha \\ \tan \beta \\ 1 \end{bmatrix}. \quad (4.11)$$

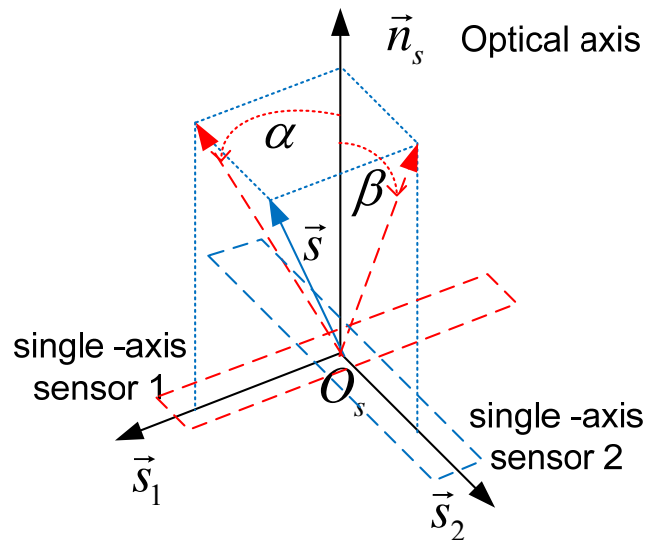


Figure 3. Geometry of two-axes (analogue) Sun sensor.

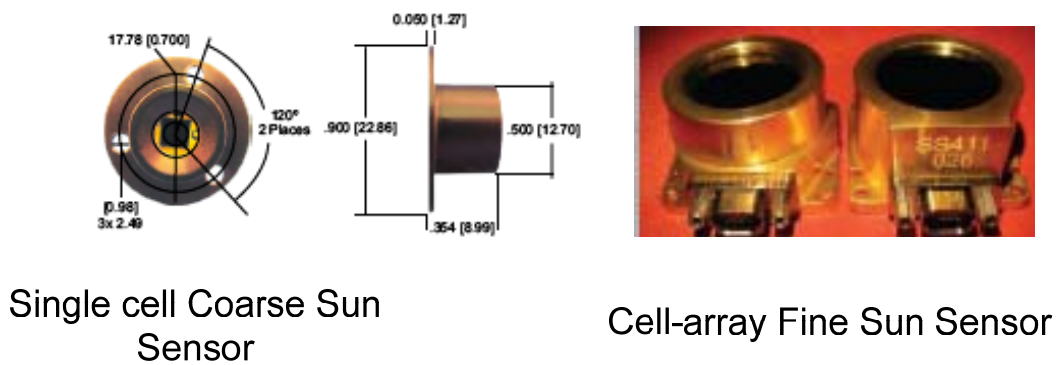


Figure 4. Coarse and fine sun sensors.

TBC with parameter table.

4.3.1.3 Digital sensor

Modern and expensive Sun sensors are the Digital Sun Sensors (DSS) that are capable of an accuracy well below 1 mrad, much better than analogue sensors. In addition they allow a larger FOV. A single axis sensor is made by a Sun entrance slit collecting light rays through an aperture $|\alpha| \leq \alpha_{\max} < \pi/2$. Below the slit there is the optical head focusing the rays on an array of photocells arranged so as to generate when lit the bits of a digital code (usually a Gray code¹). Each bit of the code is generated by a cell located on a row along axis normal to

¹ A Grey code is an equidistant code meaning that only one bit changes between two successive codes.

the projection of the entrance slit The first row is the automatic threshold adjust (ATA) slit, occupying the whole FOV, and providing the cosine sensor response (4.3), to be used as a threshold for selecting the right Gray code slits. The second row, occupying half of the FOV, is the sign cell, which is the most significant bit (MSB). Accuracy is dictated by the code size expressed in number of bits μ_s . Assuming $\alpha_{\max} = 1$ rad and $\mu_s = 14$ the sensor quantization ρ_s holds

$$\rho_s = \frac{2\alpha}{2^{\mu_s}} \cong 0.12 \text{ mrad } (0.007^\circ). \quad (4.12)$$

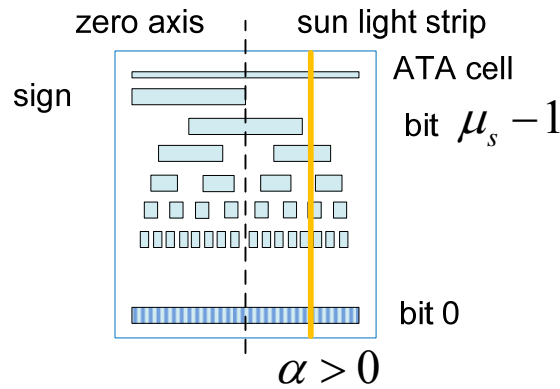


Figure 5. Gray code array.

Denoting the measured digital codes by a two-axis sensor with \tilde{y}_0 and \tilde{y}_1 , the measured angles are obtained through calibrated look-up tables

$$\begin{aligned} \tan \alpha &= f_\alpha(\tilde{y}_0) \\ \tan \beta &= f_\beta(\tilde{y}_1) \end{aligned} \quad (4.13)$$

Remarks. Two-axis digital sensors are implemented in the same way as analogue sensors (see Figure 3).

Modern sun sensors [10] are based on CCD cameras and CMOS Active Pixel Sensors¹ (APS).

TBC with parameter table

¹ An active-pixel sensor is a CMOS pixel made by a photodetector and an active amplifier, to be arranged in an array for imaging processes. They are emerging as an alternative to CCD arrays.

4.3.2 *Horizon sensors*

4.3.2.1 Introduction

The large part of space missions, except those aimed at solar system exploration, are carried on by satellites orbiting at a distance from the Earth where the viewing solid angle of the Earth disk is wide. At an altitude of 500 km the Earth solid angle is 3.9 sr (steradian), at an altitude of 5000 km the solid angle reduces to 1.1 sr, at 36000 km (geostationary orbits, GEO) the Earth solid angle is 0.072 sr. For comparison, the Sun solid angle as viewed from the Earth is about 7 msr (millisteradian), whereas the star solid angle is less than 0.1 psr (picosteradian).

At such orbits the Earth is the second among the brightest object in the sky. Since the solid angle is large, the Earth cannot be considered a point-wise, but an extended light source. This requires specific observation and measuring concepts in order to retrieve the viewing direction of extended objects. To this end, horizon is defined as the locus of points where the lines exiting from an observer in the space (considered as a point) become tangent to the planet surface. Actually, the horizon is uncertain and variable because of the reflected/emitted radiation, of the atmosphere and of surface irregularities.

Usually the Earth (planet) light can be observed either in the visible or in infrared wavelength. The infrared Earth horizon (centered on 15 μm) is more constant and allows higher accuracy. The Earth image in the visible band is very sensible to albedo (the reflected sunlight), thus affecting measurements with significant and not acceptable errors. Although the most common horizon sensors are built for the Earth observation, concept and implementation may apply to other planets.

4.3.2.2 Construction

Horizon sensors are available under various solutions; some of them are strictly dependent on the kind of satellite, as for instance to spinning spacecrafts.

The great part of the sensors consist of four parts:

- 1) a scanning mechanism,
- 2) an optical subsystem,
- 3) a detecting device,
- 4) an electronic unit.

Type and want of a scanning mechanism provide the sensors name:

- 1) static horizon sensor: they have a field-of-view larger than the entire Earth's edge (limb),
- 2) body-mounted horizon sensors: they lack the scanning mechanism since the latter is implemented by the satellite attitude motion itself (spin),

- 3) wheel-mounted horizon sensors: they are similar to body-mounted sensors, but are mounted on a wheel making the whole sensor to rotate,
- 4) non-spinning Earth sensors: only the optical head is made rotating to scan the Earth image.

The optical subsystem consists of a lens assembly focusing the planet image on the focal plane of the detecting device and of a filter for tuning the optics sensitivity to within the spectral band of interest. Depending on the type of sensor, the optical subsystem is equipped with a scanning motor. The detecting devices may vary from photodiodes, to thermopiles (a series of thermocouples) and to thermistors¹. The measurement time and the consequent measurement delay depend on the scanning time. In the case of thermal sensors (sensible to infrared spectrum) the measurement response time may become rather long due to thermodynamics, thus narrowing the sensor bandwidth.

4.3.2.3 Body-mounted sensors

They are employed by spinning satellites and their concept and implementation is the simplest one. They consist of an optical head with a narrow field-of-view (FOV) and of a detector in the visible or infrared spectral band. It is mounted in such a way that the optical axis moves along a cone about the spin axis of the satellite (see Figure 6). The sensor detects or not the Earth presence in the FOV. Given the spin rate and the Earth observation time from the in-trigger point to the out-trigger point, the observed Earth spherical chord is measured, depending on the spin axis attitude. The spin axis attitude is kept constant by keeping the sensor cone aperture constant. Accuracy constraints impose a narrow FOV, which may render difficult the Earth acquisition without other sensors.

¹ Thermistors are resistors whose resistance varies significantly (and in nonlinear way) with temperature, more than in standard resistors. They differ from resistance temperature detectors (RTD) in that the thermistor material is usually a ceramic or polymer, while RTDs use pure metal.

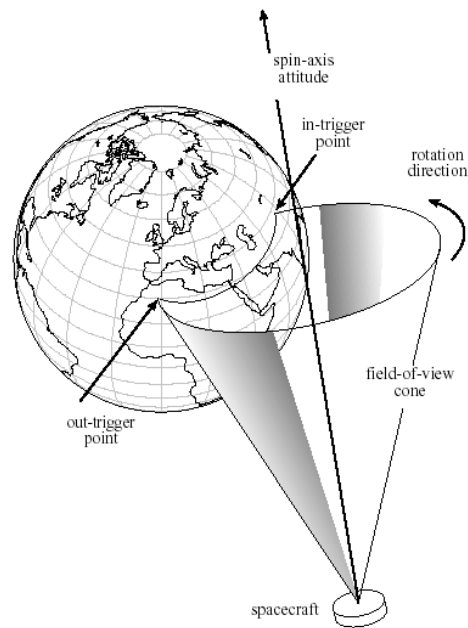


Figure 6. Body mounted horizon sensor.

4.3.2.4 Static horizon sensor

The field-of-view is larger than the entire Earth's edge (limb). An infrared detector is used, like micro-machined thermopiles. It provides two angle orientation (roll around x-axis, pitch around y-axis) with respect to local vertical (nadir, z-axis). An example is the Goodrich multi-mission horizon sensor in Figure 7.

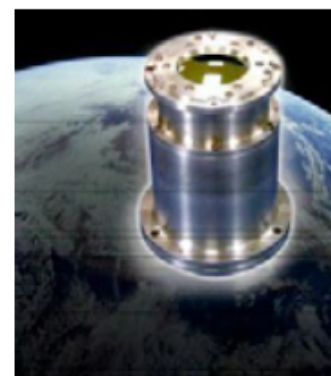
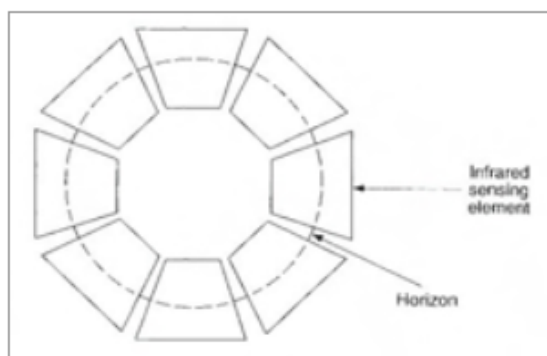


Figure 7. Static horizon sensor.

Typical performance data at a geostationary orbit are in Table 2.

Table 2. Typical performance of a multi-mission horizon sensor
--

No	Parameter	Symbol	Unit	Value	Comment
0	Infrared spectral band		μm	14 to 16	
1	Field of view		rad \times rad	0.28 \times 0.18	
2	Pitch accuracy at null (+)		mrad	0.23 (3 σ)	E-W @ GEO
3	Roll accuracy at null (+)		mrad	0.52 (3 σ)	N-S @ GEO
4	Spacecraft keep-out cone		rad	0.26	
5	Operating range: roll (pitch)		rad	± 0.035 (± 0.01)	
6	Acquisition range		rad	0.28	

4.3.2.5 Wheel mounted sensors

They consists of two telescopes mounted in a way to make orthogonal scan. Each optical system is made by scanning mirror, whereas the detecting device is a thermistor.

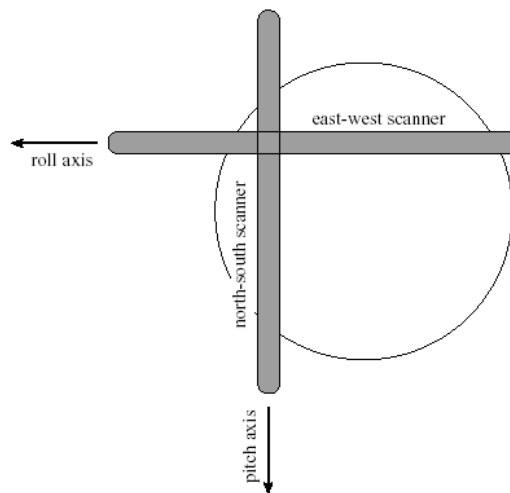


Figure 8. Wheel mounted horizon sensors: scanning principle.

4.3.3 *Magnetometers*

4.3.3.1 Introduction

Magnetometers measure magnitude and direction of magnetic field. They are light, reliable, low power; they operate in a wide temperature range and are free of moving parts. Unfortunately, uncertainty on the Earth or planet magnetic field and cross-coupling with on-board magnetic fields greatly reduce their accuracy, which is worse than 9 mrad. Moreover, since a planet (when exists) magnetic field decreases in magnitude with the altitude (it is proportional to the inverse of the distance cube from the planet CoM), there is a bound to the orbit altitude which is around 1500 to 200 km for the Earth. Magnetometers for higher altitudes can be kept as ad-hoc applications. They are commonly used for damping the angular rate of Low-Earth-Orbit satellites in the early phases of the mission. They include a sensible part (the magnetic sensor) and a detection electronics.

A pair of physical principles are employed in their construction:

- 1) Quantum (or proton precession) magnetometers employ matter properties like the nuclear magnetic resonance(NMR). They are used for very specific applications as scientific payloads since they are heavy and expensive.
- 2) Inductive magnetometers employ magnetic induction (Faraday's law) and are manufactured in two different ways: (i) search-coil and (ii) fluxgate. As they are commonly used as attitude sensors, will be fully treated here.

Magnetometers are appearing in portable mobile phones.

4.3.3.2 Quantum magnetometers

The proton precession magnetometer operates on the principle that the protons in all atoms are spinning on an axis aligned with the magnetic field. Ordinarily, protons tend to line up with the earth's magnetic field. When subjected to an artificially-induced magnetic field, the protons will align themselves with the new field. When this new field is interrupted, the protons return to their original alignment with the earth's magnetic field. As they change their alignment, the spinning protons precess, or wobble, much as a spinning top does as it slows down. The frequency at which the protons precess is directly proportional to the strength of the earth's magnetic field. This is the Proton Gyromagnetic Ratio, equal to 0.042576 Hz/nT. For example, in an area with a field strength of 60000 nT (such as on the Earth ground), the frequency of precession would be approximately 2500 Hz. A direct current flowing in an inductor creates a strong magnetic field around a hydrogen-rich fluid, causing some of the protons to align themselves with that field. The current is then interrupted, and as protons realign themselves with ambient the magnetic field, they precess at a frequency that is directly proportional to the magnetic field. This produces a weak alternating magnetic field that is

picked up by a separated inductor, amplified electronically, and fed to a digital frequency counter whose output is typically scaled and displayed directly as field strength or output as digital data. The two main sources of measurement errors are magnetic impurities in the sensor and errors in the measurement of the frequency.

4.3.3.3 Inductive magnetometers

The Faraday's law of induction states that a variable magnetic flux $\phi(\vec{b})$ produced by the induction \vec{b} and linked with a closed circuit Γ encircling a surface Σ , induces an electromotive force (energy per unit charge) E [V] having the value

$$E = -\frac{d\phi(\vec{b})}{dt}, \quad \phi(\vec{b}) = \int_{\Sigma(\Gamma)} \vec{b} \cdot \vec{n} dA, \quad (4.14)$$

where the positive sign corresponds to assuming positive current entering the positive voltage terminal (passive sign convention), \vec{n} is the normal to the surface element of area dA .

In the case of search-coil magnetometers the closed circuit is made by the N coils of a solenoid wound around a ferromagnetic nucleus with magnetic permeability μ and section area A . The induced electromotive force holds

$$E = \mu N A \frac{d(\vec{b} \cdot \vec{n})}{dt}, \quad (4.15)$$

where \vec{n} is the direction of the solenoid axis. The electromotive force depends on the variation of \vec{b} and on the rotation of \vec{n} .

They are used on spinning satellites. Assume for simplicity the solenoid axis is orthogonal to the inertial spinning axis equal to the body pole $\vec{i}_3 = \vec{b}_3$. Then the inertial components of \vec{n} and \vec{b} are

$$\mathbf{n}_i = \begin{bmatrix} \cos(\omega_s t + \psi_0) \\ \sin(\omega_s t + \psi_0) \\ 0 \end{bmatrix}, \quad \mathbf{b}_i = \begin{bmatrix} b_{i1} \\ b_{i2} \\ b_{i3} \end{bmatrix}. \quad (4.16)$$

Consequently, assuming constant \vec{b} in the rotation period $P_s = 2\pi / \omega_s$ and $\psi_0 = 0$, (4.15) becomes

$$E = \mu N A b_{i2} \omega_s \cos(\omega_s t + \psi_{12})$$

$$b_{i2} = \sqrt{b_{i1}^2 + b_{i2}^2}, \quad \tan \psi_{12} = \frac{b_{i1}}{b_{i2}}. \quad (4.17)$$

Signal phase ψ_{12} provides the magnetic field direction in the plane orthogonal to spin axis.

Fluxgate magnetometers [12] exploit saturation and hysteresis of ferromagnetic materials. They were invented in the 1930s by V. Vacquier at Gulf Research Laboratories, USA. Vacquier applied them during World War II as an instrument for detecting submarines. Fluxgate magnetometers for space applications have been produced since 1958 (Sputnik 3) and are nowadays commonly used instruments aboard satellites. Around a magnetic nucleus two coils are wound as in Figure 9.

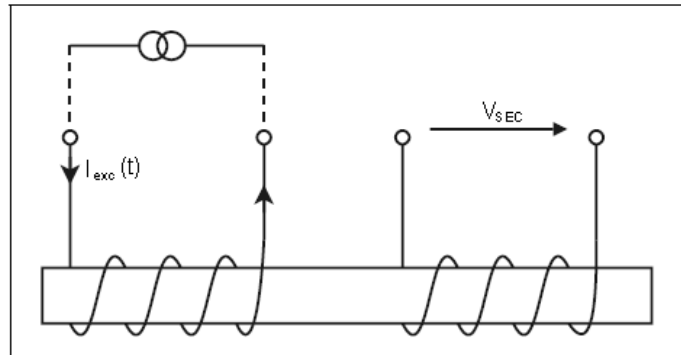


Figure 9. Basic circuit of a fluxgate magnetometer

- 1) The excitation (or drive) coil is supplied with a known sine wave of current I_{exc} such to saturate (positively and negatively) the magnetic core.
- 2) Changes of the magnetic flux through the pick-up coil give rise to induced voltage V_{sec} . This will generate a magnetic field that in turn induces a voltage in the pick-up coil, and can thus be detected. An example of the signal from the pick-up coil is found in Figure 10.

Without an external field, the magnetic flux in the core will depend only on the field created by the drive coil. The core will stay in magnetic saturation most of the time, with the two directions of saturation taking equal parts of the excitation period. Change of the magnetic flux between the two states of saturation will induce voltage in the pick-up coil. If a component of the external field is present along the core axis, the time spent in one of the saturations will increase. This will lead to shifts of the induced pick-up coil voltage as in Figure 10. This shift can be detected and from the size of the shift the external field flux density can be derived.

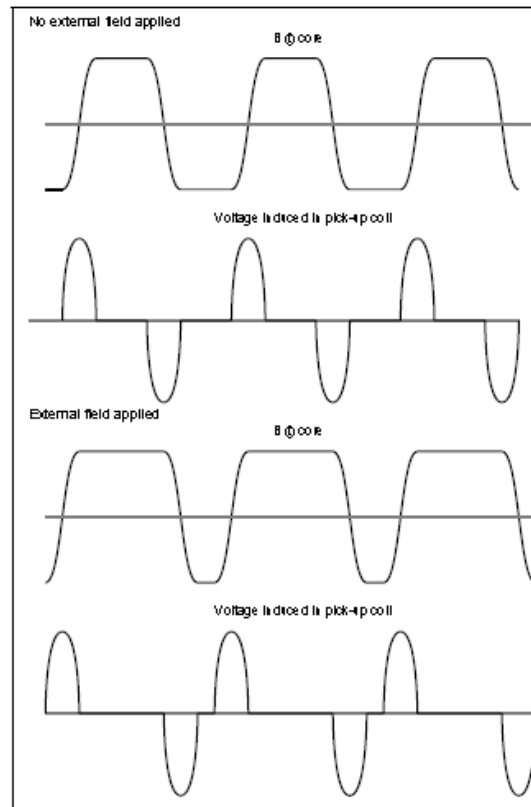


Figure 10. Signal in the pick-up coil.

To cancel the large signal generated by the excitation coil an additional core can be added within the pick-up coil with an excitation winding in the opposite direction from the first one as seen in Figure 11. To guarantee that both coils are driven by the same current the coils are connected in series.

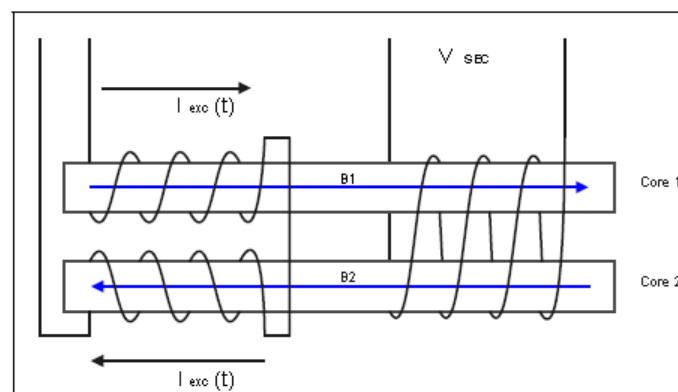


Figure 11. Compensation for the excitation field.

The coils are wound in opposite direction and are thus generating magnetic fields of opposite direction in the cores. In absence of an external magnetic field the magnetic fluxes of the

opposite cores will cancel each other exactly and no voltage will be induced in the pick-up coil (Figure 12, left). If an external field is present, the pick-up signal will arise due to the relative shifts of the flux changes in each core. It will be dominated by the second harmonic of the excitation frequency, and contain higher even harmonics (Figure 12, right).

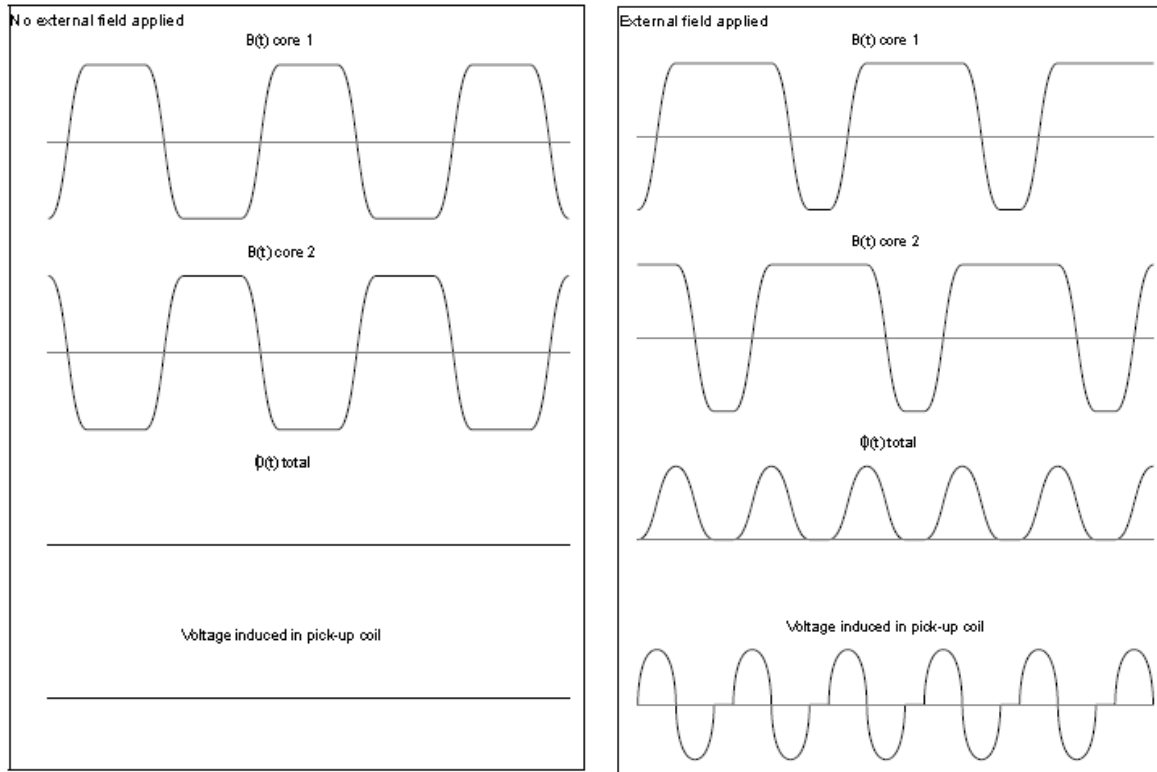


Figure 12. Voltage induced in the pick-up coil without and with external field.

An alternative view on the fluxgate mechanism is that the alternating state of saturation will cause the apparent permeability μ_a to vary between 0 and μ_{\max} . When the excitation field passes zero the signal will only depend on the external field and this can be seen as a phase/time shift from a signal depending only on the magnetization of the core. This means that information of the external field can be observed at twice the drive frequency.

The voltage induced in the pick-up coil is described by

$$V_{sec} = NA \frac{db}{dt}, \quad (4.18)$$

$$b = \mu_a b_{ex}$$

where N is the number of windings in the pick-up coil, A is the coil cross sectional area and b is proportional to the (constant) external induction b_{ex} through the apparent magnetic permeability μ_a of the core. The apparent magnetic permeability differs from the material

specific permeability μ_r due to the demagnetizing factor D , which varies with the geometry of the core and accounts for the internal magnetization of the material¹

$$\mu_a = \frac{\mu_r}{1 + D\mu_r}. \quad (4.19)$$

The demagnetizing factor of a rod of length L and diameter $d \ll L$ approaches $D \cong (L/d)^{-(1+\alpha)} \ll 1$, $\alpha < 1$. Now (4.18) can be rewritten as the basic fluxgate equation and is a function of the relative permeability of the core

$$V_{sec} = NAb_{ex} \frac{d\mu_a}{dt} = NAb_{ex} \frac{(1-D)}{(1+D\mu_r)^2} \frac{d\mu_r}{dt}. \quad (4.20)$$

4.3.3.4 Body attitude from magnetometer

TBC

4.3.4 *Star trackers*

They are the most accurate attitude sensors and in the modern versions they directly provide the attitude measurements. They consist of

- 1) optical subsystem,
- 2) semiconductor detector and the relevant readout electronics,
- 3) electronics unit.

The optical subsystem consists of a variable set of lenses (up to 7 and 8 lenses). The off-the-shelf sensors employ a matrix of photodetectors (pixel) either Charged Coupled Devices (CCD) or Active Pixel Sensors (APS). The latter are more resistant to environment radiations, need less power and use a simpler proximity electronics (conversion electronics from analogue data to digital data), which in the whole leads to a mass reduction. The assembly made by focal plane and proximity electronics is referred to as optical head.

¹ The actual exciting field H is obtained by subtracting the material (de-)magnetizing field $H_d \cong D\mu_r H$ from the external field H_{ex} , which leads to $(1 + D\mu_r)H = H_{ex}$ and to $(1 + D\mu_r)B = \mu_r B_{ex}$.

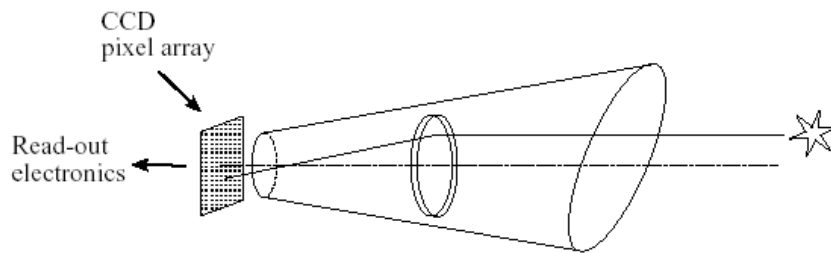


Figure 13. Sketch of a star tracker.

TBC

4.3.4.1 Construction

The first star trackers were only capable of tracking star images in the focal plane. As such they were completely served to the control unit of the AOCS, which provided with the star coordinates in the focal plane.

In the last ten years, autonomous star trackers have made available capable of providing the sensor attitude. They have exploited progress in micro-electronics and are endowed with memory and computing capabilities. Using star catalogues, autonomous star trackers permit attitude determination in a continuous and accurate way.

The optical subsystem FOV varies between $10^\circ \times 10^\circ$ and $40^\circ \times 40^\circ$. Accuracy is determined by pixel size, FOV size and the algorithms. Usually the errors range between 5 and 50 μrad for the optical axis attitude (two angles) and between 0.05 and 0.3 mrad for rotations about the optical axis.

The most recent devices include algorithms for the degraded performance in the case of higher angular rates, from 0.02 to 0.1 rad/s. In such cases attitude cannot be determined and only angular rates are provided.

Limitations of start trackers are

- 1) Stray light. Sensitivity to angular separation from bright objects like Earth, Sun and Moon. Their light entering the FOV degrades performance and may impede any measurements. The minimum angular separation from Earth and Moon is about 20° , and from the Sun about 30° .
- 2) Radiation sensitivity.
- 3) Angular speed. Accuracy depends on the angular speed. the first attitude determination is usually guaranteed for angular speed less than 0.01 to 0.02 rad/s.
- 4) Complexity. Star tracker algorithms are rather complex and comparable to the algorithms of the overall ACS.

5) Power and mass are rather high.

4.3.4.2 Principles

Star sensors may be very accurate as stars are point light sources, and their direction is practically inertial during a spacecraft mission life. The present technology is based on CCD (charge coupled device) and APS cameras [10] and accurate star Catalogues as that produced by Hipparcos mission.

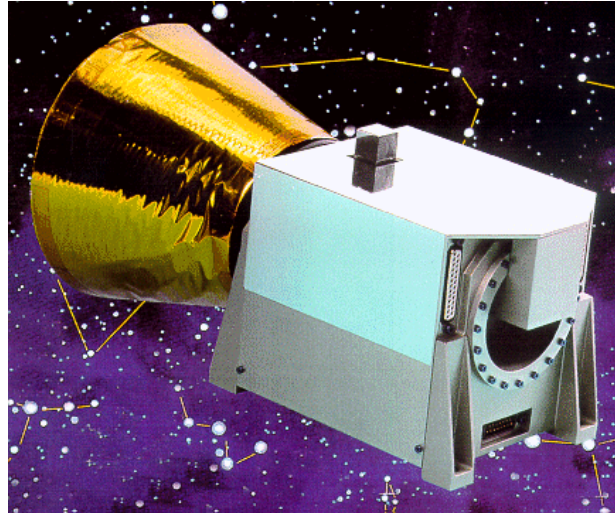


Figure 14. A typical star tracker.

Each star k differentiates because of its magnitude m_k and direction \vec{v}_k opposite to the light wavefront. Magnitude¹ is a logarithmic scale accounting for the 10 base logarithm of the ratio of the light fluxes (luminosity) S_k and S_j of two stars, as follows

$$m_j - m_k = -2.5 \log_{10} (S_j / S_k), S_k \left[\text{W/m}^2 \right]. \quad (4.21)$$

Magnitude needs to assign a reference magnitude, say zero, to a star, and then derive the other ones through (4.21). The zero apparent magnitude is assigned to star Vega (α Lyrae). The brightest star is Sirius (α Canis Major) with $m_{\text{Sirius}} = -1.49$ magnitude. Sun has magnitude $m_s = -26.8$.

Remarks. Reader must be warned that brighter stars have lower magnitude -possibly negative - than fainter stars! For instance a flux ratio of 10 corresponds to a magnitude difference of -2.5.

¹ One must distinguish absolute from apparent magnitude, the former being the latter correct through the star distance from the Earth. Apparent magnitude measure the star brightness as seen from the Earth in the visible band, denoted with V.

Star light is emitted along different spectra, which is very much dependent on star surface temperature θ ; from blue to red they are classified into main seven types denoted with O (blue, $\theta > 30000$ K), B (blue to blue white), A (white, Vega, Sirius), F (yellowish white), G (yellow, the Sun), K (orange, Aldebaran) and M (red, Betelgeuse, $\theta < 3700$ K).

Star unit vector \mathbf{v}_k coordinates are usually given in the J2000 frame, though right ascension α_k and declination δ_k

$$\mathbf{v}_k = \begin{bmatrix} \cos \alpha_k \cos \delta_k \\ \sin \alpha_k \cos \delta_k \\ \sin \delta_k \end{bmatrix}. \quad (4.22)$$

A generic star unit vector \mathbf{v}_s in the instrument frame $\mathcal{R}_s = \{O_s, \bar{\mathbf{i}}_s, \bar{\mathbf{j}}_s, \bar{\mathbf{k}}_s\}$ where the last axis is the optical axis, is written as the rotation of the optical axis around $\bar{\mathbf{j}}_s$ by θ_s and then around the rotated $\bar{\mathbf{i}}_s$ by φ_s

$$\mathbf{v}_s = X(-\varphi_s)Y(-\theta_s) \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi_s & -\sin \varphi_s \\ 0 & \sin \varphi_s & \cos \varphi_s \end{bmatrix} \begin{bmatrix} \sin \theta_s \\ 0 \\ \cos \theta_s \end{bmatrix} = \begin{bmatrix} \sin \theta_s \\ -\sin \varphi_s \cos \theta_s \\ \cos \varphi_s \cos \theta_s \end{bmatrix}. \quad (4.23)$$

The instrument FOV is defined by the largest angles entering the image: they are assumed equal

$$\text{FOV: } |\theta_s| \leq \theta_{s,\max} / 2, \quad |\varphi_s| \leq \varphi_{s,\max} / 2, \quad \theta_{s,\max} = \varphi_{s,\max} \quad (4.24)$$

The star light must be kept as an electromagnetic wave with a planar wave-front when hitting the entrance pupil of the star tracker optics. The optics role is to focus the wavefront to a point in the image or focal plane, by collecting as much energy as possible. The finite size of the entrance pupil (usually a circle of radius a) prevent ideal focusing into a point image, but creates a diffraction pattern (or point spread function, PSF) described by a Bessel function of the 1st kind (2D Fourier transform of a cylinder)

$$p(r) = 2 \left(J_1(\mathcal{G}(r)) / \mathcal{G}(r) \right)^2$$

$$\mathcal{G}(r) = \frac{2\pi r a}{\lambda f} \geq 0 \quad (4.25)$$

where r is the radial distance on the focal plane from the optical axis, λ is the light wavelength ($\lambda = 0.4 \div 0.8 \mu\text{m}$ in the visible spectrum) and f the focal length. The PSF contour has different zero-contours, the first one being called Airy disk with radius

$$r_0 \cong 1.22 \lambda f / d, \quad d = 2a, \quad (4.26)$$

inside which falls the 84% of the incoming light energy, being the index of image resolution. Figure 15 shows a radial profile of a circular aperture PSF (normalized) and the 3D profile (right) for

$$f = 0.035 \text{ m}, a = 0.0115 \text{ m}, \lambda = 0.6 \mu\text{m} \text{ (yellow)} \quad (4.27)$$

corresponding to $r_0 \cong 1.1 \mu\text{m}$. The ratio $N = f / d$ is called the f-number and the ratio is usually written as f/N . Smaller f-numbers are called faster, the fastest N of the human eye being 2.1, corresponding to $r_0 \cong 1 \mu\text{m}$ at $\lambda \cong 0.4 \mu\text{m}$ (visible blue), whereas cone spacing in the fovea is about $r_c \cong 2.5 \mu\text{m}$ corresponding to $N=5$.

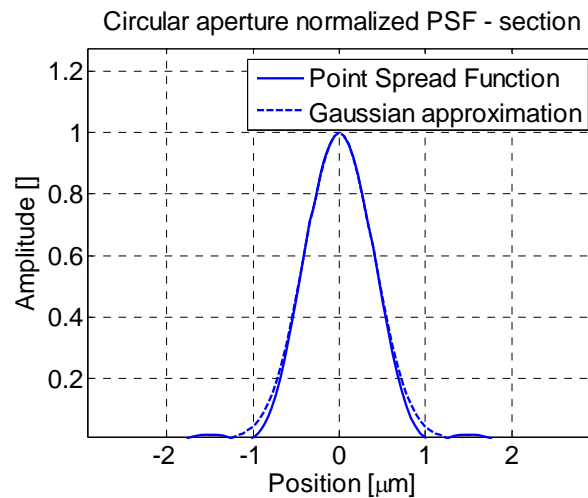


Figure 15. Normalized PSF of a circular aperture an Gaussian approximation: radial cross-section.

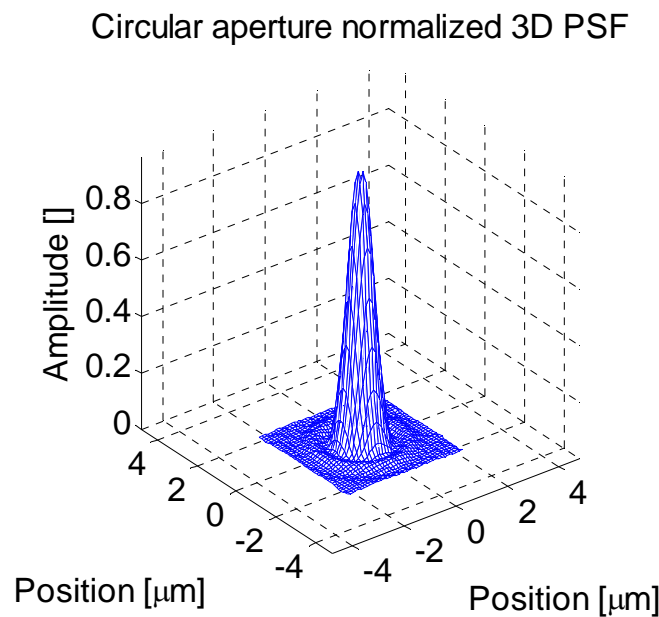


Figure 16. Normalized PSF of a circular aperture: 3D profile.

PSF can be approximated as 2D Gaussian density function

$$p(r) = e^{-\frac{r^2}{2\sigma^2}}, \quad \sigma \cong 0.44\lambda f / d \cong r_0 / 2.77. \quad (4.28)$$

Since pixel size (square pixels) of CCD and APS sensors used in star trackers is close or greater than $10 \mu\text{m}$, the above PSF should be enlarged to cover more than a single pixel, which is necessary to compute the PSF centroid and localize the star. That is obtained by defocusing the star image, so as the light spot covers at least 3×3 pixels. Defocus is obtained by displacing Δz the sensor plane away from the focal plane along the optical axis. then the image tends to become a circular spot of radius

$$r_1 = \frac{\Delta z a}{f} = \frac{1}{2} \frac{\Delta z}{N}. \quad (4.29)$$

The circular spot is called circle of confusion or blur spot.

Consider a star tracker with the image plane made by $n \times n$, $n = 1024$, square pixels of size $d_p = 15 \mu\text{m}$, which provides an image size $d_i = n d_p$. The FOV $\varphi_{s,\max}$ is then related to the image size by

$$d_i = n d_p \geq \varphi_{s,\max} f, \quad (4.30)$$

which leads to a FOV

$$\varphi_{s,\max} \leq \frac{n d_p}{f} \cong 0.44 \text{ rad } (25^\circ). \quad (4.31)$$

Moreover, given the pixel size, assuming the circle of confusion to cover at least three pixels, and assuming Gaussian shape (rough approximation)

$$2r_1 \geq 3d_p \Rightarrow \Delta z = \frac{3d_p f}{d} \cong 30 \mu\text{m}, \quad \sigma = \frac{1.5d_p}{2.77} \cong 8.1 \mu\text{m}. \quad (4.32)$$

Denote with (x_{sk}, y_{sk}) the centroid C_k of the star k PSF in the instrument cartesian coordinates. They are related with the star unit vector through focal length, by considering the unit vector passing through the optical centre F with coordinates $\mathbf{f} = [0 \ 0 \ f]^T$ and the centroid C_k . Then it holds

$$\vec{\mathbf{f}} = \overrightarrow{O_s C_k} + \vec{\mathbf{v}}_k \lambda \Rightarrow \begin{bmatrix} 0 \\ 0 \\ f \end{bmatrix} = \begin{bmatrix} x_{sk} \\ y_{sk} \\ 0 \end{bmatrix} + \lambda \begin{bmatrix} \sin \theta_{sk} \\ -\sin \varphi_{sk} \cos \theta_{sk} \\ \cos \varphi_{sk} \cos \theta_{sk} \end{bmatrix}. \quad (4.33)$$

The third coordinate gives the camera equations

$$\begin{aligned}\lambda &= \frac{f}{\cos \varphi_{sk} \cos \theta_{sk}} \\ \tan \varphi_{sk} &= -\frac{y_{sk}}{f} \\ \tan \theta_{sk} &= \frac{x_{sk}}{f} \cos \varphi_{sk}\end{aligned}\quad (4.34)$$

Under small angles

$$\begin{aligned}\lambda &\cong f \\ \varphi_{sk} &\cong -\frac{y_{sk}}{f} \\ \theta_{sk} &\cong \frac{x_{sk}}{f}\end{aligned}\quad (4.35)$$

Equations (4.35) should be modified to account for the optics and image plane distortions.

The centroid may be simply computed from the pixel intensities $z_k(i, j)$ and pixel coordinates (i, j) where

$$\begin{aligned}x_s &= (i + 0.5)d_p \\ y_s &= (j + 0.5)d_p\end{aligned}\quad (4.36)$$

implying

$$-n/2 \leq i, j < n/2. \quad (4.37)$$

Then

$$x_{sk} = \frac{\sum_{i=i_{k0}}^{i_{k0}+n_k} z_k(i, j)i}{\sum_{i=i_{k0}}^{i_{k0}+n_k} z_k(i, j)}, \quad y_{sk} = \frac{\sum_{j=j_{k0}}^{j_{k0}+m_k} z_k(i, j)j}{\sum_{j=j_{k0}}^{j_{k0}+m_k} z_k(i, j)}. \quad (4.38)$$

Pixel intensity should be cleaned by the estimated background to the sky and pixel noise itself. More sophisticated algorithms may be used involving the calibrated PSF. The centroid coordinate errors (root mean square) can be usually reduced to less than 0.1 of the pixel size, yielding an angular error (called noise equivalent angle, NEA)

$$\sigma_\theta = \sigma_\varphi = 0.1 \frac{\varphi_{\max}}{n} \cong 42 \mu\text{rad} (\cong 8.5 \text{ arcsecond}). \quad (4.39)$$

The star Catalogue to be added to star tracker depends on the FOV and the number of stars n_s to be continuously inside the FOV: at least two widely separated stars are needed to reconstruct the instrument attitude, but $n_s \geq 3$ stars are needed to recognize the sky region

when star tracker is switched-on (the 'lost-in-space' procedure). Were the star distribution in the sky be uniform at least $N_s = n_s 4\pi / (\varphi_{\max})^2$ would be necessary. Unfortunately star distribution is not uniform, because of the Milky Way, which implies a factor $\gamma_s > 1$ must be added. With the above values $N_s \geq 390 @ \gamma_s \geq 2$. The number of stars then decide the maximum magnitude to be visible by the star tracker. For instance the star density brighter than $m = 5.0$ is about 100 rad^{-2} , which is sufficient to the purpose as the total would be around $4\pi\delta_s \approx 1200$.

Table 3. Typical star tracker data					
No	Parameter	Symbol	Unit	Value	Comments
0	Mass	m	kg	4	
1	Power	W	W	7.6	max 9.5 W
2	Bias	b_s	μrad	<50	
3	Noise: across optical axis		μrad	<50	
4	Noise: optical axis		μrad	<270	
5	FOV		rad×rad	0.44×0.44	25°×25°
6	Tracking rate				

4.4 Inertial sensors - accelerometers

4.4.1 The proof-mass concept

Accelerometers are employed for measuring the non gravitational accelerations of a spacecraft. The essential implementation is that of a small mass (sometimes called proof-mass) in a cage inside the spacecraft (Figure 17), thus only subject to gravity \vec{g} acceleration, but not to the forces $\vec{F} = m\vec{a}$ acting on the satellite. That implies the proof-mass, if unrestrained, to drift in the cage and randomly bounce on the cage walls. An internal force \vec{F}_a between cage and mass such as to keep a mass-to-cage fixed distance will be clearly proportional to the spacecraft acceleration \vec{a} not sensed by the proof-mass, less some noise. The above concepts will be clarified through equations. The attitude control of the mass, if needed, will be neglected.

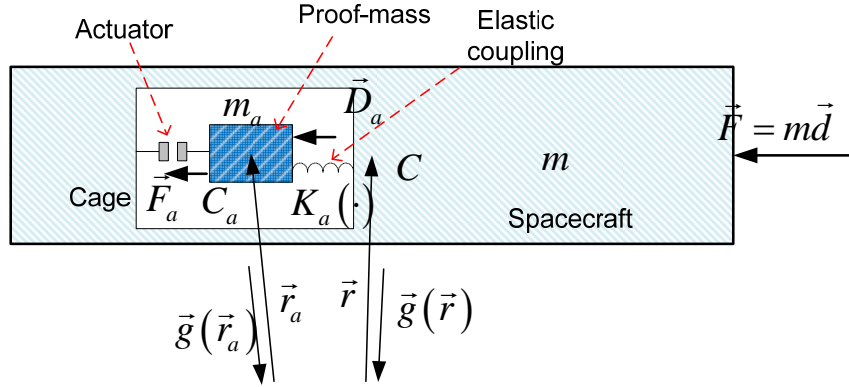


Figure 17. Layout of the proof-mass within a spacecraft.

Denote the spacecraft CoM C position and rate vector with \vec{r} and \vec{v} , the mass with m . Similar notations apply to the proof-mass CoM C_a adding the subscript a , with the constraint

$$m_a \ll m. \quad (4.40)$$

Spacecraft and mass Newton's equations hold with the aid of Figure 17

$$\begin{aligned} \dot{\vec{v}}(t) &= -\vec{g}(\vec{r}(t)) + (\vec{F}(t) - \vec{F}_a(t)) / m \\ \dot{\vec{v}}_a(t) &= -\vec{g}(\vec{r}_a(t)) + \vec{F}_a(t) / m_a + (\vec{D}_a(t) - K_a(\vec{r}_a(t) - \vec{r}(t))) / m_a, \end{aligned} \quad (4.41)$$

having partitioned the parasitic forces acting on the proof-mass, to be ideally zero, into

- 1) (nonlinear) elastic force $K_a(\vec{r}_a(t) - \vec{r}(t))$ depending on the relative position $\Delta\vec{r} = \vec{r}_a(t) - \vec{r}(t)$, whose contribution to spacecraft has been kept as negligible,
- 2) and a generic disturbance $\vec{D}_a(t)$ only perturbing the proof-mass.

Since the active force \vec{F}_a is computed along an instrument frame $\mathcal{R}_a = \{C, \vec{i}_a, \vec{j}_a, \vec{k}_a\}$, rigidly connected with the spacecraft, hence rotating at a rate $\vec{\omega}$, but in generally misaligned with the body frame, so must be the measured acceleration.

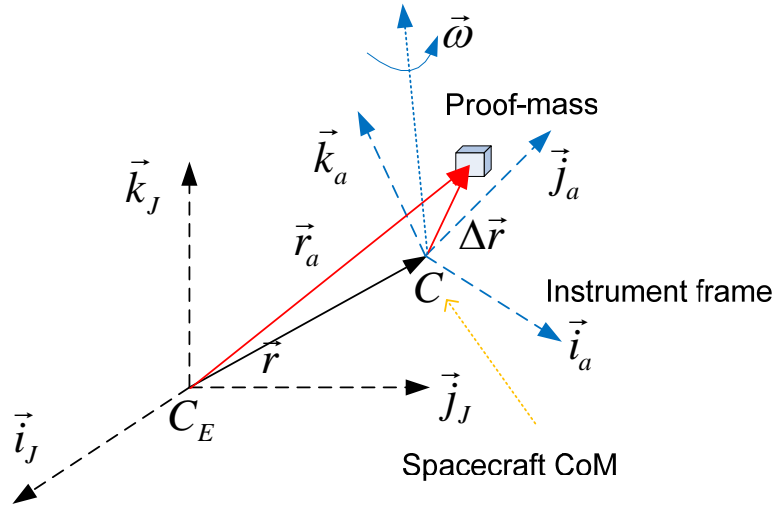


Figure 18. Inertial and instrument frames.

To this end, express $\dot{\vec{v}}_a(t)$ wrt to $\dot{\vec{v}}(t)$, with the aid of relative motion equations. First express the proof-mass relative velocity

$$\vec{v}_a(t) - \vec{v}(t) = \Delta \vec{v}(t) + \vec{\omega}(t) \times \Delta \vec{r}(t), \quad (4.42)$$

where $\Delta \vec{v} = \dot{\Delta \vec{r}}|_a$ is a ‘local’ rate, i.e. wrt to a spacecraft frame (instrument or body). the take the derivative of (4.42)

$$\dot{\vec{v}}_a(t) - \dot{\vec{v}}(t) = \Delta \dot{\vec{v}}(t) + \dot{\vec{\omega}}(t) \times \Delta \vec{r}(t) + \vec{\omega}(t) \times (\vec{\omega}(t) \times \Delta \vec{r}(t)) + 2\vec{\omega}(t) \times \Delta \vec{v}(t), \quad (4.43)$$

where the ‘local’ acceleration $\Delta \dot{\vec{v}} = \dot{\Delta \vec{v}}|_a$ has been employed. Taking the difference $\dot{\vec{v}}_a(t) - \dot{\vec{v}}(t)$ in (4.41), comparing it to LHS in (4.43), and defining the measured acceleration as

$$\mathbf{a}(t) = \mathbf{F}(t)/m + \dot{\vec{\omega}}(t) \times \Delta \mathbf{r}(t) + \vec{\omega}(t) \times (\vec{\omega}(t) \times \Delta \mathbf{r}(t)) + 2\vec{\omega}(t) \times \Delta \mathbf{v}(t), \quad (4.44)$$

the final equation follows

$$\Delta \dot{\vec{v}}(t) = (-K_a(\Delta \mathbf{r}(t)) + \mathbf{D}_a(t))/m_a - \mathbf{a}(t) + \mathbf{F}_a(t)/m_a, \quad (4.45)$$

where (4.40) has been adopted, and arrows have been dropped, since instrument coordinates have been assumed.

Note the elastic term may include the differential gravity contribution, if not kept as negligible.

One immediately verifies from (4.45) the introductory concepts: keeping the mass-to-cage distance constant, i.e. $\Delta \dot{\vec{v}}(t) = 0$ in (4.45), makes $\mathbf{F}_a(t)$ proportional to $\mathbf{a}(t)$, less some disturbance terms to be minimized or corrected. In this way, $\mathbf{F}_a(t)$ becomes at the same time

command and measurement, a concept usually referred to as servo-accelerometer. As a discrete-time measurement, it may be expressed as

$$\mathbf{y}_a(k) = \mathbf{F}_a(kT_a) / m_a, \quad (4.46)$$

where T_a is the accelerometer time unit, usually shorter than 1 ms.

As a command, control law and available measurements must be defined. To this end, fix a reference mass-to-cage position $\Delta \mathbf{r}$ to be kept, and assume a suitable 3D sensor (usually a capacitor) measuring the tracking error $\delta \mathbf{r}(i) = \Delta \mathbf{r}(i) - \Delta \mathbf{r}$, less some error $\mathbf{e}(i)$, and providing the real-valued measure

$$\mathbf{y}(i) = \delta \mathbf{r}(iT) + \mathbf{e}(i), \quad (4.47)$$

where $\mathbf{e}(i)$ includes also quantization errors and other components to be defined below, and T is the control time unit, where usually $T \leq T_a$.

4.4.2 Continuous-time control

To simplify treatment a single axis will be retained and first continuous-time feedback is treated. It is rather direct to derive the following 2nd order state equation from (4.45)

$$\begin{aligned} \begin{bmatrix} \delta \dot{r} \\ \Delta \dot{v} \end{bmatrix} (t) &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \delta r \\ \Delta v \end{bmatrix} (t) + \begin{bmatrix} 0 \\ b_a \end{bmatrix} u_a(t) + \begin{bmatrix} 0 \\ d_a(t) - a(t) \end{bmatrix}, \\ d_a(t) &= -K_a(\Delta \mathbf{r}(t)) / m_a + D_a(t) / m_a \end{aligned} \quad (4.48)$$

where $F_a / m_a = b_a u_a$, and u_a is the command voltage. Note the elastic term has been kept outside the state matrix since it is nonlinear, and has been made dependent on the whole relative position to account for interactions. Since usually a linear approximation is possible close to $\Delta \mathbf{r}$, one may rewrite

$$\begin{aligned} d_a(t) &= -s_a \nu_a^2 (\delta r(t) + \Delta \underline{r}) - \delta K_a(\Delta \mathbf{r}(t)) / m_a + D_a(t) / m_a, \\ s_a &= \pm 1 \end{aligned} \quad (4.49)$$

where a sign s_a has been added, cross-coupling, uncertainty and nonlinearity have been confined in the residual function $\delta K_a(\cdot)$, and ν_a is an angular frequency [rad/s]. A displacement sensor (capacitive) must be added, measuring deviation δr with respect to reference distance $\Delta \underline{r}$ less some error e and a scale factor k_y [V/m]

$$y(t) = k_y \delta r(t) + e(t). \quad (4.50)$$

The Laplace transform holds

$$y(s) = k_y \frac{1}{s^2} (b_a u_a(s) + d_a(s) - a(s)) + e(s). \quad (4.51)$$

Consider a simple PD control

$$\begin{aligned} u &= -k_p y - k_d (y - q) / \tau_d \\ \dot{q}(t) &= (y - q) / \tau_d \end{aligned} \quad (4.52)$$

whose transfer function holds

$$\frac{u(s)}{y(s)} = - \left(k_p + k_d \frac{s}{s\tau_d + 1} \right). \quad (4.53)$$

From (4.51) and (4.53), it is straightforward to derive the command Laplace transform

$$b_a u_a(s) = \frac{1}{\phi(s)} (N_a(s)(-d_a + a) - N_e(s)e(s)), \quad (4.54)$$

where

$$\begin{aligned} \phi(s) &= \tau_d s^3 + s^2 + s k_y b_a (k_d + k_p \tau_d) + k_y b_a k_p = \\ &= \tau_d s^3 + s^2 + 2\zeta_a \omega_a s + \omega_a^2 \\ N_a(s) &= s k_y b_a (k_d + k_p \tau_d) + b_a k_y k_p = 2\zeta_a \omega_a s + \omega_a^2 \\ N_e(s) &= s^2 N_a(s) / k_y \end{aligned} \quad (4.55)$$

In (4.55) the closed-loop angular frequency ω_a and damping $\zeta_a \cong 1$ have been defined. Assuming $\tau_d \ll \omega_a^{-1}$, the closed-loop eigenvalues, under $0 \leq \zeta_a \leq 1$, approximate

$$\begin{aligned} \lambda_{1,2} &\cong - \left(\zeta_a \pm j \sqrt{1 - \zeta_a^2} \right) \omega_a \\ \lambda_3 &\cong -1 / \tau_d \end{aligned} \quad (4.56)$$

The angular frequency $\omega_a = 2\pi f_a$ fixes the accelerometer BW. Within the BW, i.e. for $f < f_a$, (4.54) approaches

$$b_a u_a(s) = a(s) - d_a(s) - s^2 e(s) / k_y, \quad (4.57)$$

since $\lim_{s \rightarrow 0} N_a(s) / \phi(s) = 1$. Therefore the accelerometer error holds within the BW

$$n_a(s) = a(s) - b_a u_a(s) = d_a(s) + s^2 e(s) / k_y. \quad (4.58)$$

4.4.3 Typical accelerometer errors

In the frequency domain, the accelerometer error may be subdivided in four parts

- 1) bias, due to the mean value \underline{d}_a of the parasitic forces,
- 2) low-frequency drift, due to slow fluctuations of d_a (thermal fluctuations, ...),
- 3) flat noise, due to random components (white) of d_a (actuator noise),

4) high-frequency noise, which is the 2nd order derivative of the sensor noise e in (4.58).

Using spectral density formalism, noise PSD may be approximated in the mid BW by

$$S_a^2(f) \cong S_{a0}^2 (f_{a0} / f)^{2r_0} + S_{a1}^2 + S_{a2}^2 (f / f_{a2})^{2r_2} \quad (4.59)$$

with the asymptotic assumptions

$$\begin{aligned} \lim_{f \rightarrow 0} S_a^2(f) &= S_{a0}^2 \\ \lim_{f \rightarrow \infty} S_a^2(f) &\leq S_{a2}^2 \end{aligned} \quad (4.60)$$

that subsume f_{a0} and f_{a2} are the low- and high-frequency cutoffs, below and above which spectral density tends to become flat. Assuming sensor noise PSD is flat and equal to S_e within the accelerometer BW, and $f_{a2} \cong f_a$, it holds

$$\begin{aligned} r_2 &= 2 \\ S_{a2} &= S_e \omega_a^2 / k_y \end{aligned} \quad (4.61)$$

Assuming $S_e / k_y \leq 10^{-9} \text{ m}/\sqrt{\text{Hz}}$ (achievable by capacitive sensors) and $f_a \leq 100 \text{ Hz}$, yields $S_{a2} \leq 0.4 \text{ mm/s}^2/\sqrt{\text{Hz}}$, which shows the high accuracy of position sensor to limit accelerometer noise in a wide BW.

Figure 19 shows the spectral density of an ultra-sensitive scientific accelerometer (GOCE type), used to monitor from space Earth gravity. Note the ordinates are in $\mu\text{m/s}^2/\sqrt{\text{Hz}}$. x and y axes show the typical basin-like shape of the spectral density in agreement with (4.59). Axis z is designed to be less sensitive especially in the mid-frequency region above 1 mHz, which is the key mission frequency band. Bias ranges $0.1 \div 10 \mu\text{m/s}^2$.

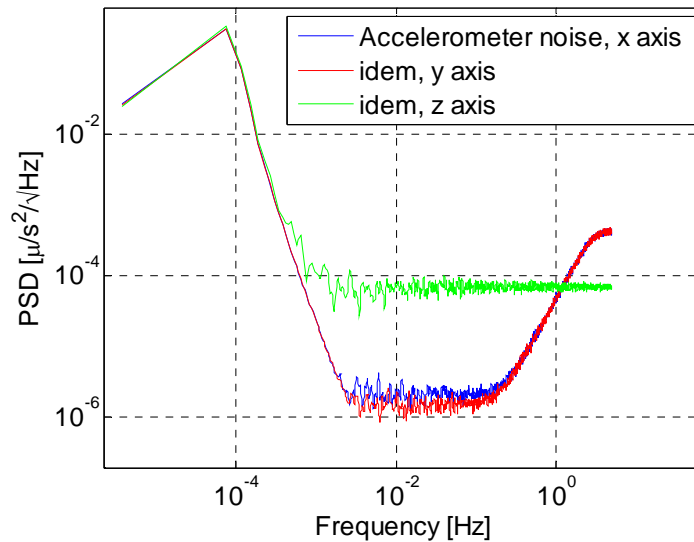


Figure 19. Spectral density of the error of a scientific 3D accelerometer.

Accelerometers mounted in standard strapdown¹ inertial measurement units (IMU) are much less sensitive, their spectral density above 0.01 mHz being flat and close to $1 \text{ mm/s}^2/\sqrt{\text{Hz}}$. Bias may range from 0.3 to 3 mm/s^2 . Below 0.01 mHz, random drift establishes.

4.4.4 Discrete-time control

The discrete-time model easily derives from (4.48) assuming sampling time T_r , command interpolation through zero-order hold and $\omega_a T \ll 1$.

$$\begin{aligned} \begin{bmatrix} \delta r \\ \delta v \end{bmatrix}(i+1) &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \delta r \\ \delta v \end{bmatrix}(i) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} a_u(i) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} d(i) \\ y(i) &= [1 \quad 0] \begin{bmatrix} \delta r \\ \delta v \end{bmatrix}(i) + e(i) = y_m(i) + e(i) \end{aligned} \quad (4.62)$$

where state variables δr , δv , command u , in voltage units, and input accelerations d_a , a have been converted to position units, say

$$\begin{aligned} \delta v(i) &= \Delta v(i)T \\ d(i) &= (d_a(i) - a(i))T^2 \\ a_u(i) &= bu(i), \quad b = \phi T^2 / m_a \end{aligned} \quad (4.63)$$

ϕ denoting the command gain [N/V]. In (4.62) distinction has been made between plant measure y and model output y_m ; the difference e is the so-called model error.

In modern synthesis [9], control objectives may be extended from simple position tracking, to velocity and acceleration ‘stabilization’, i.e. to

$$\begin{aligned} \delta r(i) &= \delta v(i) = 0 \\ \delta v(i+1) - \delta v(i) &= a_u(i) + d(i) = 0 \end{aligned} \quad (4.64)$$

The first two conditions require equation (4.62) to be stabilized which is guaranteed by a state feedback; the last condition just implies the command to repeat the negative disturbance terms, as follows

$$a_u(i) = -(k_r \delta r(i) + k_v \delta v(i) + d(i)). \quad (4.65)$$

Before affording the problem of how retrieving the three components in (4.65), an expression for the two feedback gains, the position gain k_r and the rate gain k_v must be found. It follows

¹ Original IMU consisted of gimbaled gyro-stabilized platforms carrying accelerometers. Modern IMU no gimbal exists, but accelerometers are strapped to vehicle.

by assigning asymptotically stable eigenvalues to the closed-loop state matrix $A - BK$, which in this case, becomes

$$A - BK = \begin{bmatrix} 1 & 1 \\ -k_r & 1 - k_v \end{bmatrix}. \quad (4.66)$$

Given two complementary eigenvalues $\gamma_{ck} = 1 - \lambda_{ck}$, $k = 0, 1$, such as $|\lambda_{ck}| < 1$, the gains easily follow from the coefficients of the characteristic polynomials, i.e. from

$$k_r = \gamma_{c0}\gamma_{c1}, \quad k_v = \gamma_{c0} + \gamma_{c1}, \quad (4.67)$$

which may be proved by the reader.

Exercise 1. Prove that assuming real eigenvalue and $0 \leq \lambda_{ck} \leq 1 - \varepsilon$, with $\varepsilon > 0$, it holds

$$\begin{aligned} \varepsilon^2 &\leq k_r \leq 1 \\ \varepsilon &\leq k_r / 2 \leq 1 \end{aligned} \quad (4.68)$$

4.4.4.1 State observer

The way to retrieve components in (4.65) passes through a state-observer fed by measurement y and the command a_u . Preliminary is the definition of a disturbance state x_d capable of estimating d . The simplest way as in [9] is to model d as the combination of a random drift and noise

$$\begin{aligned} d(i) &= x_d(i) + w(i) \\ x_d(i+1) &= x_d(i) + w_d(i) \end{aligned} \quad (4.69)$$

Combining (4.62) and (4.69) a 3rd order state equation, called the Embedded Model, which allows to know at any time all state variables.

$$\begin{aligned} \begin{bmatrix} \delta r \\ \delta v \\ x_d \end{bmatrix} (i+1) &= \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \delta r \\ \delta v \\ x_d \end{bmatrix} (i) + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} a_u(i) + \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} w \\ w_d \end{bmatrix} (i), \quad \begin{bmatrix} \delta r \\ \delta v \\ x_d \end{bmatrix} (0) = \begin{bmatrix} \delta r_0 \\ \delta v_0 \\ x_{d0} \end{bmatrix} \\ y(i) &= [1 \quad 0 \quad 0] \begin{bmatrix} \delta r \\ \delta v \\ x_d \end{bmatrix} (i) + e(i) = y_m(i) + e(i) \end{aligned} \quad (4.70)$$

Equation (4.70) immediately shows all the state variables are known at every time, when the initial state and the noise vector are known, assuming command to be known being internal to control unit. Contribution of the initial state goes to zero (free response) if the eigenvalues are asymptotically stable, which is not the case of (4.70). Stabilization may be obtained through

output-to-state feedback which is typical of state observers. The feedback is driven by the ‘actual’ model error defined in (4.70), i.e. by

$$\bar{e}(i) = y(i) - \hat{y}_m(i), \quad (4.71)$$

where a pair of notations have been introduced:

- 1) the ‘hat’ notation, denoting ‘one-step prediction’ and meaning that the corresponding variable, say \hat{y}_m , only depends on past measurements $y(i-k)$, $k > 0$,
- 2) the ‘bar’ notation, denoting ‘estimation’ or ‘correction’, and meaning the corresponding variable, say \bar{e} , depends on past and present measurements $y(i-k)$, $k \geq 0$.

The question is now where to apply the error $\bar{e}(i)$: one way is to pass through noise vector, thus disposing of a pair of feedback paths, as in the following closed-loop equation

$$\begin{aligned} \begin{bmatrix} \delta \hat{r} \\ \delta \hat{v} \\ \hat{x}_d \end{bmatrix} (i+1) &= \begin{bmatrix} 1 & 1 & 0 \\ -l_0 & 1 & 1 \\ -l_1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \delta \hat{r} \\ \delta \hat{v} \\ \hat{x}_d \end{bmatrix} (i) + \begin{bmatrix} 0 \\ b \\ 0 \end{bmatrix} u(i) + \begin{bmatrix} 0 \\ l_0 \\ l_1 \end{bmatrix} y(i), \quad \begin{bmatrix} \delta \hat{r} \\ \delta \hat{v} \\ \hat{x}_d \end{bmatrix} (0) = \begin{bmatrix} \delta \hat{r}_0 \\ \delta \hat{v}_0 \\ \hat{x}_{d0} \end{bmatrix} \\ y(i) &= [1 \quad 0 \quad 0] \begin{bmatrix} \delta \hat{r} \\ \delta \hat{v} \\ \hat{x}_d \end{bmatrix} (i) \end{aligned} \quad (4.72)$$

The reader is asked to prove that (4.72) cannot be stabilized, whichever be the gain values. To overcome the impasse, according to [9], a dynamic feedback must be added through a simple 1st order dynamics cascaded to model error as in

$$\begin{aligned} \begin{bmatrix} \delta \hat{r} \\ \delta \hat{v} \\ \hat{x}_d \\ q \end{bmatrix} (i+1) &= \begin{bmatrix} 1 & 1 & 0 & 0 \\ -l_0 & 1 & 1 & m_0 \\ -l_1 & 0 & 1 & m_1 \\ -1 & 0 & 0 & 1-\beta \end{bmatrix} \begin{bmatrix} \delta \hat{r} \\ \delta \hat{v} \\ \hat{x}_d \\ q \end{bmatrix} (i) + \begin{bmatrix} 0 \\ b \\ 0 \\ 0 \end{bmatrix} u(i) + \begin{bmatrix} 0 \\ l_0 \\ l_1 \\ 0 \end{bmatrix} y(i), \quad \begin{bmatrix} \delta \hat{r} \\ \delta \hat{v} \\ \hat{x}_d \\ q \end{bmatrix} (0) = \begin{bmatrix} \delta \hat{r}_0 \\ \delta \hat{v}_0 \\ \hat{x}_{d0} \\ q_0 \end{bmatrix} \\ \hat{y}_m(i) &= [1 \quad 0 \quad 0 \quad 0] \begin{bmatrix} \delta \hat{r} \\ \delta \hat{v} \\ \hat{x}_d \\ q \end{bmatrix} (i) \end{aligned} \quad (4.73)$$

Four feedback channels are now available and five gains, including the 1st order dynamics eigenvalue β ; they may be easily related to the assigned characteristic polynomial

$$P_m(\gamma) = \gamma^4 + c_{m3}\gamma^3 + c_{m2}\gamma^2 + c_{m1}\gamma + c_{m0} = \prod_{k=0}^3 (\gamma - \gamma_{mk}). \quad (4.74)$$

It is left to the reader to prove the following equalities

$$\begin{aligned}
 c_{m3} &= \sum_{k=0}^3 \gamma_{mk} = \beta \\
 c_{m2} &= \gamma_{m0} \sum_{k=1}^3 \gamma_{mk} + \gamma_{m1} \sum_{k=2}^3 \gamma_{mk} + \gamma_{m2} \gamma_{m3} = l_0 \\
 c_{m1} &= \gamma_{m0} \gamma_{m1} \sum_{k=2}^3 \gamma_{mk} + \gamma_{m2} \gamma_{m3} \sum_{k=0}^1 \gamma_{mk} = m_0 + \beta l_0 + l_1 \\
 c_{m0} &= \prod_{k=0}^3 \gamma_{mk} = m_1 + \beta l_1
 \end{aligned} \tag{4.75}$$

Equations *equation reference goes here* though nonlinear can be directly solved, through more than one solution exist due to five unknowns. Of them, m_0 and l_1 show to be alternative. The choice comes from comparing the signals they multiply:

- 1) m_0 multiplies the ‘filtered’ error q ,
- 2) l_1 multiplies the raw error \bar{e} , affected by the measurement errors.

Reducing measurement error contribution to state prediction is a fundamental directive in view of the acceleration measurement, that suggest to set $l_1 = 0$. The choice may be formally demonstrated.

4.4.4.2 Transfer functions

Harmonic analysis is a way to derive the relation between accelerometer errors and measure, especially when the spectral density is concerned. Start from the block-diagram in Figure 20, where cross-coupling $K(\cdot)$ and error dynamics $\partial P(\cdot)$ have been connected through dashed lines, since they will be neglected. The relevant Z-transforms can be obtained from (4.70) and (4.73). First the command to position holds

$$M(z) = \frac{\delta r(z)}{a_u(z)} = (z-1)^{-2}. \tag{4.76}$$

The state-predictor feedback channels have been separated into predictable disturbance and noise

$$\begin{aligned}
 H_w(z) &= \bar{w} / \bar{e} = (z-1+\beta)^{-1} m_0 + l_0 \\
 H_d(z) &= \hat{d} / \bar{e} = (z-1+\beta)^{-1} (z-1)^{-1} m_1
 \end{aligned} \tag{4.77}$$

Finally the feedback command holds

$$C(z) = \frac{a_c}{\hat{y}_m}(z) = k_r + (z-1)k_v. \tag{4.78}$$

$$\begin{aligned}\delta r &= M \left(d - C(z)(\delta r - \hat{e}) - \hat{d} - \bar{w} + \bar{w} \right) \\ (I + MC) \delta r &= (I + MC) \hat{e} + M\bar{w} \quad , \\ \delta r &= \hat{e} + S_c M \bar{w}\end{aligned}\tag{4.82}$$

upon definition of the control sensitivity S_c . Now express prediction error and estimated noise \bar{w} in terms of model error e and output disturbance $d_y = Md$, using the equality $y - \hat{y}_m = e + \hat{e}$. Prediction error holds

$$\begin{aligned}\hat{e} &= d_y - MH(y - \hat{y}_m) = d_y - MH(\hat{e} + e) \\ \hat{e} &= S_m d_y - (I - S_m)e = S_m d_y - V_m e\end{aligned}\tag{4.83}$$

upon definition of $H = H_w + H_d$, and of the predictor sensitivity $S_m = (I + HM)^{-1}$ and its complement $V_m = I - S_m$. Likewise the noise is obtained

$$\bar{w} = H_w \bar{e} = H_w(e + \hat{e}).\tag{4.84}$$

The final equation of the accelerometer tracking error follows

$$\begin{aligned}\delta r &= \hat{e} + S_c MH_w(e + \hat{e}) = (I + S_c MH_w)(S_m d_y - V_m e) + S_c MH_w e = \\ &= S d_y - V e\end{aligned}\tag{4.85}$$

upon definition of the extended sensitivity and complement

$$\begin{aligned}S &= S_m + S_c MH_w S_m \\ V &= V_m - S_c MH_w S_m \quad . \\ V &= I - S\end{aligned}\tag{4.86}$$

Remarks. One must remember that S is a high-pass transfer function and V low-pass.

Remarks. Equation (4.85) is one of the performance equations of the designed control, and says model error contributes at lower frequencies, whereas input accelerations at higher frequencies. Tuning feedback and state predictor gains should aim to reduce some norm of the tracking error below a design bound.

The instrument error is obtained as the difference between the measure \hat{d} and the ‘model value’ $d / T^2 = -a + d_a$

$$\begin{aligned}\hat{d} - d &= H_d(\hat{e} + e) - d = H_d(S_m Md - V_m e + e) - d = \\ &= (H_d S_m M - I)d + H_d S_m e = -S_d d + V_d M^{-1} e\end{aligned}\tag{4.87}$$

which shows the key property of only depending on the state predictor design. Now, separating the acceleration to be measured from parasitic terms in d , and using acceleration units, the measurement error amounts

$$n_a = S_d a + V_d (d_a + M^{-1} e / T^2). \quad (4.88)$$

4.4.4.3 Harmonic analysis

Assuming to stay below the Nyquist frequency the following approximation will be adopted

$$z - 1 = e^{j\omega T} - 1 \cong j2\pi fT \quad (4.89)$$

Compute the noise transfer functions

$$\begin{aligned} V_d &= \frac{H_d M}{I + HM} = \frac{m_1}{(z-1)^3 (z-1+\beta) + m_0 (z-1) + l_0 (z-1)(z-1+\beta) + m_1} \\ I - V_d &= \frac{(z-1)^3 (z-1+\beta) + m_0 (z-1) + l_0 (z-1)(z-1+\beta)}{(z-1)^3 (z-1+\beta) + m_0 (z-1) + l_0 (z-1)(z-1+\beta) + m_1} \end{aligned} \quad (4.90)$$

and their asymptotes

$$\begin{aligned} \lim_{z \rightarrow 1} V_d(z) &= 1, \quad \lim_{z \rightarrow \infty} V_d(z) = V_{d\infty}(z) = m_1 (z-1)^{-4} \\ \lim_{z \rightarrow 1} S_d(z) &= \frac{m_0 + l_0 \beta}{m_1} (z-1), \quad \lim_{z \rightarrow \infty} S_d(z) = 1 \end{aligned} \quad (4.91)$$

Introducing frequency and taking the amplitude

$$\begin{aligned} |V_{d\infty}(jf)| &= (f_{d\infty} / f)^4, \quad 2\pi f_{d\infty} T = m_1^{1/4} \\ |S_{d0}(jf)| &= f / f_{d0}, \quad 2\pi f_{d0} T = m_1 / (m_0 + l_0 \beta) \end{aligned} \quad (4.92)$$

which only depends predictor eigenvalues, and therefore allow to decouple accelerometer measurement from position tracking. Note the sharp attenuation at higher frequencies guaranteed by the fourth power in (4.91).

Compute now the position transfer functions upon the following ones

1) state predictor sensitivity

$$S_m = \frac{1}{I + HM} = \frac{(z-1)^3 (z-1+\beta)}{(z-1)^3 (z-1+\beta) + m_0 (z-1) + l_0 (z-1)(z-1+\beta) + m_1} \quad (4.93)$$

2) feedback sensitivity

$$S_c = \frac{(z-1)^2}{(z-1)^2 + k_v (z-1) + k_r} \quad (4.94)$$

Therefore

$$\begin{aligned}
 S &= \frac{(z-1)^3 (z-1+\beta)}{(z-1)^3 (z-1+\beta) + m_0 (z-1) + l_0 (z-1)(z-1+\beta) + m_1} \times \\
 &\times \left(1 + \frac{m_0 + l_0 (z-1+\beta)}{\left((z-1)^2 + k_v (z-1) + k_r \right) (z-1+\beta)} \right) = \\
 &= \frac{(z-1)^3}{(z-1)^3 (z-1+\beta) + m_0 (z-1) + l_0 (z-1)(z-1+\beta) + m_1} \times \\
 &\times \frac{m_0 + (z-1+\beta) \left((z-1)^2 + k_v (z-1) + k_r + l_0 \right)}{\left((z-1)^2 + k_v (z-1) + k_r \right)}
 \end{aligned} \tag{4.95}$$

and

$$\begin{aligned}
 I - S &= I - \frac{(z-1)^3}{(z-1)^3 (z-1+\beta) + m_0 (z-1) + l_0 (z-1)(z-1+\beta) + m_1} \times \\
 &\times \frac{m_0 + (z-1+\beta) \left((z-1)^2 + k_v (z-1) + k_r + l_0 \right)}{\left((z-1)^2 + k_v (z-1) + k_r \right)} = \\
 &= \frac{m_1 (z-1)^2 + (m_0 (z-1) + l_0 (z-1)(z-1+\beta) + m_1) (k_v (z-1) + k_r)}{\left((z-1)^3 (z-1+\beta) + m_0 (z-1) + l_0 (z-1)(z-1+\beta) + m_1 \right)} \times \\
 &\times \frac{1}{\left((z-1)^2 + k_v (z-1) + k_r \right)}
 \end{aligned} \tag{4.96}$$

Asymptotes

$$\begin{aligned}
 \lim_{z \rightarrow -1} V(z) &= V_\infty(z) = l_0 k_v (z-1)^{-3} \\
 \lim_{z \rightarrow -1} S(z) &= S_0 = \frac{m_0 + (l_0 + k_r) \beta}{m_1 k_r} (z-1)^3 = \frac{m_0 + l_0 \beta}{m_1} \left(\frac{1}{k_r} + \frac{1}{m_0 / \beta + l_0} \right).
 \end{aligned} \tag{4.97}$$

TBC

4.5 Angular rate sensors

4.5.1 Mechanical gyroscope

A typical configuration is the following using the same 3-1-3 frame of Chapter 2.

- 1) There is an outer gimbal rotating around the inertial axis \vec{i}_3 of the angle ψ (precession) and carrying around the first gimbal axis \vec{g}_1 , and the intermediate axis \vec{i}_{21} . The first gimbal axis becomes the rotation axis of the inner gimbal.
- 2) The inner gimbal carries a spinning body around the third gimbal axis $\vec{g}_3 = \vec{b}_3$, orthogonal to \vec{g}_1 ; the rotation angle of the inner gimbal about \vec{g}_1 is denoted by θ (nutation). The intermediate axis \vec{i}_{21} when rotated by θ is denoted with \vec{g}_2 : it is the second axis of the inner gimbal, and defines together with \vec{g}_1 the plane orthogonal to $\vec{g}_3 = \vec{b}_3$, where the body axes (\vec{b}_1, \vec{b}_2) rotate by the angle ϕ (body spin).
- 3) The body centre of mass may be shifted by a length l along the inner gimbal axis \vec{g}_3 to become a spinning top.

The gimbal to inertial transformation holds

$$R_g^i = Z(\psi) X(\theta) \quad (4.98)$$

and the body to gimbal transformation holds

$$R_b^g = Z(\phi). \quad (4.99)$$

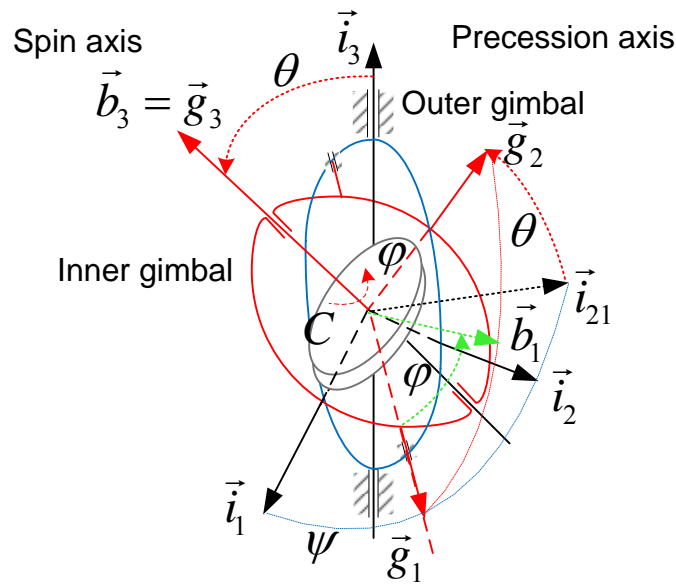


Figure 21. Gyroscope geometry.

The angular rate vector of the rotor is given by

$$\vec{\omega} = \dot{\psi} \vec{i}_3 + \dot{\theta} \vec{g}_1 + \dot{\phi} \vec{g}_3. \quad (4.100)$$

Applying the inverse of (4.98) to \vec{i}_3 , the angular rate coordinates in the gimbal frame hold

$$\boldsymbol{\omega}_g = \begin{bmatrix} \dot{\theta} \\ 0 \\ \dot{\phi} \end{bmatrix} + X(-\theta) \begin{bmatrix} 0 \\ 0 \\ \dot{\psi} \end{bmatrix} = \begin{bmatrix} \dot{\theta} \\ \dot{\psi} \sin \theta \\ \dot{\phi} + \dot{\psi} \cos \theta \end{bmatrix}. \quad (4.101)$$

TBC

4.6 Position sensors

4.6.1 The Global Navigation Satellite Systems (GNSS)

4.6.1.1 Introduction

Satellite navigation is a method to accurately determine position and time anywhere on the Earth by means of a satellite and ground network, designed and maintained as a Global Navigation Satellite System.

The NAVSTAR GPS (NAVigation System with Time and Ranging Global Positioning System) is the only fully operational GNSS system, as of 2009. The GPS is a satellite-based radio navigation, positioning and time-transfer system, initiated in 1973 and fully operational since 1995. It consists of three subsystems (or segments):

- 1) The space segment consists of 24 spacecrafts orbiting in six orbit planes $k = 0, \dots, 5$ with right ascension $\Omega_k = \Omega_0 + k\pi/3$ and equal inclination $i = 0.96 \text{ rad } (55^\circ)$; each orbit contains 4 satellites evenly spaced; the orbit is close to be circular ($e \cong 0.003$) and the semimajor axis is about $a \cong 26560 \text{ km}$, corresponding to a flight altitude $h_s = 20180 \text{ km}$ and an orbit period of a semi-sidereal day ($86164.091/2 \text{ s}$), about 2 minutes shorter than half standard solar day ($86400/2 \text{ s}$).
- 2) The control segment consists of several ground stations.
- 3) Two user segments operate: the ‘civil’ Standard Positioning Service (SPS) provides positioning and timing information worldwide, the Precise Positioning Service (PPS) provides data to authorized users.

The GPS satellites transmit microwave carrier signals at two frequencies L1 and L2. The second frequency allows measurement and correction of ionosphere signal delay. Three binary codes are modulated on the carriers.

- 1) The Coarse Acquisition (C/A) code is a short pseudorandom number (PRN) code of 1023 bits (1 ms duration) modulated on L1 carrier and specific of each satellite; it is the basis for the civil service. It is repeated equal 20 times, and corresponds to a single bit of the navigation message (Figure 22). Bit counter n is called the clock epoch or fundamental time frame (FTF) and is counted from GPS time origin, reset each week.
- 2) The precise code (P) is a PRN code long about 6×10^{12} bits, equivalent to one week. It modulates both L1 and L2 carrier phases and is the basis for the authorized service. When Anti-Spoofing (AS) mode is active, the P code is encrypted (Y code).
- 3) The navigation data are transmitted in a 50 bits/s stream, added to C/A and P(Y) codes on L1 carrier. The message consists of 25 frames, each made of 5 subframes, a sub-frame including 10 words of 30 bits each: in total 37500 bits each long 20 ms, repeated every

12.5 minutes. The message starts with the telemetry frame (TLM) of 32 bits (0.64 s) which allows to recognize the message onset. Then it contains low-accuracy (almanac data) and high-accuracy (broadcast ephemeris) satellite orbit data and clock corrections.

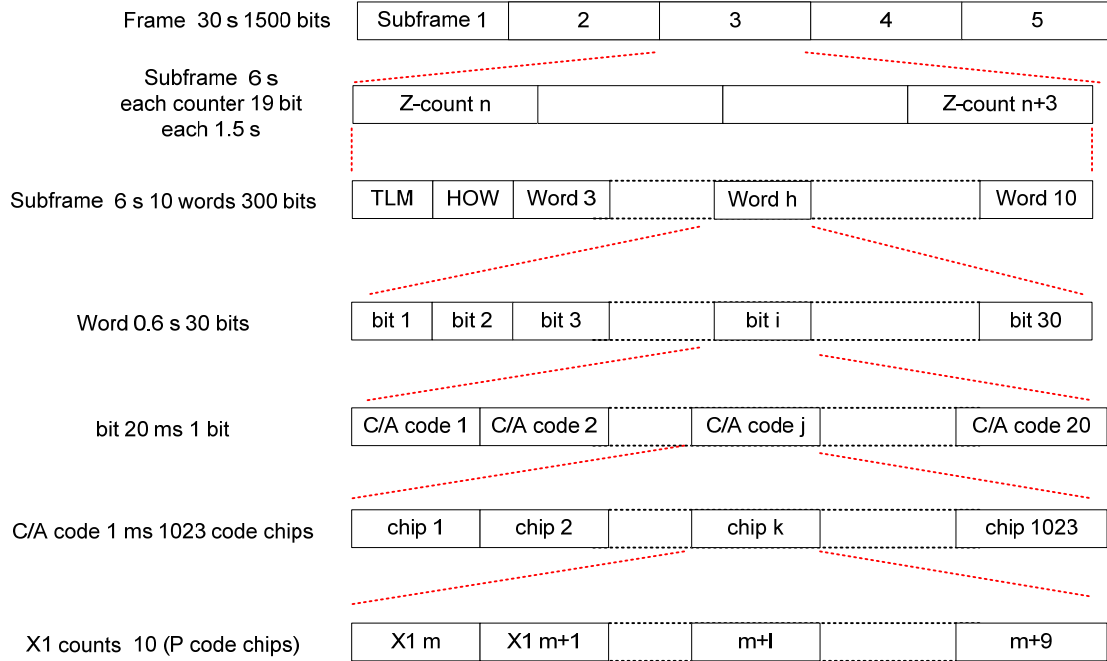


Figure 22. GPS C/A code, timing and counters.

4.6.1.2 Pseudo-range computation

GPS-based positioning relies essentially on the measurement of biased one-way range data, called pseudo-range. Without entering details of the digital electronics performing the measurement [1], the following principles are followed.

- 1) A common GPS time t is maintained by GPS control segment. The time encoded in the message is called time-of-the-week (TOW) and may be written¹ as

$$\begin{aligned}
 t_w(t) &= nT_w + \text{mod}(mT_c) \\
 T_w &= 1.5 \text{ s}, 0 \leq n < 403200 \\
 T_c &= 1 \text{ ms}, 0 \leq m < 1500
 \end{aligned}
 \tag{4.102}$$

The word HOW (handover words) contains the TOW n of the next subframe. The same expression as (4.102) applies to time t_f in a subframe with $0 \leq n < 4$. Each subframe always starts with the same code in the word TLM.

¹ Actually the X1 counts should be used instead of the C/A chip counts as in (4.102).

- 2) Satellite (local) clock time $t_{l,s}$ and receiver clock time $t_{l,r}$ are affected with respect to global times t_s and t_r by time varying offsets (errors) δt_s and δt_r , respectively, due to different sources, such that

$$\begin{aligned} t_{l,s} &= t_s + \delta t_s \\ t_{l,r} &= t_r + \delta t_r \end{aligned} \quad (4.103)$$

- 3) A C/A bit (20 ms length) transmitted at local time t_s is received at time $t_{l,r} = t_{l,s} + \Delta t$ where $\Delta t \leq c \Delta r_{\max} \cong 76$ ms, having denoted the maximum distance from a GPS satellite with Δr_{\max} .
- 4) The receiver must recover the transmission time t_s from the received code. To this end the receiver produces replica of the satellite PRN (typical of each satellite) and makes the replica to align in time with the received code. Actually alignment must be continuously kept (tracking) since the delay Δt is variable. Alignment must be achieved in a time less than $T_c = 1$ ms, which is the length of the C/A code chip. Denote with $0 \leq \Delta t_c(n) < T_c$ the time shift, simplified to be positive, detected by the receiver at TOW n . Note C/A code chip alignment does not imply full alignment, as a bit contains 20 repetitions of the C/A code. The corresponding distance error (ambiguity) is a multiple of $\delta d_{\min} = c T_c \cong 300$ km. This implies the delay Δt to hold

$$\Delta t = \Delta t_c + p T_c, \quad p \geq 0, \quad (4.104)$$

where p must be detected.

- 5) The above ambiguity is solved when transmitter TOW $t_{sw} = n_s T_w$ is known, which occurs at $t_{rw} = n T_w + q T_c + \Delta t_c$. When $n = n_s$, $p = q$ solves ambiguity. When $n = n_s + 1$, $p = 1500 + q$ solves ambiguity, which implies $q < 0$.

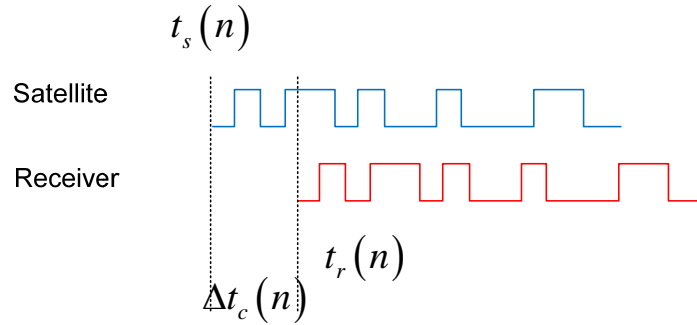


Figure 23. Delay between transmitted and received code when the delay is less than 1 ms.

Then, using (4.103), the pseudo-range ρ provided by a generic GPS satellite, is defined as

$$\Delta \rho = c \Delta t = c(t_r - t_s) + c(\delta t_r - \delta t_s) = \Delta r + c \delta t, \quad (4.105)$$

where the prefix ‘pseudo’ make distinction between the erroneous $\Delta\rho$, the measurement, and the ‘true’ range Δr , and $c = 299.8 \times 10^6$ m/s is the speed of light in vacuum.

4.6.2 Position estimation

4.6.2.1 Measurement equations

Consider a low-Earth satellite $h \cong 300$ km (the user), much less than GPS altitude $h_s = 20180$ km, whose CoM coordinates in the inertial frame are denoted by \mathbf{r} (Figure 24). Denote the CoM coordinates of a visible GPS satellite as \mathbf{s}_k , $k = 1, \dots, n_s$, where $n_s \geq 4$, and define the relative position

$$\Delta \mathbf{r}_k = \mathbf{s}_k - \mathbf{r} = \begin{bmatrix} x_k - r_x \\ y_k - r_y \\ z_k - r_z \end{bmatrix}. \quad (4.106)$$

Then equation (4.105), appending subscript k , becomes

$$\Delta \rho_k = |\mathbf{s}_k - \mathbf{r}| + c\delta t_k = \sqrt{(x_k - r_x)^2 + (y_k - r_y)^2 + (z_k - r_z)^2} + c\delta t_k. \quad (4.107)$$

A pair of changes must be done in (4.107) so as to make it viable

- 1) time error is decomposed into an unknown common error δt mainly due to user clock, a satellite clock correction δt_k , and a variable error $e_k = c(\delta t_k - \delta t)$ including all error sources,
- 2) a user inaccurate position $\hat{\mathbf{r}}$ is known either from previous measurements or from other sources; in any case a starting procedure must be defined, to be afforded below.

The second assumption allows to write

$$\delta \mathbf{r} = \mathbf{r} - \hat{\mathbf{r}} = \begin{bmatrix} \delta r_x \\ \delta r_y \\ \delta r_z \end{bmatrix}, \quad \Delta \hat{\mathbf{r}}_k = \mathbf{s}_k - \hat{\mathbf{r}} = \begin{bmatrix} \Delta \hat{x}_k \\ \Delta \hat{y}_k \\ \Delta \hat{z}_k \end{bmatrix}, \quad \Delta \hat{r}_k = |\Delta \hat{\mathbf{r}}_k|, \quad (4.108)$$

and assuming $|\delta \mathbf{r}| \ll |\mathbf{r}|$, to rewrite (4.107) in the approximate form

$$\Delta \rho_k = \sqrt{(x_k - \hat{r}_x)^2 + (y_k - \hat{r}_y)^2 + (z_k - \hat{r}_z)^2} - \frac{\Delta \hat{x}_k}{\Delta \hat{r}_k} \delta r_x - \frac{\Delta \hat{y}_k}{\Delta \hat{r}_k} \delta r_y - \frac{\Delta \hat{z}_k}{\Delta \hat{r}_k} \delta r_z + c\delta t + e_k, \quad (4.109)$$

which is left to prove by the reader. Equation (4.109) may be rewritten in compact form

$$\Delta \rho_k = \Delta \hat{\rho}_k + H_k \delta \mathbf{p} + e_k, \quad (4.110)$$

with the notations

$$\delta \mathbf{p} = \begin{bmatrix} \delta \mathbf{r} \\ c\delta t \end{bmatrix} = \begin{bmatrix} \delta r_x \\ \delta r_y \\ \delta r_z \\ c\delta t \end{bmatrix}, H_k = \begin{bmatrix} -\frac{\Delta \hat{x}_k}{\Delta \hat{r}_k} & -\frac{\Delta \hat{y}_k}{\Delta \hat{r}_k} & -\frac{\Delta \hat{z}_k}{\Delta \hat{r}_k} & 1 \end{bmatrix} = \begin{bmatrix} -\mathbf{h}_k^T & 1 \end{bmatrix}. \quad (4.111)$$

Remarks. Approximation error must be considered as part of e_k .

Remarks. The known position $\hat{\mathbf{r}}$ and the subsequent range vector $\Delta \hat{\mathbf{r}}_k = \mathbf{s}_k - \hat{\mathbf{r}}$ may be thought of the output of a prediction algorithm, to be corrected upon a new measure $\Delta \rho_k$.

Remarks. Satellite position \mathbf{s}_k is obtained from broadcast orbit parameters (ephemerides) as in Chapter 1.

By disposing of n different down-links the following measurement equations can be established

$$\Delta \mathbf{p} = \Delta \hat{\mathbf{p}} + H \delta \mathbf{p} + \mathbf{e}, \quad (4.112)$$

where notations in (4.110) and below

$$\Delta \mathbf{p} = \begin{bmatrix} \Delta \rho_1 \\ \Delta \rho_2 \\ \dots \\ \Delta \rho_n \end{bmatrix}, \Delta \hat{\mathbf{p}} = \begin{bmatrix} \Delta \hat{\rho}_1 \\ \Delta \hat{\rho}_2 \\ \dots \\ \Delta \hat{\rho}_n \end{bmatrix}, H = \begin{bmatrix} H_1 \\ H_2 \\ \dots \\ H_n \end{bmatrix}, \mathbf{e} = \begin{bmatrix} e_1 \\ e_2 \\ \dots \\ e_n \end{bmatrix} \quad (4.113)$$

have been used.

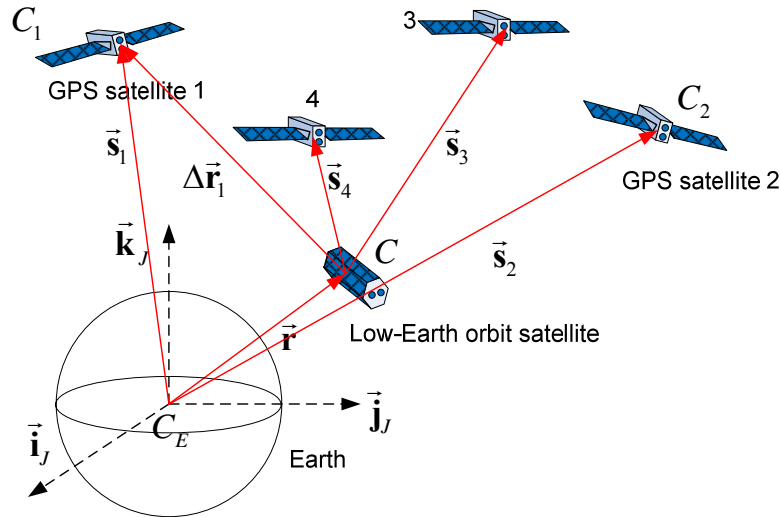


Figure 24. Geometry of satellite-to-satellite position measurement.

Error in (4.112) is usually approximated to be a zero-mean random variable with known covariance, i.e.

$$\mathcal{E}\{\mathbf{e}\} = 0, \mathcal{E}\{\mathbf{e}\mathbf{e}^T\} = I\sigma_e^2, \quad (4.114)$$

where σ_e^2 is the variance of the user-equivalent range error (UERE). TBV

Statistics in (4.114) may be justified as follows.

- 1) Zero-mean value justifies because of (4.109) where the mean clock error has been extracted as unknown. In addition \mathbf{e} may be assumed independent from past measurements, thus giving it white-noise properties.
- 2) Covariance in (4.114) is approximated to be diagonal, which is not completely true as propagation errors along different links are correlated.
- 3) Deviation from such assumptions will not be treated.

4.6.2.2 Least-square estimation and performance

Under (4.112) and (4.114) the ‘least square’ solution minimizing the square error $\mathbf{e}^T \mathbf{e}$ with respect to $\Delta \mathbf{p}$ must satisfy

$$\nabla_{\delta \mathbf{p}} \mathbf{e}^T \mathbf{e} = \nabla_{\delta \mathbf{p}} (\Delta \mathbf{p} - \Delta \hat{\mathbf{p}} - H \delta \bar{\mathbf{p}})^T (\Delta \mathbf{p} - \Delta \hat{\mathbf{p}} - H \delta \bar{\mathbf{p}}) = -2H^T (\Delta \mathbf{p} - \Delta \hat{\mathbf{p}}) + 2H^T H \delta \bar{\mathbf{p}} = 0, \quad (4.115)$$

thus leading to the solution

$$\delta \bar{\mathbf{p}} = \begin{bmatrix} \delta \bar{\mathbf{r}} \\ c \delta \bar{t} \end{bmatrix} = (H^T H)^{-1} H^T (\Delta \mathbf{p} - \Delta \hat{\mathbf{p}}) = K (\Delta \mathbf{p} - \Delta \hat{\mathbf{p}}), \quad (4.116)$$

under assumption of $H^T H$ being invertible.

Remarks. Matrix $H^T H$ is invertible provided the unit vectors \mathbf{h}_k in (4.111), $k=1, \dots, n$, form a 3D basis, their tips do not lay in a plane and $n \geq 4$.

Proof. Assume first $n=4$, which implies $H^T H$ is invertible if and only if H it is. Then partition H as

$$H = \begin{bmatrix} H_3 & \mathbf{e}_3 \\ \mathbf{h}_4 & 1 \end{bmatrix} \quad (4.117)$$

and compute the determinant $\det H = \det(H_3 - \mathbf{e}_3 \mathbf{h}_4^T)$. Now assume $\det H_3 \neq 0$, implying the unit vectors \mathbf{h}_k , $k=1, 2, 3$ are a basis. In this case $\mathbf{h}_4^T = \mathbf{q}_4^T H_3$ and

$$\det(H_3 - \mathbf{e}_3 \mathbf{h}_4^T) = \det H_3 \det(I - \mathbf{e}_3 \mathbf{q}_4^T). \quad (4.118)$$

The first determinant in (4.118) is non zero if and only H_3 is a 3D basis; the second determinant can be shown non zero if and only if $\mathbf{e}_3^T \mathbf{q}_4 \neq 1$, i.e. the tip of \mathbf{q}_4 does not lay in the plane defined by three Cartesian unit vectors, or which is the same the tip of \mathbf{h}_4 does not lay in the plane defined by the unit vectors in H_3^T . Since \mathbf{h}_k is a unit vector, they lay on a unit

sphere and provided they constitute a 3D basis, and $n \geq 4$, H is always full rank and $H^T H$ invertible.

Statistical properties of the solution derive from (4.116) and (4.114). Specifically $\delta \bar{\mathbf{p}}$ is unbiased and minimum variance

$$\begin{aligned} \mathcal{E}\{\delta \bar{\mathbf{p}}\} &= \delta \mathbf{p} \\ \mathcal{E}\{(\delta \bar{\mathbf{p}} - \delta \mathbf{p})(\delta \bar{\mathbf{p}} - \delta \mathbf{p})^T\} &= (H^T H)^{-1} \sigma_e^2 = S_{\delta \mathbf{p}}^2. \end{aligned} \quad (4.119)$$

The covariance result easily follows from (4.112) and writing the estimation error as follows

$$\delta \bar{\mathbf{p}} - \delta \mathbf{p} = (H^T H)^{-1} H^T ((\Delta \mathbf{p} - \Delta \hat{\mathbf{p}}) - H \delta \mathbf{p}) = (H^T H)^{-1} H^T \mathbf{e}, \quad (4.120)$$

and the equalities

$$\mathcal{E}\{(\delta \bar{\mathbf{p}} - \delta \mathbf{p})(\delta \bar{\mathbf{p}} - \delta \mathbf{p})^T\} = (H^T H)^{-1} H^T \mathcal{E}\{\mathbf{e} \mathbf{e}^T\} H (H^T H)^{-1} = (H^T H)^{-1} \sigma_e^2. \quad (4.121)$$

It is usually preferable to extract scalar performance indices from covariance $S_{\delta \mathbf{p}}^2$. The most useful is the covariance trace $\sigma_{\delta \mathbf{p}}^2 = \text{trace} S_{\delta \mathbf{p}}^2$ corresponding to the sum of the covariance diagonal, and therefore defined as

$$\sigma_{\delta \mathbf{p}}^2 = \text{trace} S_{\delta \mathbf{p}}^2 = \mathcal{E}\{(\delta \bar{\mathbf{p}} - \delta \mathbf{p})^T (\delta \bar{\mathbf{p}} - \delta \mathbf{p})\}, \quad (4.122)$$

which is the sum of the parameter variance and also the sum of the covariance matrix eigenvalues¹. Then a performance index only depending on link geometry is the so called ‘geometric dilution of precision’ (GDOP)

$$s_{\delta \mathbf{p}} = \frac{\sqrt{\text{trace} S_{\delta \mathbf{p}}^2}}{\sigma_e} = \sqrt{\text{trace} (H^T H)^{-1}}. \quad (4.123)$$

The ‘position dilution of precision’ (PDOP) is restricted to position variance and holds

$$s_{\delta \mathbf{r}} = \frac{\sqrt{\mathcal{E}\{(\delta \bar{\mathbf{r}} - \delta \mathbf{r})^T (\delta \bar{\mathbf{r}} - \delta \mathbf{r})\}}}{\sigma_e}, \quad (4.124)$$

where $\delta \bar{\mathbf{r}} - \delta \mathbf{r}$ is the random uncertainty (error) around the true ‘user’ position vector. The PDOP assess the 1-sigma uncertainty radius normalized by pseudo-range standard deviation.

¹ $\sqrt{\text{trace} S_{\delta \mathbf{p}}^2}$ is also referred to as Frobenius norm.

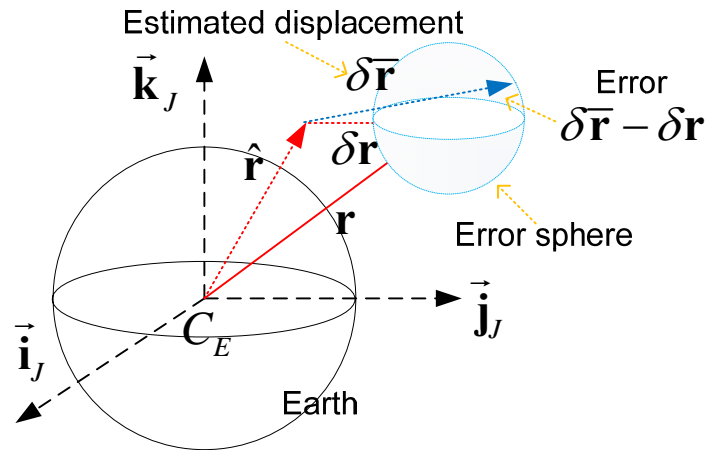


Figure 25. Predicted, true and position error.

Example 1. Consider the quadruple made by 3 Cartesian axes and their bisector, i.e.

$$H = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} & 1 \end{bmatrix}, \quad \det H = 1 - 3/\sqrt{3} = 1 - \sqrt{3}. \quad (4.125)$$

Then GDOP and PDOP hold

$$GDOP = 3.6, \quad PDOP = 3.1, \quad (4.126)$$

which are low figures.

Remarks. GDOP decreases increasing the number of visible satellites and their angular distance as seen by the user. Since H is time varying, so are GDOP and PDOP.

4.6.3 Pseudo-range errors

Pseudo-range errors entering e_k in (4.109) are of the following types

- 1) satellite clock errors entering δt_k in (4.107),
- 2) ephemeris errors affecting the satellite position \mathbf{s}_k recovered from navigation message,
- 3) atmospheric errors due to signal propagation in the ionosphere (time advance) and troposphere (time delay),
- 4) errors due to receiver noise,
- 5) multipath errors.

Table 4. GPS precise (standard) positioning typical user error
--

No.	Segment source	Error source	RMS error [m]	Comments
0	Space/control	Broadcast clock	1.1	
1		L1 C/A group delay	(0.3)	
2		Broadcast ephemeris	0.8	
3	User	Ionospheric delay	0.1 (7.0)	Highly correlated among satellites
4		Tropospheric delay	0.2	
5		Receiver noise and resolution	0.1	
6		Multipath	0.2	
7	System	Total RMS	1.4 (7.1)	

4.6.3.1 Satellite clock errors

GPS satellites contain highly stable atomic clocks, but deviation between satellite time and GPS time, maintained by the master control station (MCS), may reach 1 ms (300 km range). Clock error correction is periodically uplinked by MCS at a reference time \underline{t} and re-broadcast by each satellite. Correction is expressed as a 2nd order polynomial

$$\delta \underline{t}_k = a_{0k} + a_{1k} (t - \underline{t}) + a_{2k} (t - \underline{t})^2 + \delta \underline{t}_k, \quad (4.127)$$

where a_{0k} [s] is the clock bias, a_{1k} [s/s] is the clock drift, a_{2k} [s/s²] is the clock aging and $\delta \underline{t}_k$ accounts for relativistic effects. Residual clock error $c\delta t$ may increase between two subsequent corrections from 0.8 to 4 m. Clock errors are expected to continually decrease pushed by technology improvements.

4.6.3.2 Atmospheric effects

Propagation speed of electromagnetic waves in a medium depends on the index of refraction n defined as the ratio of the wave's propagation speed c in the free space to that in the medium v

$$n = c / v. \quad (4.128)$$

The medium is said to be dispersive if the index of refraction n , hence the wave's speed v , depends on the wave's frequency f (think of a rainbow caused by a dispersive prism). Modulated wave travelling in space may be represented as the composition of close, but different, frequency waves, in the simplest case a pair at $f_1 \neq f_2$ as follows

$$S(t, x) = S_0 \left(\sin 2\pi f_1 (t - x/v_1) + \sin 2\pi f_2 (t - x/v_2) \right), \quad (4.129)$$

where waves have been assumed to travel in space with different (phase) speeds $v_1 \neq v_2$ to account for dispersive medium and to have the same amplitude S_0 . Trigonometric identities allow to rewrite (4.129) in the modulated form

$$S(t, x) = 2S_0 \cos 2\pi \Delta f (t - x/v_m) \sin 2\pi f (t - x/v), \quad (4.130)$$

where

- 1) $\Delta f = (f_1 - f_2)/2$ is the beat or modulation frequency, whereas $f = (f_1 + f_2)/2$ is the carrier frequency,
- 2) v_m and v are the modulation (o group) and carrier phase speeds, with expressions

$$\begin{aligned} v_m &= \frac{f_1 - f_2}{f_1/v_1 - f_2/v_2} = \frac{v_1/\lambda_1 - v_2/\lambda_2}{1/\lambda_1 - 1/\lambda_2} = v_1 - \lambda_1 \frac{v_2 - v_1}{\lambda_2 - \lambda_1}, \\ v &= \frac{f_1 + f_2}{f_1/v_1 + f_2/v_2} \end{aligned} \quad (4.131)$$

where wavelengths $\lambda_k = v_k / f_k$ have been introduced.

When as in GPS transmission frequencies and speeds are very close, (4.131) may be approximated as

$$v_m \cong v \left(1 - \frac{\lambda}{v} \frac{dv}{d\lambda} \right) = v \left(1 - \frac{d \ln v}{d \ln \lambda} \right), \quad (4.132)$$

showing the group velocity (carrying the information) equals the carrier speed corrected by the fractional variation of the speed with respect to wavelength.

5 Actuators

The treatment below is mainly functional, and aims to capture what is of interest for the attitude control, i.e. from the standpoint of the ACS designer. Therefore technology and constructive aspects will be barely mentioned.

Space type actuators can be classified into:

- 1) Force actuators: mainly chemical and electrical thrusters. They are used both for CoM motion control (orbit control) both for attitude control. To this end force direction must not pass through the spacecraft CoM.
- 2) Torque actuators: mainly inertia wheels, magnetic torquers, control moment gyros. They only allow attitude control.

5.1 Propulsion systems

5.1.1 Introduction and thrust equations

Propulsion systems must deliver forces and torques acting on the spacecraft body. They divide into three main categories

- 1) cold gas,
- 2) chemical: solid and liquid,
- 3) electrical.

Chemical propulsion is said energy limited, because available energy is stored in the propellant, the reactants having a fixed energy per mass unit. However because propellants are their own energy source, the rate at which energy may be supplied to propellant is independent of the propellant mass (only limited by reaction kinetics), so very high power and thrust levels can be attained. Thus chemical propulsion offers high thrust-to-mass ratios.

Electric propulsion is instead power limited, as the rate of energy supply is bounded by the power system mass. This has the effect of limiting the maximum thrust for a given spacecraft mass, and therefore of reducing the thrust to mass ratio, and therefore the vehicle acceleration. On the other hand, an arbitrary amount of energy (from solar panels, nuclear power system) can be supplied to a propellant mass, thus achieving high exhaust speed (or specific impulse per unit mass), much higher than chemical propulsion, thus allowing long missions with small propellant mass.

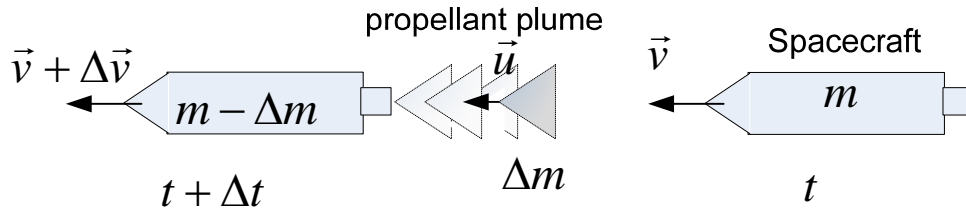


Figure 26. Propellant ejection and momentum variation.

Consider a spacecraft of mass m and velocity \mathbf{v} at time t . At time $t + \Delta t$, assume propellant of mass Δm has been ejected at velocity \mathbf{u} (Figure 26). Denote the spacecraft velocity at time $t + \Delta t$ with $\mathbf{v} + \Delta \mathbf{v}$. Newton's equation may be written by expressing the momentum variation of the propellant and spacecraft system during Δt

$$(m - \Delta m)(\mathbf{v} + \Delta \mathbf{v}) + \Delta m \mathbf{u} - m\mathbf{v} = \Delta t \sum_k \mathbf{F}_k, \quad (5.1)$$

where \mathbf{F}_k is an external force. Momentum variation has the unit of a force impulse [Ns], and if divided times the mass [kg] has the unit of speed [m/s]; velocity may thus be considered as the specific impulse per unit mass. Converting to derivative and neglecting 2nd order terms yields

$$m \frac{d\mathbf{v}}{dt} = \frac{dm}{dt}(\mathbf{v} - \mathbf{u}) + \sum_k \mathbf{F}_k. \quad (5.2)$$

By defining the exhaust velocity, or specific impulse per unit mass, as the relative velocity

$$\mathbf{v}_r = \mathbf{u} - \mathbf{v} \quad (5.3)$$

and the thrust (TBR)

$$\mathbf{F}_t = -\frac{dm}{dt} \mathbf{v}_r = -\dot{m} \mathbf{v}_r = -\sum_{k=1}^m \dot{m}_k \mathbf{v}_{rk}, \quad (5.4)$$

where the usual m-dot notation has been employed, (5.2) becomes

$$m(t) \dot{\mathbf{v}}(t) = \mathbf{F}_t(t) + \sum_k \mathbf{F}_k. \quad (5.5)$$

Remarks. Equation (5.4) is the thrust equation, showing thrust to be opposite to exhaust velocity, and the absolute value to be proportional to ejected mass flow \dot{m} and exhaust velocity amplitude $|\mathbf{v}_r|$. Propulsion technology aims to increase exhaust velocity.

Since propellant is usually stored under high pressure P_e , there will be a pressure difference $(P_e - P_0)$ with respect to ambient where it is ejected, and a further thrust $A_e(P_e - P_0)$ through the ejection section (nozzle). Usually this further thrust is accounted for in (5.4), by defining the effective exhaust velocity \mathbf{v}_{ef} such that

$$\mathbf{F}_t = -\dot{m} \mathbf{v}_r + A_e(P_e - P_0) = -\dot{m} \mathbf{v}_{ef}. \quad (5.6)$$

In airplane propulsion the atmosphere is used as a propellant, that needs to be injected. In this case (5.6) must be further corrected to

$$\mathbf{F}_t = -\dot{m}\mathbf{v}_r + \dot{m}_0\mathbf{v}_0 + A_e(P_e - P_0) \cong -\dot{m}\mathbf{v}_r + A_e(P_e - P_0) = -\dot{m}\mathbf{v}_e, \quad (5.7)$$

assuming $\dot{m}_0\mathbf{v}_0$ is small.

5.1.2 Delta-V maneuver and Tsiolkovsky equation

Assume the external forces \mathbf{F}_k in (5.5) to be negligible, the effective exhaust velocity \mathbf{v}_e to be constant and integrate (5.5) and (5.6) from $t = t_0$ to $t = t_1$, which yields the Delta-V equation

$$\int_{t_0}^{t_1} d\mathbf{v}(\tau) = -\mathbf{v}_e \int_{t_0}^{t_1} \frac{dm(\tau)}{m(\tau)} \Rightarrow \Delta\mathbf{v}(t_0, t_1) = \mathbf{v}_e \ln\left(\frac{m_0}{m_1}\right). \quad (5.8)$$

Taking the velocity vector magnitude in (5.8), Tsiolkovsky¹ equation follows

$$\begin{aligned} |\Delta\mathbf{v}(t_0, t_1)| &= |\mathbf{v}_e| \ln\left(\frac{m_0}{m_1}\right) \Rightarrow \Delta v(t_0, t_1) = v_e \ln\left(\frac{m_0}{m_1}\right), \\ m_1 &= m_0 e^{-\Delta v/v_e} \end{aligned} \quad (5.9)$$

showing spacecraft mass decrease to accomplish a Delta-V manoeuvre, increasing velocity magnitude by Δv . The propellant mass m_p ejected during the manoeuvre holds

$$m_p = m_0 - m_1 = m_0(1 - e^{-\Delta v/v_e}). \quad (5.10)$$

Clearly equation (5.10) imposes high exhaust (ejection) speed v_e to save propellant: this is the reason of the development of electric propulsion. When $\Delta v \ll v_e$, Tsiolkovsky equation simplifies to proportionality

$$\frac{m_p}{m_0} \cong \frac{\Delta v}{v_e}. \quad (5.11)$$

5.1.3 Specific impulse

Often the exhaust velocity, which is the specific impulse per unit mass, is replaced by the specific impulse per unit weight (simply, specific impulse). To this end exhaust velocity is

¹ K. Tsiolkovsky was a Russian high-school mathematics teacher, considered the father of theoretical astronautics (1957-1935).

referred to equatorial Earth gravity acceleration $g_0 = 9.807 \text{ m/s}^2$, providing the key propulsion parameter: the specific impulse I_{sp} in time units

$$I_{sp} = v_e / g_0 \text{ [s]}. \quad (5.12)$$

To understand the meaning, rewrite thrust magnitude from (5.4), make explicit the mass flow, and use (5.12)

$$g_0 \dot{m}(t) = \frac{F_t(t)}{I_{sp}}, \quad (5.13)$$

showing the propellant flow (in weight units N/s) is inversely proportional to specific impulse.

Remarks. Note specific impulse usually depends on thrust F_t .

Assuming constant I_{sp} , propellant consumption during a time interval H holds

$$c(H) = \frac{1}{I_{sp} g_0} \int_0^H F_t(\tau) d\tau. \quad (5.14)$$

Exercise 2. Compute propellant consumption given $I_{sp} = 60 \text{ s}$ (typical of cold gas propulsion), $F_t(t) = F_{t0} + F_{t1} \sin \omega t$, with $F_{t1} = 5 \text{ mN} < F_{t0} = 10 \text{ mN}$, and $H = 1 \text{ y}$.

5.1.4 The rocket nozzle equation

The primary function of a nozzle is to channel and accelerate combustion products in chemical propulsion or pressurized gas in cold gas propulsion, in such a way to maximize the exhaust velocity to supersonic velocity (Figure 27). The familiar rocket nozzle, known as convergent-divergent, or deLaval nozzle, accomplishes this feat by simple geometry, in other words by just varying the cross-sectional area!

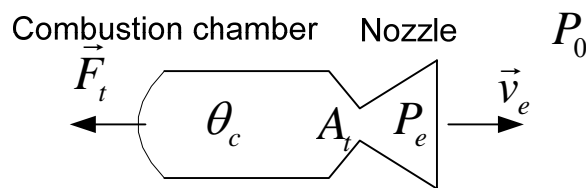


Figure 27. Convergent-divergent nozzle.

Equations leading to exhaust velocity expression simplify under the assumption of steady, one-dimensional compressible fluid flow of an ideal gas.

The final result to be proved below is the following

$$v_e = \sqrt{2 \frac{\theta_c R}{M} \frac{k}{k-1} \left(1 - \left(\frac{P_e}{P_c} \right)^{\frac{k-1}{k}} \right)}, \quad R = 8131 \text{ J/(kmol}\times\text{K)}, \quad (5.15)$$

where

- 1) θ_c is the (combustion) temperature of the propellant,
- 2) M [kg/kmol] is the molecular weight of the propellant,
- 3) $k = c_p / c_v$ is the specific heat ratio, at constant pressure and at constant volume, and is called the isentropic exponent; considering specific heat unit as [J/(kmol×K)], specific heats of an ideal gas are related by $c_p = c_v + R$; then $\gamma = (k - 1) / k$, may be expressed as

$$\gamma = \frac{k}{k-1} = \frac{\frac{c_v}{R} + 1}{\frac{c_v}{R}} = \frac{c_v}{R} + 1 > 1, \quad (5.16)$$

where $c_v / R \geq 1.5$.

- 4) P_c and P_e are the (combustion) chamber pressure and nozzle exit pressure.

Assuming $P_e / P_c \ll 1$, which is a favourable condition, (5.15), converted to specific impulse, simplifies to

$$I_{sp} \cong \frac{1}{g_0} \sqrt{2\gamma \frac{\theta_c R}{M}}. \quad (5.17)$$

Example 2. Assume a cold-gas propulsion with N₂ (Nitrogen), with $M = 28$ kg/kmol, and $c_v = 29100$ J/(kmol×K), and room chamber temperature $\theta_c \cong 300$ K. Then $\gamma \cong 4.57$ and $I_{sp} \cong 92$, which is rather optimistic.

Remarks. Equation (5.17) shows specific impulse in chemical propulsion can only be effectively increased by increasing temperature, which is obtained by combustion. Of course there are engineering bounds, which prevents temperature larger than a few thousands Kelvin, corresponding to specific impulses lower than 450 s, in case of bipropellant propulsion.

5.1.4.1 Rocket equation proof.

TBC

5.1.5 Electric propulsion

The idea of electric propulsion dates back to propulsion pioneers as R. Goddard (1906¹) and K. Tsiolkovsky (1911), still now is not routine, but must be kept under development due to incredible spread of alternative principles and technologies, and realization difficulties.

The basic principles is to create and accelerate electric charges, usually ions, to ejection speeds much higher than chemical propulsion. Whereas in chemical propulsion the main difficulty comes from achieving large thermal energy, here the main difficulty comes from disposing, in space, of large electric power.

The basic equation comes from the equality between kinetic energy of an electric charge (usually ion) of molar mass M_i [kg/kmol] and the electrostatic energy provided to the molar charge Q_i [C/kmol] by an accelerating potential difference V . It results

$$I_{sp} = \frac{v_e}{g_0} = \frac{1}{g_0} \sqrt{\frac{2Q_i V}{M_i}} \quad (5.18)$$

$$Q_i = N_A e = 0.6022 \times 10^{27} \times 0.1602 \times 10^{-18} = 96.5 \times 10^6 \text{ C/kmol}$$

Example 3. Consider for instance Xenon ion Xe^+ usually employed in ion propulsion and assume $V = 1 \text{ kV}$. Since $M_i(Xe^+) = 132 \text{ kg/kmol}$, one obtains

$$I_{sp}(Xe^+) = 3900 \text{ s}, \quad (5.19)$$

about 10 times the chemical propulsion values. As a consequence the propellant mass is much lower than other propulsion technology, thus making suitable to long-term missions. On the other hand, large electric power must be supplied to achieve high thrust values.

Electric propulsion technology differentiates how to produce ions.

5.1.6 ON-OFF and proportional thrusters

Thrusters may provide thrust either at a fixed max value $F_{t,\max}$ or at variable values $F_t \geq F_{t,\min} > 0$. In the first case, ON-OFF, thrust ideally takes only two values 0 (OFF) and $F_{t,\max}$ (ON). Actually in both cases thrust is variable due to noise and dynamics. If the ON time $t_t(t_k)$ starting at time t_k may be kept as small with respect to spacecraft time constants, then the finite time thrust profile may be integrated to and assigned to an ideal pulse located in the profile centroid $t_k + \tau_t(t_k)$, where τ_t denotes the delay between thrust firing command and pulse time

¹ R. Goddard (1882-1945), American professor of physics, was a pioneer of liquid-fueled rocketry, launching the world's first liquid-fueled rocket on March 26, 1926, among contemporary derision.

$$\underline{F}_t(t_k + \tau_t) = \int_t^{t+\gamma T_t} F_t(\tau) d\tau \delta(t - t_k - \tau_t) \quad (5.20)$$

and

$$\tau_t(t_k) = \frac{\int_{t_k}^{t_k + \gamma T_t} \tau F_t(\tau) d\tau}{\int_{t_k}^{t_k + \gamma T_t} F_t(\tau) d\tau}. \quad (5.21)$$

Under an ideal pulse of magnitude $F_{t,\max}$ and duration $T_t(t_k)$, it holds

$$\underline{F}_t(t_k + T_t(t_k)/2) = F_{t,\max} T_t(t_k) \delta(t - t_k - T_t(t_k)/2), \quad (5.22)$$

where $P_t(t_k) = F_{t,\max} T_t(t_k)$ is the commanded thrust pulse in Ns.

Remarks. Integral (5.20) should be taken up to γT_t , $\gamma > 1$, since thruster response is not ideal and decay after OFF command.

Remarks. There will be a minimum thrust ON duration $T_{t,\min}$, close to 10 ms for $1 \div 10$ N thrusters.

Therefore response under ON-OFF thrusters may be studied by considering a sequence of ideal pulses of different amplitudes, established by pulse duration T_t .

The error

$$d_t(t_k) = F_{t,\max} - \frac{1}{T_t(t_k)} \int_t^{t+\gamma T_t(t_k)} F_t(\tau) d\tau \quad (5.23)$$

may be kept as the ON-OFF thruster disturbance. Thruster quantization corresponds to ON duration quantization. Assuming $T_{t,\min} \leq T_t \leq T_{t,\max}$, where $T_{t,\max}$ is imposed by control design, and time quantization denoted with $\Delta T_t \leq T_{t,\min}$, thrust pulse bit holds

$$\mu_t = \log_2 \frac{T_{t,\max}}{\Delta T_t}. \quad (5.24)$$

Proportional thrusters allow to command variable thrust from a minimum $F_{t,\min}$ to a maximum. TBC

5.1.7 Thruster types and application

Table 5 shows the range of the deliverable force and the specific impulse.

Table 5. Force and specific impulse for different thrusters				
No	Thruster type	Force [N]	Specific impulse [s]	Comments

Table 5. Force and specific impulse for different thrusters				
No	Thruster type	Force [N]	Specific impulse [s]	Comments
Chemical				
0	Solid propellant	$50 \div 5 \times 10^6$	$200 \div 300$	
1	Liquid propellant: monopropellant	$0.05 \div 5$	$200 \div 250$	
2	Liquid propellant: bi-propellant	$5 \div 5 \times 10^6$	$300 \div 350$	
3	Cold gas (nitrogen)	$0.001 \div 200$	60	
Electrical				
4	Ion thrusters	$0.002 \div 0.5$	$2000 \div 4000$	
5	Filed-effect electrical propulsion (FEPP)	$10^{-6} \div 0.00015$	$10000 \div 15000$	

Thrusts of the order of 1 to 100 N are needed for orbital control purposes. For attitude control the required thrust ranges from 0.0001 mN to about 1 N, and depend on the environment, mass properties and control requirements. Chemical thrusters are used in ON-OFF mode: force is delivered along much shorter intervals with respect to the interval between successive firing. Electrical and cold gas are fired in a continuous mode. Driving performance for chemical propulsion is the max required thrust and the specific impulse. For continuous mode operation, bandwidth, slew-rate and repeatability of the static response (actuated force versus commanded force) are equally important.

The typical application of the various thruster types is in Table 6.

Table 6. Application of the different thrusters				
No	Thruster type	Orbit insertion	Orbital control	Attitude control
Chemical				
0	Solid propellant	X		

Table 6. Application of the different thrusters				
No	Thruster type	Orbit insertion	Orbital control	Attitude control
1	Liquid propellant: monopropellant		X	X
2	Liquid propellant: bi-propellant	X	X	X
3	Cold gas (nitrogen)		X	X
Electrical				
4	Ion thrusters		X	X
5	Filed-effect electrical propulsion (FEEP)			X

Orbital insertion as for instance into geostationary orbit is made through solid propellant thrusters, like the apogee motors. The fine orbit and trajectory control can be made through ion and cold gas thrusters. The attitude thruster assembly is usually referred to as reaction control system (RCS), to be served to the ACS.

5.1.8 Thruster geometry

Thrusters deliver thrust along a unit vector whose body coordinates are denoted by \vec{v}_t , passing through an application point denoted with \vec{a}_t in body coordinates.

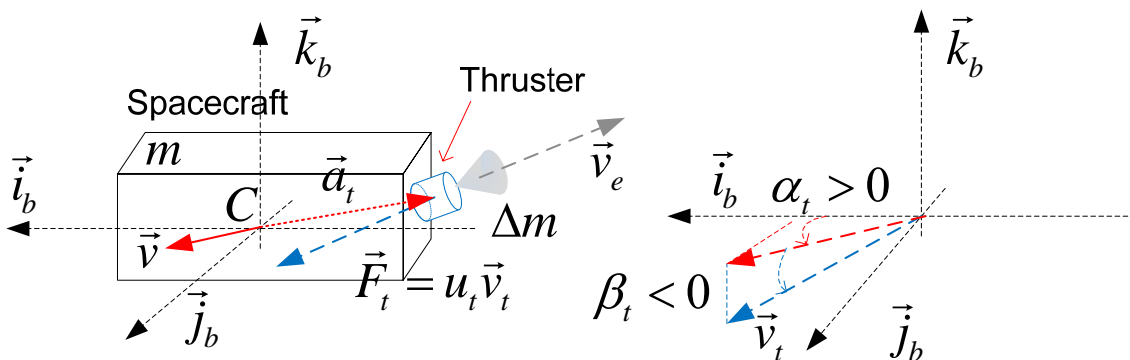


Figure 28. Elements of thruster geometry.

Now consider a thruster assembly $j = 0, \dots, m-1$ and separate the thrust $0 < u_{t,\min} \leq u_j \leq u_{t,\max}$ delivered by single thrusters from forces and torques applied to spacecraft. It holds for the spacecraft command force

$$\mathbf{F}_t = \begin{bmatrix} \mathbf{v}_{t,0} & \mathbf{v}_{t,1} & \cdots & \mathbf{v}_{t,m-1} \end{bmatrix} \mathbf{u}_t = V_t \mathbf{u}_t$$

$$\mathbf{u}_t = \begin{bmatrix} u_{t0} \\ u_{t1} \\ \vdots \\ u_{t,m-1} \end{bmatrix}, \quad (5.25)$$

and for the command torque

$$\mathbf{M}_t = \begin{bmatrix} \mathbf{a}_{t,0} \times \mathbf{v}_{t,0} & \mathbf{a}_{t,1} \times \mathbf{v}_{t,1} & \cdots & \mathbf{a}_{t,m-1} \times \mathbf{v}_{t,m-1} \end{bmatrix} \mathbf{u}_t = C_t \mathbf{u}_t. \quad (5.26)$$

The thrust unit vector may be rewritten using azimuth, $Z(\alpha_{ik})$, and elevation $Y(-\beta_{ik})$ angles as follows

$$\mathbf{v}_{ij} = \begin{bmatrix} \cos \alpha_{ij} \cos \beta_{ij} \\ \sin \alpha_{ij} \cos \beta_{ij} \\ \sin \beta_{ij} \end{bmatrix}. \quad (5.27)$$

It is immediate to establish, at least formally, from (5.25) and (5.26) the constrained dispatching laws

$$\mathbf{u}_t = (V_t)^{-1} \mathbf{F}_t, \quad Iu_{t,\min} \leq \mathbf{u}_t \leq Iu_{t,\max}$$

$$\mathbf{u}_t = (C_t)^{-1} \mathbf{M}_t, \quad Iu_{t,\min} \leq \mathbf{u}_t \leq Iu_{t,\max}, \quad (5.28)$$

where the inverse in (5.28), through existing, may depart from the square-matrix inverse, as V_t and A_t may be non square.

When the same thrusters are used to provide force and torques, equation (5.25) and (5.26) unifies into

$$\begin{bmatrix} \mathbf{F}_t \\ \mathbf{M}_t \end{bmatrix} = \begin{bmatrix} V_t \\ C_t \end{bmatrix} \mathbf{u}_t = B_t \mathbf{u}_t. \quad (5.29)$$

When the thruster size m is larger than force/torque size $n = 3 \div 6$, then the inverse may be obtained by solving an optimization problem aiming at minimizing some thrust performance, like thrust peak $\max_j u_{tj}$ or total thrust $u_{t0} + u_{t1} + \dots + u_{t,m-1}$, the latter being related to instantaneous propellant consumption.

5.1.8.1 Drag-free propulsion assembly

The goal of the thruster assembly is to contrast non-gravitational forces, which are dominate at low-Earth orbits by aerodynamic forces \mathbf{F}_d . To this end, rewrite spacecraft CoM equation from Chapter 1,

$$\begin{aligned}\dot{\mathbf{r}}(t) &= \mathbf{v}(t), \quad \mathbf{r}(0) = \mathbf{r}_0 \\ \dot{\mathbf{v}}(t) &= \mathbf{g}(\mathbf{r}) + \frac{R_b(\mathbf{q})\mathbf{F}_b(t)}{m_1}, \quad \mathbf{v}(0) = \mathbf{v}_0,\end{aligned}\quad (5.30)$$

where gravitational acceleration is denoted with $\mathbf{g}(\mathbf{r})$, and the external force \mathbf{F}_b is given in body coordinates, to be converted into inertial by the body-to-inertial transformation $R_b(\mathbf{q})$ depending on the body quaternion \mathbf{q} . To be simple, the external force vector is the sum of a disturbance force (mainly aerodynamics, ‘drag’) $-\mathbf{F}_d$, defined in Chapter 1, and a command force \mathbf{F}_t (due to thruster assembly) defined in (5.25). It leads to

$$\dot{\mathbf{v}}(t) = \mathbf{g}(\mathbf{r}) + \frac{R_b(\mathbf{q})(-\mathbf{F}_d(t) + V_t \mathbf{u}_t)}{m_1}, \quad \mathbf{v}(0) = \mathbf{v}_0. \quad (5.31)$$

Drag-free control aims to leave the spacecraft CoM free-falling, only subject to gravitational acceleration, which ideally implies

$$V_t \mathbf{u}_t = \mathbf{F}_d = \begin{bmatrix} F_{dx} \\ F_{dy} \\ F_{dz} \end{bmatrix}, \quad (5.32)$$

but subject to constraints (5.28). Assuming body axes are ideally aligned with the Local Orbit Frame (LORF), x-body axis coincides with the velocity unit vector $\vec{i}_b \cong \vec{v}/v$, which implies the following drag components constraints

$$F_{dx} > 0, \quad F_{dx} \gg \max(|F_{dy}|, |F_{dz}|). \quad (5.33)$$

A natural geometry is therefore to dispose of a single thruster aligned with \vec{i}_b to contrast the along-track drag $-F_{dx}\vec{i}_b$, and two pairs of opposite thrusters aligned with \vec{j}_b and \vec{k}_b to contrast cross-track and radial components, as they do not possess a definite sign. Direction matrix holds

$$V_t = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}. \quad (5.34)$$

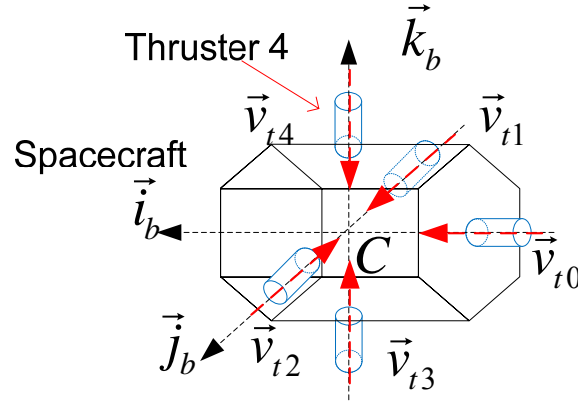


Figure 29. Propulsion assembly.

Application matrix A_t may be ideally assumed to be zero, to avoid parasitic torques: thrust vectors must pass through the CoM. Then rewriting decomposing drag into non negative components

$$\mathbf{F}_d = \begin{bmatrix} F_{dx} \\ F_{dy+} - F_{dy-} \\ F_{dz+} - F_{dz-} \end{bmatrix}, F_{dx} > 0, F_{dy+} \geq 0, F_{dy-} \geq 0, F_{dz+} \geq 0, F_{dz-} \geq 0, \quad (5.35)$$

the dispatching law, inverting (5.32) and respecting (5.28) holds

$$\begin{bmatrix} u_{t0} \\ u_{t1} \\ u_{t2} \\ u_{t3} \\ u_{t4} \end{bmatrix} = \begin{bmatrix} u_{t,\min} + \min(\max(F_{dx} - u_{t,\min}, 0), u_{t,\max} - u_{t,\min}) \\ \min(u_{t,\min} + F_{dy+}, u_{t,\max}) \\ \min(u_{t,\min} + F_{dy-}, u_{t,\max}) \\ \min(u_{t,\min} + F_{dz+}, u_{t,\max}) \\ \min(u_{t,\min} + F_{dz-}, u_{t,\max}) \end{bmatrix}. \quad (5.36)$$

The reader should verify the solution. Note thrust bias on u_{tj} , $j > 0$, which allows respecting minimum thrust, without affecting lateral force components. Unfortunately bias will create extra propellant consumption, which may not be acceptable over long times.

Assuming constant I_{sp} consumption holds from (5.14)

$$c(H) = \frac{H}{I_{sp} g_0} (\underline{F}_{dx} + 4u_{t,\min} + \underline{F}_{dy+} + \underline{F}_{dy-} + \underline{F}_{dz+} + \underline{F}_{dz-}), \quad (5.37)$$

$$\underline{F}_{dx} = \frac{1}{H} \int_0^H F_{dx}(t) d\tau, \dots$$

where underline means time average.

Note a single thruster has been assumed along-track, since $\min_t F_{dx}(t) > 0$. Of course, $u_{t,\min}$ must be compatible with it.

A further remark is that equal sized thrusters as implicit in (5.28) are oversized. Lateral thrusters $j=1, \dots, m-1$ can be sized to lateral drag components. Of course when thrust limits are touched, drag-free requirements are lost! Do other geometries exist allowing propellant saving? Since lateral (y, z) components are much less than along-track, they could be accommodated by lateral components of slanted thrusters, mounted on the back side. A layout of this kind may be appealing since avoids thrusters on lateral faces, usually devoted to solar panels. Next geometry is in this direction.

5.1.8.2 Drag-free pyramidal assembly

Consider four nominally equal-size thrusters mounted on the back side of a spacecraft in a pyramidal layout, such that

$$V_t = \begin{bmatrix} \cos \beta & \cos \alpha & \cos \beta & \cos \alpha \\ 0 & \sin \alpha & 0 & -\sin \alpha \\ \sin \beta & 0 & -\sin \beta & 0 \end{bmatrix}, \quad (5.38)$$

with $0 < \alpha < \pi/2$, $0 < \beta < \pi/2$.

missing

Figure 30. thruster pyramidal assembly.

Note V_t is not square and there are more thrusters than force components as in the geometry above. To solve (5.32), group thrusters into an arbitrary triple $\mathbf{u}_{t3}^T = [u_{tj} \ u_{tk} \ u_{th}]$ and a single thruster u_t , and denote the corresponding directions as V_{t3} and \mathbf{v}_t . Since V_{t3} , whichever be the triple, is always invertible, one may solve as follows

$$\mathbf{u}_{t3} = V_{t3}^{-1} (\mathbf{F}_t - \mathbf{v}_t u_t), \quad (5.39)$$

and u_t is a free thrust which allows to observe constraints.

E.g. assume $\mathbf{u}_{t3}^T = [u_{t0} \ u_{t1} \ u_{t2}]$, $u_t = u_{t3}$ and $\alpha = \beta$. It holds

$$V_{t3}^{-1} = \frac{1}{2 \sin \alpha} \begin{bmatrix} \tan \alpha & -1 & 1 \\ 0 & 2 & 0 \\ \tan \alpha & -1 & -1 \end{bmatrix} \quad (5.40)$$

and

$$\mathbf{u}_{t3} = \frac{1}{2 \sin \alpha} \left(\begin{bmatrix} \tan \alpha \\ 0 \\ \tan \alpha \end{bmatrix} F_{dx} + \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix} F_{dy} + \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} F_{dz} \right) + \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix} u_{t3}. \quad (5.41)$$

Now u_{t3} may only recover a single constraint, and specifically the thrust component u_{t1} in (5.41) that may become negative because of F_{dy} . Specifically

$$u_{t3} = u_{t,\min} + \max(0, -F_{dy} / \sin \alpha). \quad (5.42)$$

Equation (5.42) only guarantees u_{t1} and u_{t3} to respect minimum thrust. The other six constraints must be guaranteed by design, and through command saturation (to be possibly avoided). Note u_{t0} and u_{t2} in (5.41) to have the same expression less F_{dz} sign, which allows to treat them in the same way. Considering the maximum thrust

$$u_{t,\max} \geq \frac{1}{2 \sin \alpha} (\tan \alpha F_{dx,\max} + F_{dy-, \max} + F_{dz,\max}) - u_{t,\min}, \quad (5.43)$$

$$F_{dz,\max} = \max(F_{dz-, \max}, F_{dz+, \max})$$

inequality allows either to design cant angle α or the maximum thrust. E.g. design the cant angle as

$$\tan \alpha \cong \frac{F_{dy-, \max} + F_{dz,\max}}{F_{dx,\max}}, \quad (5.44)$$

which immediately provides

$$u_{t,\max} + u_{t,\min} \geq F_{dx,\max} / \cos \alpha. \quad (5.45)$$

A similar inequality holds for the minimum thrust

$$u_{t,\min} \leq \frac{1}{2 \sin \alpha} (\tan \alpha F_{dx,\min} - F_{dy+, \max} - 2F_{dy-, \max} - F_{dz,\max}) - u_{t,\min}, \quad (5.46)$$

which implies

$$\tan \alpha \geq \frac{F_{dy+, \max} + 2F_{dy-, \max} + F_{dz,\max} - 4 \sin \alpha u_{t,\min}}{F_{dx,\min}}, \quad (5.47)$$

fixing a lower bound to cant angle. Note the this bound may lead to unfeasible layout as soon as it approaches $\tan \alpha = 1$, as cant angle must allow thrust direction to pass through CoM. Therefore denoting the radius of the application points on the back side with r_a and the axial coordinate of the back side with $-x_a$, the layout constraint exists

$$\tan \alpha \leq r_a / x_a. \quad (5.48)$$

Assuming for instance $r_a = 0.5$, $x_a = 2$, $\tan \alpha \leq 0.25$ corresponding to $\alpha \leq 0.245$ rad ($\cong 15^\circ$). As a conclusion, pyramidal solution is effective only when lateral aerodynamic forces are much lower than along-track force!

Remarks. This is not the case in rocket or descent module propulsion, as the along track force is dominated by gravity to be compensated, not as in drag-free control.

Coming back to propellant consumption, equations (5.41) and (5.42) lead to

$$c(H) = \frac{H}{I_{sp} g_0} \frac{F_{dx}}{\cos \alpha}. \quad (5.49)$$

Propellant saving wrt to (5.37) is the following

$$\Delta c(H) = \frac{H}{I_{sp} g_0} \left(4u_{t,\min} + F_{dy+} + F_{dy-} + F_{dz+} + F_{dz-} - F_{dx} \left(\frac{1}{\cos \alpha} - 1 \right) \right), \quad (5.50)$$

which again suggests to minimize α in agreement with (5.48), and allows to select the layout at least from propellant standpoint.

Remarks. A smaller number of thrusters, say 3, lack degrees-of-freedom so as to respect constraints, which should be guaranteed by design.

5.1.9 Thruster dynamics and noise

Thruster dynamics determines the time response of the thrust to (voltage) command. The latter must be considered both for ON-OFF and proportional thrusters a step-wise signal, which in the proportional case has a fixed time unit T . Rarely manufacturers provide step-response profiles, which is surrogated by the response time τ_s , defined as the time interval from the differential step command $\Delta V_t(t_k) = V_t(t_k) - V_t(t_{k-1})$ application instant $t = t_k$ to the achievement of the 90% of the steady differential thrust $\Delta u_t(t_k)$. Note the response time may depend on magnitude and sign of $\Delta V_t(t_k)$. Conversion from τ_s to transfer function or state equation requires other data. In case of small steps $|\Delta u_t(t_k)| < u_{t,\max}$, a linear time-invariant dynamics usually applies, 1st or 2nd order, plus a delay $\tau_t < \tau_s$.

Assuming 2nd order dynamics it holds

$$\Delta u_t(s) = e^{\tau_t s} \frac{\omega_t^2}{s^2 + 2\zeta_t \omega_t s + \omega_t^2} K_t \Delta V_t(s), \quad (5.51)$$

where ΔV_t is the voltage command step, ω_t is the angular frequency, ζ_t is the damping coefficient, and K_t is the thruster scale factor [N/V]. The 2nd order step response holds

$$\Delta u_t(t) = K \Delta V_t \left(1 - \cos \left(\omega_t \sqrt{1 - \zeta_t^2} (t - \tau_t) - \sin^{-1}(\zeta_t) \right) e^{-\zeta_t \omega_t (t - \tau_t)} \right), \quad t \geq \tau_t. \quad (5.52)$$

The first transient response holds

$$\Delta u_t(t) = K \Delta V_t \frac{\omega_t^2}{2} (t - \tau_t)^2, t > \tau_t \quad (5.53)$$

which implies

$$\frac{\sqrt{1.8}}{\omega_t} < \tau_s - \tau_t < \frac{\pi}{\omega_t \sqrt{1 - \zeta_t^2}}. \quad (5.54)$$

TBC

Thruster noise may include

- 1) variable bias due to scale factor error,
- 2) high-frequency noise, usually with flat PSD $S_{th}(f)$
- 3) low-frequency noise, with a PSD higher than high-frequency, and connected by a 1st order drift as follows

$$S_{tl}^2(f) \cong \frac{S_{t0}^2}{1 + (f / f_{t0})^2}, \quad (5.55)$$

- 4) thrust quantization: given the number of bits μ_t and $u_{t,\max}$, quantization level (QL) holds

$$\rho_t = u_{t,\max} / 2^{\mu_t} \quad (5.56)$$

and should not be confused with $u_{t,\min}$. Then assuming a command time unit T , the quantization noise spectral density (root) is assumed to be flat and equal to

$$S_{t\rho}(f) = \frac{\rho_t}{\sqrt{12}} \sqrt{T}. \quad (5.57)$$

5.2 Torque actuators

5.2.1 Inertia wheels

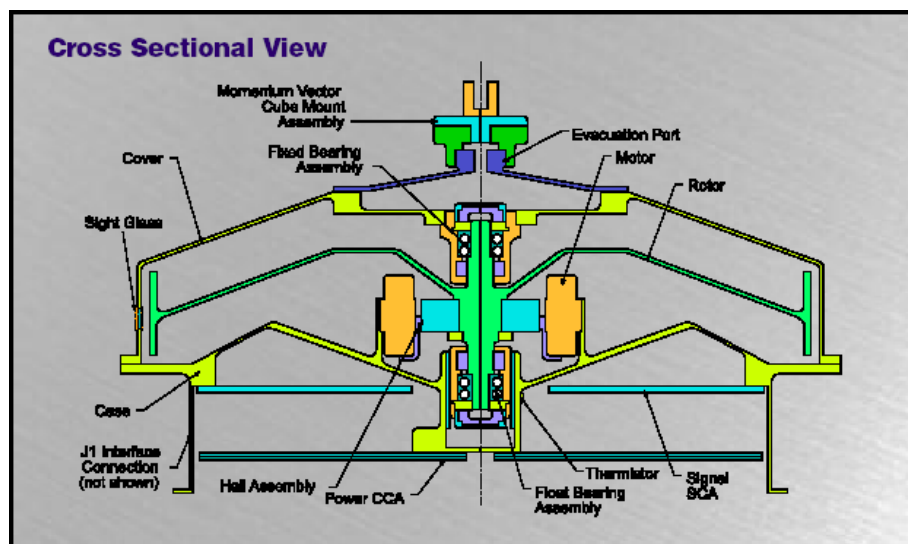
They are made by a wheel or disk made rotating by a motor. They receive from and transfer to the space vehicle the angular momentum.

Torque application to the rotor of the wheel creates an equal and opposite to the space vehicle where the motor stator is rigidly connected. The spacecraft angular momentum is reduced or increased at the expense of the wheel angular momentum which is either increased or reduced. The presence of internal friction torques (viscous and Coulombian friction on the

bearings) and of environment secular¹ torques, makes the magnitude of the angular momentum of the wheel to increase beyond the manufacturer's limits while keeping the spacecraft momentum bounded. The wheel speed cannot increase beyond a well defined bound, less earlier wearing of the bearings and mechanical vibrations because of the wheel unbalance. It is therefore mandatory to periodically unload the wheel, bringing the angular momentum to the nominal value. The operation must be necessarily made with the help of external command torques like those provided by thrusters or by magnetic torquers coupling with the Earth magnetic field. In this way the spacecraft angular momentum is kept unaltered.

Depending on the velocity vector of the wheel, whether it is variable in magnitude and direction, different types exist:

- 1) Momentum wheel: they work with variable speed around a mean non null value, whereas direction remains fixed in the body frame. They provide spin stability.
- 2) Reaction wheel (Figure 31, Table 7): they work with a variable speed around the zero value, direction remains fixed. They are used to compensate periodic torques, to drive slew maneuvers with a limited-time accumulation of angular momentum. Their assembly is typically made by 4 wheels: 3 of them are normally used, the 4th one allow tolerating the first failure. The layout is pyramidal.
- 3) Control moment gyro (CMG): it is a wheel rotating about an axis that can be oriented along different directions in the body frame. The axis can be re-oriented with one or two degrees-of-freedom. They are employed on the International Space Station.



¹ Secular component refers to a signal increasing or decreasing (not necessarily in a smooth way) without bounds.

Figure 31. Cross-sectional view of a reaction wheel.

They are made rotating by electric motors: AC motors are used to drive momentum wheels, brushless Dc motors are used in reaction wheels and CMG. They may be commanded to track a reference torque or speed. The wheel speed is measured by an incremental encoder.

Reaction wheels are typically used as precision actuators. Before the maturity of the electric thrusters and their routinely use on-board, reaction wheels were the torque actuator capable of providing the highest-resolution torques. Their limitation in favour of fine propulsion is due to the part movement, subject to friction, static and dynamic unbalance. Wear of bearings because of unbalance and speed progressively degrades performance and reduces lifetime.

Table 7. Typical reaction wheel parameters					
No	Parameter	Symbol	Unit	Value	Comments
0	Angular momentum at max speed		Nms	4 to 12	
1	Output torque at max speed		Nm	0.055	
2	Peak power at max speed		W	<80	
3	Power holding at max speed		W	<15	
4	Power bus voltage		V	14 to 23	
5	Wheel speed		rad/s	628	
6	Mass		kg	3.6 to 5.0	
7	Outer diameter		m	0.27	
8	Height		m	0.12	
9	Integrated electronics				Yes
10	Life		years	>10	
11	Radiation hard		krad		
12	Part screening				

Table 7. Typical reaction wheel parameters					
No	Parameter	Symbol	Unit	Value	Comments
13	Bearing				
14	Temperature range		°C	-15 to 60	
15	Vibration				
16	Motor type			19.8	DC
17	Interface				Analogue
18	Static unbalance		°	<0.2	
19	Dynamic unbalance		°	<3.1	

5.2.2 Magnetic torquers (MTR)

In Chapter 2, interaction between the residual moment of magnetic dipole of a spacecraft with the Earth magnetic field was studied and showed to create perturbing torques. In the same way, it is possible to implement an ad-hoc dipole capable of delivering the requested torques.

The dipole consists of a ferromagnetic nucleus and a coil around it, where current flows. Given the flowing current I , the dipole depends on the magnetic permeability μ of the nucleus, the number N of windings, the area A of the coil section. Neglecting losses in the magnetic circuit, the resulting dipole holds

$$\mathbf{m} = \mu N A I \mathbf{e}_m \quad (5.58)$$

where \mathbf{e}_m is the direction orthogonal to the winding cross section. The measuring unit is Am^2 . The equivalent electric circuit contains a series of resistance and inductance. Depending on the application, the flowing current can be regulated or not through a suitable current loop. No regulation is provided when the required torque needs not be accurate, which is the case when magnetic torquers only have the task of unloading inertia wheels (auxiliary actuators).

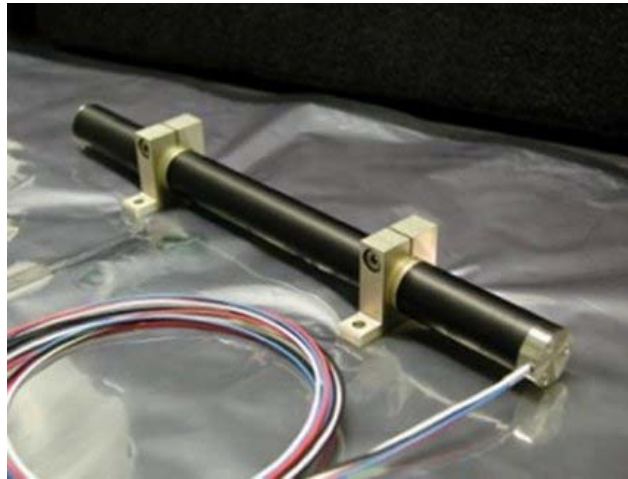


Figure 32. A typical magnetic torquer.

The command torque holds

$$\mathbf{M}_m = \mathbf{m} \times \mathbf{b}_E . \quad (5.59)$$

The torque can only be orthogonal to the Earth magnetic field. In this way magnetic torquers do not allow 3-axis controllability of the spacecraft attitude. The latter is a key limitation, that restricts magnetic control to polar orbit satellites, where the continuous rotation of the magnetic field direction allows an average 3 axis controllability during the orbit: roll and pitch close to the poles, yaw and pitch close to the Equator. Anyway magnetic control is limited to low Earth orbits, with a height not greater than 1500 to 2000 km, since above the magnetic field intensity is too low.

6 References

- [1] E. D. Kaplan and C. J. Hegarty eds *Understanding GPS. Principles and applications*. Artech House, Boston, 2006.
- [2] G. F. Franklin, J. D. Powell and M. L. Workman *Digital control of dynamic systems*, Addison-Wesley, Reading (MA), 1990.
- [3] D. T. Greenwood *Principles of dynamics*, Prentice-Hall, Englewood Cliffs, 1965.
- [4] J. R. Wertz *Spacecraft attitude determination and control*, D. Reidel Pu. Co. , Dordrecht, 1978
- [5] B. Wie *Space vehicle dynamics and control*, AIAA Education Series, AIAA inc, Reston, 1988.
- [6] M. J. Sidi *Spacecraft dynamics and control. A practical engineering approach*. Cambridge Univ. Press, 1997.
- [7] P. C. Hughes, Peter C., *Spacecraft Attitude Dynamics*. New York, Dover Publications, Inc, 2004.
- [8] M. H. Kaplan, *Modern Spacecraft Dynamics & Control*. New York : John Wiley & Sons, 1976.
- [9] E. Canuto, *Controlli automatici. Parte II. Controllo digitale*, Celid, Torino, 2002.
- [10] F. Boldrini, E. Monnini and D. Procopio “Applications of APS detector to GNC sensors”, in *Proc. 5th ESA International Conference on Spacecraft Guidance, Navigation and Control Systems*, 22-25 October 2002, Frascati, Italy. K. Fletcher and R.A. Harris, eds. ESA SP-516.
- [11] F.L. Markley and D. Mortari “Quaternion attitude estimation using vector observations”, *Journal of the Astronautical Sciences*, Vol. 48, No. 2/3, 2000, pp. 359–380.
- [12] Å. Forslund “Designing a miniaturized fluxgate magnetometer”, MS thesis in Physical Electro-technology, Div. of space and Plasma Physics, School of Electrical Engineering, Royal Inst. of Technology, Stockholm, April 2006, Doc. XR-EE-ALP-2006:002.

Content

1	Sensors and actuators for space applications	1
1.1	Objectives	1
1.2	Sensors.....	1
1.3	Attitude sensors	2
1.3.1	Sun sensors	2
1.3.2	Horizon sensors	8
1.3.3	Magnetometers	12
1.3.4	Star trackers.....	17
1.4	Inertial sensors - accelerometers.....	24
1.4.1	The proof-mass concept	24
1.4.2	Continuous-time control.....	27
1.4.3	Typical accelerometer errors	28
1.4.4	Discrete-time control.....	30
1.5	Angular rate sensors	37
1.5.1	Mechanical gyroscope.....	37
1.6	Position sensors	40
1.6.1	The Global Navigation Satellite Systems (GNSS).....	40
1.6.2	Position estimation	43
1.6.3	Pseudo-range errors.....	47
2	Actuators	50
2.1	Propulsion systems	50
2.1.1	Introduction and thrust equations.....	50
2.1.2	Delta-V maneuver and Tsiolkovsky equation.....	52
2.1.3	Specific impulse	52
2.1.4	The rocket nozzle equation	53

2.1.5	Electric propulsion	55
2.1.6	ON-OFF and proportional thrusters	55
2.1.7	Thruster types and application	56
2.1.8	Thruster geometry	58
2.1.9	Thruster dynamics and noise.....	64
2.2	Torque actuators	65
2.2.1	Inertia wheels	65
2.2.2	Magnetic torquers (MTR)	68
3	References	70
Content	71